

UNIVERSIDAD HISPANOAMERICANA
FACULTAD DE INGENIERÍA
ESCUELA INGENIERÍA INFORMÁTICA

Propuesta de automatización de procesos mediante una base de conocimiento integrada con inteligencia artificial y analítica de datos en Power BI para la plataforma VMware vCenter en Artificial Skills (ASK)

Sustentante:

José Hernán Delgado Arce

Julio 2024

TABLA DE CONTENIDO

TABLA DE CONTENIDO	2
TABLA DE FIGURAS	7
INDICE DE TABLAS.....	10
DEDICATORIA	12
CARTA DE ACEPTACIÓN DEL PATROCINADOR DEL PROYECTO	14
.....	15
DECLARACION JURADA.....	16
CARTA DEL TUTOR.....	18
CARTA DEL LECTOR	20
.....	21
LICENCIA DE AUTORIZACIÓN	22
.....	23
CAPÍTULO I: PROBLEMA DEL PROYECTO	24
1.1 ANTECEDENTES Y JUSTIFICACION DEL PROYECTO.	25
1.1.1 Marco de referencia empresarial y contextual.	26
1.1.2 Justificación del proyecto	27
1.2 DEFINICION DEL PROBLEMA.	28
1.2.1 Problemática	28
1.2.2 Problema General	30

1.2.3	Problemas específicos	31
1.3	OBJETIVOS DEL PROYECTO	31
1.3.1	Objetivo general.....	31
1.3.2	Objetivos específicos.	32
1.4	ALCANCES Y LIMITACIONES	32
1.4.1	Alcances	32
1.4.2	Limitaciones	34
1.5	CRONOGRAMA DE ACTIVIDADES	34
CAPÍTULO II: MARCO TEÓRICO		36
2.1	FUNDAMENTOS TEÓRICOS Y CONCEPTUALES.	37
2.1.1	Scrum.	37
2.1.2	Análisis de requerimientos.....	38
2.1.3	Requerimientos funcionales.	38
2.1.4	Requerimientos no funcionales.	39
2.1.5	VMware vCenter y la gestión de las máquinas virtuales.	39
2.1.6	Interfaz de programación de aplicaciones (API, por sus siglas en inglés)...	42
2.1.7	Automatización de procesos.....	46
2.1.8	Bases de conocimiento como sistema de soporte a toma de decisiones.	51
2.1.9	Inteligencia artificial en la automatización de procesos.....	55
2.1.10	Analítica de datos.	60

2.1.11	Aplicación de Power BI en la analítica de datos.	63
	CAPÍTULO III: MARCO METODOLÓGICO	67
3.1	TIPO DE INVESTIGACIÓN	68
3.1.1	Enfoque de la investigación.	68
3.2	FUENTES Y SUJETOS DE INFORMACIÓN	69
3.2.1	Fuentes primarias.	69
3.2.2	Fuentes secundarias.	70
3.2.3	Sujetos de información.	71
3.3	TÉCNICAS Y HERRAMIENTAS DE RECOLECCIÓN DE DATOS	73
3.3.1	Cuestionarios o encuestas.	73
3.3.2	Observación.	74
3.4	VARIABLES DE INVESTIGACIÓN	75
3.5	DISEÑO DE LA INVESTIGACIÓN	77
3.6	MATRIZ DE COHERENCIA	80
	CAPÍTULO IV: DIAGNÓSTICO	83
4.1	DESCRIPCIÓN DE LA SITUACIÓN ACTUAL	84
4.1.1	Diagnostico operativo.	85
4.1.2	Diagnostico técnico.	91
4.1.3	Diagnóstico de percepción.	94
4.1.4	Determinación de brechas.	107

CAPÍTULO V: DISEÑO Y DESARROLLO DEL PROYECTO	114
5.1 DESARROLLO DE LA PROPUESTA	115
5.1.1 Contexto de la situación actual.....	115
5.1.2 Actores.....	116
5.1.3 Historias de usuario.....	117
5.1.4 Requerimientos funcionales.	118
5.1.5 Requerimientos no funcionales.	121
5.1.6 Especificación de requerimientos.	122
5.1.7 Casos de uso	132
5.1.8 Descripción casos de uso	133
5.1.9 Diagramas de casos de uso.	141
5.1.10 Diseño de la interfaz.	151
5.1.11 Herramientas de inteligencia artificial y analítica de datos.	154
5.1.12 Diseño de la base de conocimiento propuesta.....	166
5.1.13 Propuesta de automatización.	174
CAPÍTULO VI: CONCLUSIONES Y RECOMENDACIONES	182
6.1 CONCLUSIONES	183
6.1.1 Conclusión 1: Análisis del proceso actual.	183
6.1.2 Identificación herramientas de inteligencia artificial y analítica de datos.	183
6.1.3 Base de conocimiento estructurada.	184

6.1.4	Propuesta de integración de inteligencia artificial y analítica de datos.	185
6.2	RECOMENDACIONES	185
6.2.1	Implementación de la propuesta	186
6.2.2	Implementación y capacitación en herramientas de inteligencia artificial y analítica de datos	186
6.2.3	Implementación de una base de conocimiento centralizada	187
6.2.4	Propuesta de integración	187
	BIBLIOGRAFÍA	189

TABLA DE FIGURAS

Figura 1 Diagrama causa y efecto	30
Figura 2 El ciclo de vida de Scrum	38
Figura 3 Ilustración de virtualización.....	40
Figura 4 ¿Cómo funciona una API?	44
Figura 5 PLC System	48
Figura 6 Tecnologías Industria 4.0.....	51
Figura 7 Base de conocimiento	54
Figura 8 Ejemplos de donde podría utilizar inteligencia artificial.	57
Figura 9 Gráfico de barras de William Playfair	61
Figura 10 Modelo MapReduce.....	62
Figura 11 Cuadrante mágico de Gartner para plataformas de análisis e inteligencia empresarial	65
Figura 12 Etapas del diseño de investigación.....	78
Figura 13 Personal involucrado proceso actual.....	85
Figura 14 Vista RVTools	86

Figura 15	Diagrama de flujo proceso actual	89
Figura 16	Herramientas tecnológicas utilizadas.....	93
Figura 17	Pregunta 1	95
Figura 18	Pregunta 2	96
Figura 19	Pregunta 3	97
Figura 20	Pregunta 4	98
Figura 21	Pregunta 6	101
Figura 22	Pregunta 7	102
Figura 23	Pregunta 10	106
Figura 24	Diagrama de caso de uso CU-1	142
Figura 25	Diagrama de caso de uso CU-2.....	143
Figura 26	Diagrama de caso de uso CU-3.....	144
Figura 27	Diagrama de caso de uso CU-4.....	145
Figura 28	Diagrama de caso de uso CU-5.....	146
Figura 29	Diagrama de caso de uso CU-6.....	147
Figura 30	Diagrama de caso de uso CU-7.....	148
Figura 31	Diagrama de caso de uso CU-8.....	149
Figura 32	Diagrama caso de uso CU-9	150
Figura 33	Página de login.....	151

Figura 34	Página de inicio.....	152
Figura 35	Página generación de reportes.....	152
Figura 36	Página visualización de reportes	153
Figura 37	Página asistente virtual.....	154
Figura 38	Integración API PyVmomi.....	156
Figura 39	Diagrama de flujo extracción de datos.....	157
Figura 40	Integración ChatGPT gestión de consultas	160
Figura 41	Integración API ChatGPT	161
Figura 42	Integración PowerBi, visualización de datos	163
Figura 43	Integración Power BI	165
Figura 44	Diagrama Entidad-Relación Conceptos	167
Figura 45	Tablas obtención de datos vCenter	171
Figura 46	Tablas filtrado de datos VMs	171
Figura 47	Tablas cálculos costos.....	172
Figura 48	Flujo base de conocimiento.....	173
Figura 49	Ciclo de vida de minería de datos	175
Figura 50	Propuesta base de conocimiento integrada.	176
Figura 51	Diagrama de flujo aplicando propuesta de automatización.	178
Figura 52	Ejemplos tareas programadas Windows.	179

Figura 53 Diagrama de flujo automatización mediante base de conocimiento.....	181
--	-----

INDICE DE TABLAS

Tabla 1 Cronograma	35
Tabla 2 Sujetos de información	72
Tabla 3 Tabla de operacionalización de variables	76
Tabla 4 Matriz de coherencia	81
Tabla 5 Lista de datos incluidos en el informe.	88
Tabla 6 Cálculo de tiempo por actividad.....	91
Tabla 7 Listado de respuestas pregunta 5	100
Tabla 8 Listado de respuestas pregunta 8.	103
Tabla 9 Listado de respuestas pregunta 9.	105
Tabla 10 Listado respuestas pregunta 11.....	107
Tabla 11 Tabla de actores del sistema.	117
Tabla 12 Historias de usuario.	117
Tabla 13 Requerimientos funcionales.	120
Tabla 14 Requerimientos no funcionales.	122
Tabla 15 Requerimiento RF-01.	123

Tabla 16 Requerimiento RF-02.	124
Tabla 17 Requerimiento RF-03.	125
Tabla 18 Requerimiento RF-04.	126
Tabla 19 Requerimiento RF-05.	127
Tabla 20 Requerimiento RF-06.	128
Tabla 21 Requerimiento RF-07.	129
Tabla 22 Requerimiento RF-08.	130
Tabla 23 Requerimiento RF-09.	131
Tabla 24 Listado casos de uso.	132
Tabla 25 Caso de uso CU-1.	133
Tabla 26 Caso de uso CU-2.	134
Tabla 27 Caso de uso CU-3.	135
Tabla 28 Caso de uso CU-4.	136
Tabla 29 Caso de uso CU-5.	137
Tabla 30 Caso de uso CU-6.	138
Tabla 31 Caso de uso CU-7.	139
Tabla 32 Caso de uso CU-8.	140
Tabla 33 Caso de uso CU-9.	141
Tabla 34 Tabla de categorías.	168

Tabla 35 Tabla de conceptos169

DEDICATORIA

Primeramente, quiero agradecerle a Dios por permitirme llevar a cabo la elaboración de este proyecto, sin el sosteniéndome a lo largo del proceso no hubiera sido posible llegar a este punto de conclusión de esta licenciatura, adicionalmente le doy gracias infinitas por colocar en mi vida a personas sumamente importantes que de forma directa o indirecta me han apoyo a concluir con mis estudios.

Seguidamente quiero agradecerle a mi excompañero de trabajo y a mi gran amigo Román Garay por permitirme desarrollar la propuesta de este proyecto en base a su compañía y adicionalmente por siempre estar atento a cualquier apoyo requerido; más que un promotor para mí ha sido un amigo que me ha brindado su brazo en cuanto a lo he necesitado y agradezco infinitamente todo su apoyo en este proceso tan importante para mí. Asimismo, le agradezco a su compañía Artificial Skills por permitirme basar este proyecto en base a sus procesos y por siempre brindarme cualquier información requerida de primera mano.

También agradezco enormemente a todas las personas que de una u otra forma han sido parte del proceso, a mis ex compañeros de universidad, a mis compañeros de trabajo, a mi familia y amigos que me han acompañado a lo largo del camino universitario. Mención especial a mi perrito Bruno por siempre acompañarme en las noches largas de estudio.

Y finalmente quiero agradecerle y dedicarle este logro a Elena Quesada; la persona más importante para mí en este proceso, la persona que siempre estuvo detrás mío impulsándome a salir adelante y a continuar con mis estudios, a quien siempre estuvo motivándome y recordándome que era capaz de lograrlo, a la persona que a pesar de no llevar mí misma sangre siempre se sacrificó por mí para que pudiera concluir mis estudios, hoy soy lo que soy gracias a todo lo que ella me ha enseñado. Soy consciente de que una dedicatoria no basta para agradecerle todo lo que ha hecho por mí, pero sí de alguien es este logro es de ella.

**CARTA DE ACEPTACIÓN DEL PATROCINADOR DEL
PROYECTO**

16 de julio de 2024

Señores(as)
Universidad Hispanoamericana

Estimados señores(as)

Reciba un cordial saludo. Por este medio, me permito manifestar el interés de la ARTIFICIAL SKILLS en el proyecto denominado "*Propuesta de uso de un sistema experto basado en inteligencia artificial para la automatización del proceso de analítica de datos de VMware vCenter en Artificial Skills (ASK)*" presentado por el estudiante José Hernán Delgado Arce, cédula de identidad 1-1741- 0308. Este proyecto se realizará en el área de Análisis de Datos y es de suma importancia para nuestra empresa.

Explorar la implementación de una inteligencia artificial en uno de nuestros procesos resulta crucial para la mejora en la toma de decisiones basadas en datos, así como para obtener una ventaja competitiva en el área de análisis de datos de nuestros clientes, la cual es el área en la que nuestra empresa se desenvuelve. Por esta razón, permitir el desarrollo de este proyecto en nuestra institución es de vital importancia para nosotros. Nos comprometemos a brindarle al estudiante todos los insumos necesarios para su desarrollo.

Quedo atento a cualquier consulta o detalle adicional.

Atentamente,

Roman Alexander Andrade Garay

Director de Datos e Inteligencia Artificial
Artificial Skills

Cedula Jurídica: 3-102-777255

Correo electrónico roman.garay@artificial-skills.com

A handwritten signature in black ink, appearing to be the name Roman Alexander Andrade Garay, is written over a large, faint, stylized orange and yellow graphic that resembles a network or data flow diagram.

DECLARACION JURADA

Declaración Jurada

Yo José Hernán Delgado Arce mayor de edad, portador de la cédula de identidad número 1-1741-0308 egresado de la carrera de Ingeniería en sistemas de la Universidad Hispanoamericana, hago constar por medio de éste acto y debidamente apercibido y entendido de las penas y consecuencias con las que se castiga en el Código Penal el delito de perjurio, ante quienes se constituyen en el Tribunal Examinador de mi trabajo de tesis para optar por el título de Licenciatura con Énfasis en Sistemas de la Información juro solemnemente que mi trabajo de investigación titulado: " Propuesta de automatización de procesos mediante una base de conocimiento integrada con inteligencia artificial y analítica de datos en Power BI para la plataforma VMware vCenter en Artificial Skills (ASK)", es una obra original que ha respetado todo lo preceptuado por las Leyes Penales, así como la Ley de Derecho de Autor y Derecho Conexos número 6683 del 14 de octubre de 1982 y sus reformas, publicada en la Gaceta número 226 del 25 de noviembre de 1982; incluyendo el numeral 70 de dicha ley que advierte; artículo 70. Es permitido citar a un autor, transcribiendo los pasajes pertinentes siempre que éstos no sean tantos y seguidos, que puedan considerarse como una producción simulada y sustancial, que redunde en perjuicio del autor de la obra original. Asimismo, quedo advertido que la Universidad se reserva el derecho de protocolizar este documento ante Notario Público.

En fe de lo anterior, firmo en la ciudad de San José, a los 03 días del mes de abril del año dos mil veinticinco.

**JOSE HERNAN
DELGADO ARCE
(FIRMA)**



Firmado digitalmente por
JOSE HERNAN DELGADO
ARCE (FIRMA)
Fecha: 2025.04.03 10:13:01
-06'00'

Firma del estudiante

Cedula 1-1741-0308

CARTA DEL TUTOR

CARTA DEL TUTOR

San José, 04 de abril de 2025

Kattia Isabel Huertas Elizondo
Directora
Ingeniería Informática
Universidad Hispanoamericana
Sede Llorente

Estimada señora:

El estudiante **JOSÉ HERNÁN DELGADO ARCE**, cédula de identidad número **1-1741-0308**, me ha presentado, para efectos de revisión y aprobación, el trabajo de investigación denominado **"Propuesta de automatización de procesos mediante una base de conocimiento integrada con inteligencia artificial y analítica de datos en Power BI para la plataforma VMware vCenter en Artificial Skills (ASK)"**, el cual ha elaborado para optar por el grado académico de Licenciatura en Ingeniería Informática.

En mi calidad de tutor, he verificado que se han hecho las correcciones indicadas durante el proceso de tutoría y he evaluado los aspectos relativos a la elaboración del problema, objetivos, justificación; antecedentes, marco teórico, marco metodológico, tabulación, análisis de datos; conclusiones y recomendaciones.

De los resultados obtenidos por el postulante, se obtiene la siguiente calificación:

a) Original del tema	10%	10%
b) Cumplimiento de entrega de avances	20%	20%
c) Coherencia entre los objetivos, los instrumentos aplicados y los resultados de la investigación	30%	30%
d) Relevancia de las conclusiones y recomendaciones	20%	20%
e) Calidad, detalle del marco teórico	20%	20%
TOTAL		100%

En virtud de la calificación obtenida, se avala el traslado al proceso de lectura.

Atentamente,

JULIAN RAQUEL
CORDOBA
SANABRIA (FIRMA)

Firmado digitalmente por
JULIAN RAQUEL CORDOBA
SANABRIA (FIRMA)
FIRMA DE CALIDAD TECNICA
2025

Julián Córdoba Sanabria

Cédula 109640134

CARTA DEL LECTOR

San José,

Universidad Hispanoamericana
Sede Llorente
Carrera de Informática

Estimado señor

El estudiante José Hernán Delgado Arce, cédula de identidad 1-1741-0308, me ha presentado para efectos de revisión y aprobación, el trabajo de investigación denominado “Propuesta de automatización de procesos mediante una base de conocimiento integrada con inteligencia artificial y analítica de datos en Power BI para la plataforma VMware vCenter en Artificial Skills (ASK)”.

He revisado y he hecho las observaciones relativas al contenido analizado, particularmente lo relativo a la coherencia entre el marco teórico y análisis de datos, la consistencia de los datos recopilados y la coherencia entre éstos y las conclusiones; asimismo, la aplicabilidad y originalidad de las recomendaciones, en términos de aporte de la investigación.

Por consiguiente, este trabajo cuenta con mi aval para ser presentado en la defensa pública.

Atte.

**Randall
Vargas
Villalobos** Firmado digitalmente
por Randall Vargas
Villalobos
Fecha: 2025.05.22
08:43:18 -06'00'

Firma

Randall Vargas Villalobos
Cédula: 1-1140-0113

LICENCIA DE AUTORIZACIÓN

**UNIVERSIDAD HISPANOAMERICANA
CENTRO DE INFORMACION TECNOLOGICO (CENIT)
CARTA DE AUTORIZACIÓN DE LOS AUTORES PARA LA CONSULTA, LA
REPRODUCCION PARCIAL O TOTAL Y PUBLICACIÓN ELECTRÓNICA
DE LOS TRABAJOS FINALES DE GRADUACION**

San José, 30 de mayo del 2025

Señores:
Universidad Hispanoamericana
Centro de Información Tecnológico (CENIT)

Estimados Señores:

El suscrito (a) JOSÉ HERNÁN DELGADO ARCE con número de identificación 1-1741-0308 autor (a) del trabajo de graduación titulado "Propuesta de automatización de procesos mediante una base de conocimiento integrada con inteligencia artificial y analítica de datos en Power BI para la plataforma VMware vCenter en Artificial Skills (ASK)" presentado y aprobado en el año 2025 como requisito para optar por el título de LICENCIATURA EN INGENIERIA INFORMATICA; SI autorizo al Centro de Información Tecnológico (CENIT) para que con fines académicos, muestre a la comunidad universitaria la producción intelectual contenida en este documento.

De conformidad con lo establecido en la Ley sobre Derechos de Autor y Derechos Conexos N° 6683, Asamblea Legislativa de la República de Costa Rica.

Cordialmente,

JOSE HERNAN DELGADO ARCE (FIRMA)
 Firmado digitalmente por
JOSE HERNAN DELGADO
ARCE (FIRMA)
Fecha: 2025.05.30 11:34:40
-06'00'

Ced. 1-1741-0308
Firma y Documento de Identidad

CAPÍTULO I: PROBLEMA DEL PROYECTO

1.1 ANTECEDENTES Y JUSTIFICACION DEL PROYECTO.

Introducción

En el competitivo y vertiginoso entorno tecnológico actual, las empresas dedicadas a la prestación de servicios de tecnologías de la información y comunicación (TI) enfrentan desafíos constantes para mantenerse a la vanguardia; el constante aumento en la complejidad de las infraestructuras virtualizadas debido al aumento de nuevas tecnologías, combinada con la necesidad de tomar decisiones rápidas basadas en datos precisos para la mejora en el uso de recursos y principalmente en la gestión de costos, obliga a las organizaciones a buscar soluciones que optimicen sus operaciones y les proporcionen una ventaja competitiva.

En este contexto, la automatización de procesos y la integración de herramientas avanzadas de analítica de datos e inteligencia artificial emergen como enfoques estratégicos fundamentales para mejorar la eficiencia y capacidad de respuesta de las empresas que, como Artificial Skills (ASK) que operan en este sector; se encuentra ante este agudo desafío de explorar nuevas herramientas y tecnologías que impulsen su posición en los temas de la analítica de datos e inteligencia artificial en infraestructuras virtualizadas.

Actualmente, la posibilidad de contar con datos precisos en tiempo real es fundamental para tomar decisiones estratégicas en la dirección de TI, que permita tomar acciones acertadas fundamentadas en cuanto al uso de recursos y la gestión de costos; para ASK implica reaccionar de manera ágil a los cambios del mercado y anticiparse a ellos. La maximización del uso de la información recopilada a lo largo del tiempo, junto con una comprensión profunda del entorno empresarial, proporciona una visión holística que permite identificar oportunidades de mejora y limitaciones con mayor claridad para la toma de decisiones.

Debido a toda esta situación surge la necesidad de utilizar tecnologías sobresalientes como una base de conocimiento estructurada en inteligencia artificial que permita una gestión automatizada y eficiente de los datos de los recursos virtuales de la plataforma de VMware vCenter. Esta propuesta representa un paso crucial para Artificial Skills en su cometido de ofrecer soluciones avanzadas que no solo satisfagan las necesidades presentes, sino que también anticipen y aborden los desafíos futuros en el ámbito de la analítica de datos e inteligencia artificial.

1.1.1 Marco de referencia empresarial y contextual.

El presente proyecto será llevado a cabo en Artificial Skills empresa fundada en el año 2021, dedicada a la prestación de servicios de analítica de datos e inteligencia artificial (IA) mediante una plataforma que se integra a la infraestructura tecnológica de diferentes clientes a nivel de la región, la cual centra su enfoque principalmente a la infraestructuras virtualizadas en la plataforma de VMware vCenter; desarrollando métricas y análisis personalizados que mediante técnicas avanzadas de IA, generan insights valiosos para la toma de decisiones estratégicas.

La misión de ASK es “Crear servicios que impulsen el éxito de nuestros clientes a través de soluciones inteligentes y orientadas al futuro, desbloquearemos el potencial oculto en sus datos y transformarlos en acciones concretas y estratégicas”. (ARTIFICIAL SKILLS, s.f.)

En cuanto a la visión de ASK es “Aspiramos a crear servicios muy innovadores en la región, transformando información en conocimiento valioso que inspire decisiones inteligentes.” (ARTIFICIAL SKILLS, s.f.).

1.1.2 Justificación del proyecto

El presente proyecto de investigación tiene como objetivo principal proponer una base de conocimientos automatizada que utilice inteligencia artificial y analítica de datos (Power BI), permitiendo la gestión eficiente de los diferentes datos de consumos de las máquinas virtuales en la plataforma VMware vCenter de Artificial Skills (ASK). Esta iniciativa busca dotar a las áreas clave de toma de decisiones, como las gerencias y direcciones de TI, de acceso a datos precisos en tiempo real; un aspecto fundamental en un entorno altamente dinámico y competitivo; tal como se afirma, “El análisis de la información facilita a las empresas la comprensión de sus negocios, mercados y a tomar decisiones empresariales a tiempo.” (Mamani, 2018, pág. 1).

Actualmente, la información sobre las infraestructuras virtualizadas depende de los departamentos de TI, lo que genera retrasos y posibles desactualizaciones de los datos disponibles, además de la posibilidad de errores derivados del factor humano; esta situación limita la capacidad de tomar decisiones de manera ágil y oportuna ya que no se tiene una base centralizada de estos datos. Al integrar un sistema automatizado, basado en inteligencia artificial, se optimizará el procesamiento y uso de los grandes volúmenes de datos generados por VMware vCenter; de este modo, se reducirá la dependencia de las áreas operativas y se garantizará que la información esté constantemente actualizada y accesible de forma inmediata.

Este proyecto permitirá generar una propuesta de automatización del análisis y presentación de los datos, y facilitará la identificación de patrones y tendencias que mejoren la toma de decisiones sobre el uso eficiente de los recursos; además, proporcionará a ASK una ventaja competitiva al optimizar la gestión de recursos virtuales y permitir una toma de decisiones estratégicas con datos precisos. En este sentido, ASK podrá anticiparse a los cambios del

mercado y responder con mayor rapidez, fortaleciendo su posición dentro del sector de la analítica de datos e inteligencia artificial aplicada a infraestructuras virtualizadas.

Por otro lado, la adopción de este enfoque ofrecerá beneficios tanto a nivel interno como externo; la empresa podrá mejorar sus operaciones internas, minimizando tiempos de respuesta y optimizando el uso de sus recursos, mientras que su cartera de clientes se beneficiará de soluciones más eficientes y personalizadas donde el análisis de datos se ha convertido en un motor de transformación empresarial, impulsando el desarrollo de estrategias personalizadas, la mejora de la eficiencia operativa y la creación de nuevas oportunidades de negocio. (Arboleda Mazo & Orozco Carvaja, 2018).

1.2 DEFINICION DEL PROBLEMA.

1.2.1 Problemática

La empresa Artificial Skills (ASK) enfrenta actualmente la necesidad urgente de optimizar el acceso y la gestión de datos de sus infraestructuras virtualizadas, con el fin de proporcionar información precisa y oportuna a las diferentes áreas de toma de decisiones; entre ellas, gerencias y direcciones de TI. En el contexto actual, la dependencia de los departamentos de tecnología para obtener la información necesaria genera retrasos considerables en la actualización de los datos disponibles; esta situación se debe, en gran parte, a procesos manuales que afectan negativamente la productividad y la capacidad de análisis de la información. Además, dicha dependencia aumenta el riesgo de errores humanos, lo que, a su vez, puede comprometer la calidad y exactitud de las decisiones estratégicas.

Un problema adicional es que los datos generados por las infraestructuras virtualizadas no se almacenan estructurados; en cambio, se guardan en archivos planos de Excel, lo que trae consigo inconvenientes: la dificultad de análisis profundos; la falta de integración eficiente de la información y, más importante, la pérdida de conocimiento en el tiempo. Este tipo de almacenamiento no solo limita la posibilidad de extraer valor de los datos, sino que también representa un riesgo en términos de sostenibilidad del conocimiento; esto impide que ASK tome decisiones basadas en un análisis integral de su infraestructura tecnológica, lo que compromete su agilidad y competitividad.

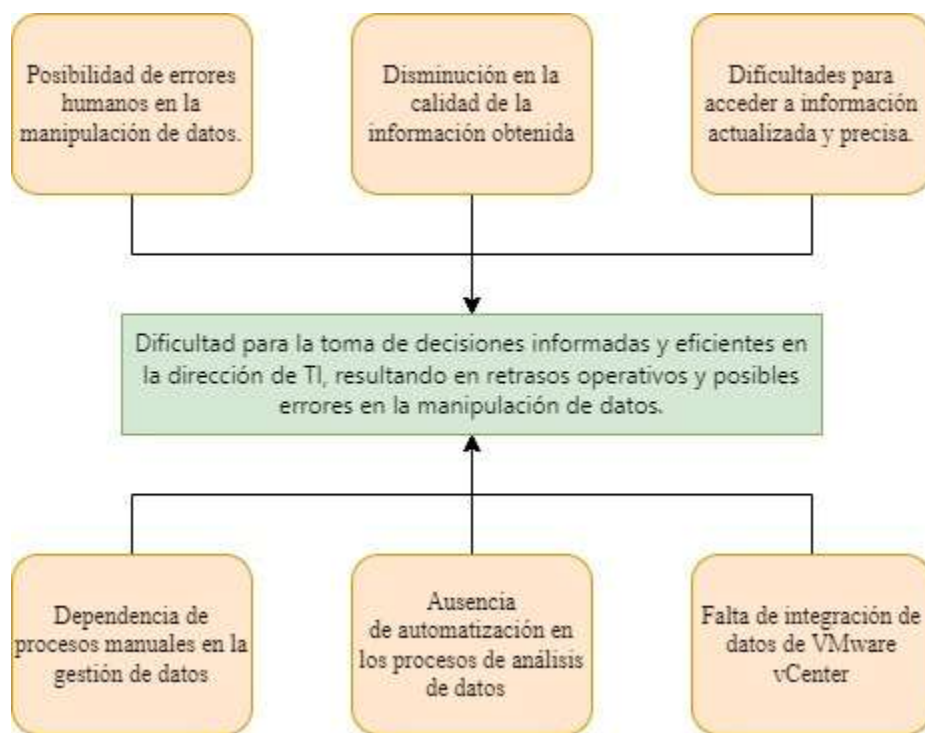
La propuesta del proyecto surge como una oportunidad estratégica para abordar esta problemática de manera integral; al implementar un sistema experto basado en inteligencia artificial, capaz de procesar y gestionar los datos de VMware vCenter de manera automatizada, se logrará agilizar el manejo de grandes volúmenes de información. Este enfoque permitirá ofrecer a las gerencias y direcciones de TI acceso a datos actualizados, veraces y ágiles, facilitando la toma de decisiones informadas (basadas en el comportamiento de su infraestructura tecnológica y patrones identificados en tiempo real).

Además, la creación de una base de conocimiento permitirá reducir drásticamente la dependencia de procesos manuales y de los departamentos de tecnología, asegurando la actualización constante de la información y proporcionando acceso autónomo a datos precisos y estructurados. Esta capacidad de acceder rápidamente a la información se traducirá en una ventaja competitiva para ASK; al mejorar la toma de decisiones (más rápidas y acertadas), se beneficiarán tanto la empresa como su cartera de clientes. Asimismo, la implementación de este proyecto no solo optimizará la eficiencia operativa de ASK, sino que también fortalecerá su

posición en el mercado, permitiéndole adaptarse mejor a un entorno empresarial cada vez más competitivo y exigente.

Figura 1

Diagrama causa y efecto



Fuente: Elaboración propia, 2024.

1.2.2 Problema General

1. ¿Podría la automatización del proceso de gestión de datos de máquinas virtuales en VMware vCenter mediante un sistema basado en inteligencia artificial y analítica de datos representar una ventaja competitiva para ASK, al mejorar la toma de decisiones, reducir tiempos operativos y mitigar errores humanos?

1.2.3 Problemas específicos

1. ¿Cuáles son los procesos actuales de gestión de los datos de las máquinas virtuales en la plataforma VMware vCenter que pueden ser automatizadas mediante inteligencia artificial y analítica de datos?
2. ¿Cómo se puede estructurar una base de conocimiento centralizada que optimice la gestión y acceso a la información para la integración y gestión de los datos de las máquinas virtuales en VMware vCenter?
3. ¿Qué herramientas de inteligencia artificial y analítica de datos son las más adecuadas para la automatización de procesos en la gestión de los datos de las máquinas virtuales en VMware vCenter?
4. ¿Cómo se pueden integrar de manera efectiva las herramientas de inteligencia artificial y analítica de datos (como Power BI) para automatizar los procesos de gestión de los datos de las máquinas virtuales en VMware vCenter?

1.3 OBJETIVOS DEL PROYECTO

1.3.1 Objetivo general.

1. Desarrollar una propuesta de automatización de procesos construyendo una base de conocimiento integrada con herramientas de inteligencia artificial y de analítica de datos con Power BI para la integración y gestión de datos de las máquinas virtuales en la plataforma VMware vCenter en la empresa Artificial Skills (ASK).

1.3.2 Objetivos específicos.

1. Analizar los procesos actuales de integración y gestión de máquinas virtuales en la plataforma VMware vCenter identificando áreas que puedan ser automatizadas o mejoradas mediante inteligencia artificial y analítica de datos.
2. Identificar las herramientas de inteligencia artificial y analítica de datos que pueden ser propuestas para la automatización de procesos en la gestión de los datos de las máquinas virtuales en la plataforma de VMware vCenter.
3. Proponer una base de conocimiento estructurada que permita centralizar y gestionar la información relacionada con la gestión de los datos de las máquinas virtuales en VMware vCenter.
4. Desarrollar una propuesta de integración de información aplicando herramientas de inteligencia artificial (IA) y analítica de datos (Power BI) para la automatización de procesos en la gestión de los datos de máquinas virtuales en la plataforma VMware vCenter.

1.4 ALCANCES Y LIMITACIONES

En este apartado se mencionarán los diferentes alcances y limitaciones que contemplará el presente proyecto de investigación.

1.4.1 Alcances

El presente proyecto será llevado a cabo con el respaldo del departamento de desarrollo y análisis de datos de Artificial Skills (ASK), con el objetivo de automatizar el proceso de gestión de los datos generados por la infraestructura de VMware vCenter. Para ello se desarrollará una base de conocimiento centralizada que permita almacenar toda la información de las máquinas

virtuales donde con la aplicación de herramientas de analíticas de datos e inteligencia artificial, estos datos puedan ser utilizados eficientemente para la toma de decisiones de TI.

El primer entregable del proyecto estará conformado por un diagnóstico detallado donde se evaluarán los procesos actuales de la gestión de los datos generados por las diferentes máquinas virtuales de la plataforma de VMware vCenter en Artificial Skills; con el fin de identificar las áreas de oportunidad de mejora mediante la automatización de procesos, enfocándose en tiempos operativos, cuellos de botella, pérdida de información y errores humanos que puedan ser mitigados mediante herramientas de inteligencia artificial y analítica de datos.

Seguidamente se investigarán diferentes herramientas de inteligencia artificial y analítica de datos que puedan ser utilizadas dentro el entorno de ASK; con el propósito de reducir los tiempos operativos manuales, optimizando los procesos de análisis y reportería de los datos generados por la plataforma de VMware vCenter.

Como tercer entregable se presentará una propuesta de creación de una base de conocimiento centralizada con el propósito de almacenar y gestionar eficientemente los grandes volúmenes de datos de las diferentes máquinas virtuales generados por la plataforma de VMware vCenter en ASK asegurando que estos sean precisos y actualizados continuamente.

Como fase final de la investigación, se desarrollará una propuesta formal de automatización de procesos a través de una base de conocimiento integrada con herramientas de inteligencia artificial y analítica de datos en la plataforma de VMware vCenter, estará propuesta expondrá los procesos que pueden ser automatizados, los resultados esperados de la automatización en cuanto a reducción de tiempos operativos y mitigación de errores; esta propuesta proporcionará a la dirección de TI de ASK herramientas para la toma de decisiones informadas en tiempo real, consolidando así una ventaja competitiva para la compañía.

Cabe mencionar que dicho proyecto no incluye la implementación de las herramientas propuestas dentro del entorno productivo, limitándose únicamente al desarrollo de la propuesta de la automatización con los estudios previos para dicha recomendación.

1.4.2 Limitaciones

Para realizar este proyecto se alterará o encriptará por políticas internas de la empresa relacionadas con la certificación vigente de la normativa internacional ISO 27001, cuyo objetivo es proteger la confidencialidad, integridad y disponibilidad de los datos manejados por la organización, los sistemas y aplicaciones que los procesan, divulgar nombres de clientes, cifras exactas y alguna otra información confidencial entre ASK y sus clientes. Adicionalmente cada documento entregable de este proyecto con relación a la empresa Artificial Skills (ASK) deberá seguir el lineamiento interno relacionado a la certificación ISO 27001, con el fin de cumplir con los estándares de la seguridad de la información.

1.5 CRONOGRAMA DE ACTIVIDADES

En el siguiente apartado se presenta el cronograma de actividades que establece una secuencia lógica y organizada de las tareas necesarias para cumplir con los objetivos planteados de esta investigación.

Tabla 1 Cronograma

NOMBRE TAREA	SEMANAS																			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Planificación del proyecto.																				
Justificación del proyecto.	■																			
Definición problema general y específicos.		■	■																	
Definición objetivos general y específicos.		■	■																	
Alcances y limitaciones.				■																
Total: 3 Semanas		■	■	■																
Marco Teórico.																				
Terminología.				■																
Definición temas principales.					■	■														
Total: 2 Semanas				■	■															
Marco metodológico.																				
Enfoque de la investigación.						■														
Fuentes de información.							■													
Técnicas de recolección de datos.								■	■											
Total: 3 semanas						■	■	■												
Análisis de la situación actual.																				
Diagnóstico inicial.									■											
Diagnostico operativo.										■	■									
Conclusiones.											■	■								
Total: 3 Semanas										■	■	■								
Propuesta del proyecto																				
Análisis de requerimientos.												■								
Modelado del sistema													■	■						
Desarrollo de la funcionalidad														■	■					
Pruebas del sistema																■				
Total: 6 semanas													■	■	■	■	■	■		
Conclusiones y recomendaciones.																				
Desarrollo de conclusiones.																		■	■	
Desarrollo de recomendaciones																			■	■
Total: 3 Semanas																		■	■	■

Fuente: Elaboración propia, 2024.

CAPÍTULO II: MARCO TEÓRICO

2.1 FUNDAMENTOS TEÓRICOS Y CONCEPTUALES.

Introducción

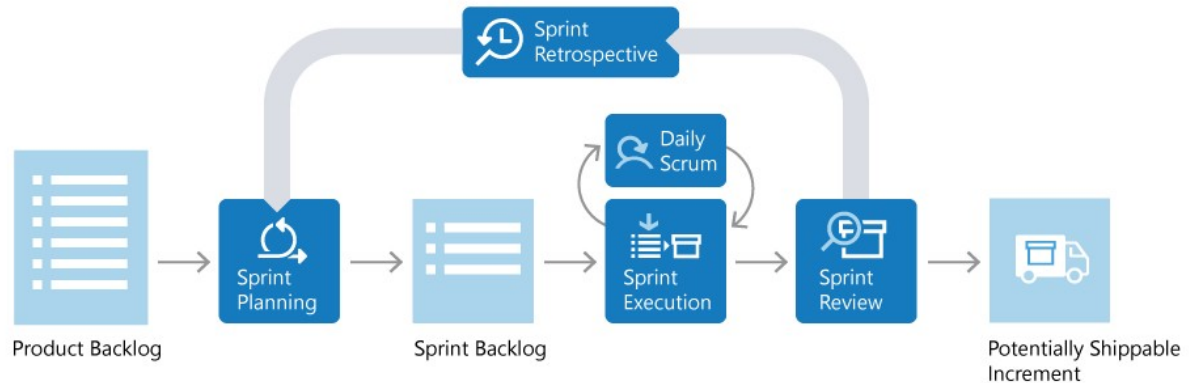
El presente capítulo pretende proporcionar los fundamentos teóricos y conceptuales necesarios para sustentar el desarrollo de este proyecto que tiene como objetivo principal la automatización de procesos de gestión de datos de las máquinas virtuales que residen en la plataforma de VMware vCenter en Artificial Skills, mediante la utilización de diferentes herramientas tanto de inteligencia artificial como de analítica de datos; en este sentido se expondrán las teorías, conceptos y tecnologías que permitirán entender el entorno y la viabilidad del uso de herramientas avanzadas para optimizar los procesos actuales dentro de la empresa Artificial Skills.

2.1.1 Scrum.

Scrum es un marco de trabajo que es empleado por los diferentes equipos de trabajo para gestionar tareas y abordar problemas de manera colaborativa en intervalos breves. Este enfoque permite adaptarse a cambios, mejorar continuamente y optimizar la entrega de valor. Basado en los principios de Agile, Scrum se estructura a través de un conjunto bien definido de artefactos, prácticas y roles, como el scrum master, el product owner y el equipo de desarrollo. Su objetivo principal es fomentar la transparencia, la inspección y la adaptación, facilitando así una gestión eficiente y flexible del trabajo (Microsoft Learn, 2023).

Figura 2

El ciclo de vida de Scrum



Fuente: Adaptado de *¿Qué es Scrum?*, por Microsoft Learn, 2023, Microsoft Learn (<https://learn.microsoft.com/es-es/devops/plan/what-is-scrum>)

2.1.2 Análisis de requerimientos.

El análisis de requerimientos es la etapa inicial en el desarrollo de software, en la que se llevan a cabo una serie de tareas para identificar, definir y documentar las necesidades o condiciones que debe cumplir un nuevo producto o una modificación de uno existente. Esta fase es fundamental para garantizar que el software desarrollado responda a las expectativas de los usuarios y los objetivos del negocio, permitiendo una planificación adecuada y la reducción de posibles errores en etapas posteriores del proyecto (Mendez, 2020).

2.1.3 Requerimientos funcionales.

Los requerimientos funcionales especifican cómo debe interactuar el sistema con su entorno, lo que incluye la participación de usuarios, la comunicación con otros sistemas, respuestas automatizadas y la ejecución de procesos predefinidos. Estos requerimientos detallan las funcionalidades y servicios que el software debe proporcionar para cumplir con su propósito.

En algunos casos, también pueden indicar explícitamente acciones que el sistema no debe realizar, siempre que estas restricciones contribuyan a garantizar un comportamiento funcional acorde con las expectativas de los usuarios o de otros sistemas con los que interactúa (Mendez, 2020).

2.1.4 Requerimientos no funcionales.

A diferencia de los requerimientos funcionales de los que se mencionó anteriormente, estos describen qué debe hacer el sistema, los requerimientos no funcionales definen los criterios que determinan la calidad y eficiencia de su operación. Estos requisitos son clave para evaluar si el sistema es adecuado para su uso, asegurando que funcione correctamente bajo diversas condiciones y satisfaga las expectativas de los usuarios y de la organización (Mendez, 2020).

2.1.5 VMware vCenter y la gestión de las máquinas virtuales.

Virtualización

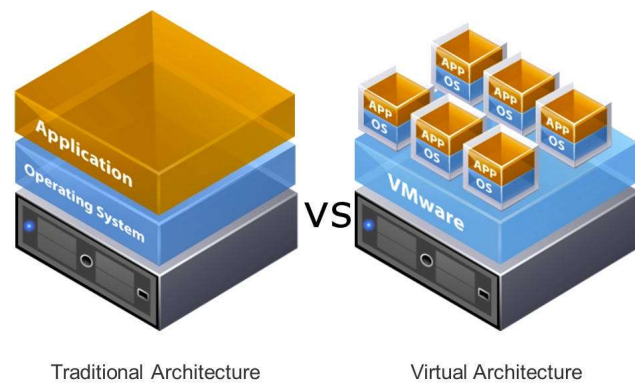
Para poder ampliar el concepto de VMware vCenter y el papel que este desempeña en el proceso de gestión de máquinas virtuales; primero debemos remontarnos unas épocas atrás donde surge el concepto de virtualización, centrándonos especialmente en la virtualización de servidores; este ha sido un concepto clave en la evolución de las diferentes tecnologías de la información; su origen lo podemos remontar a la década de 1960 donde IBM desarrollo varios modelos avanzados de su servidor modelo IBM 7044; uno de los modelos más destacado en ese momento fue el modelo 67; el cual destacó al virtualizar todas las interfaces de hardware a través de un Virtual Machine Monitor (VMM), conocido actualmente como “hipervisor”, que posee la capacidad de ejecutar sistemas operativos dentro de otros sistemas operativos.

Estos avances permitieron que los primeros sistemas operativos en máquinas virtuales, conocidos en aquel entonces como Conversational Monitor Systems (CMS) pudiesen ser ejecutados en entornos virtualizados; sentando un precedente para las futuras tecnologías de virtualización (Gómez López, Villar Fernández, & Alcayde García, 2011). En la actualidad podemos definir la virtualización como “Un proceso que permite una utilización más eficiente del hardware físico de la computadora y es la base de la computación en la nube” (IBM, 2023).

En el contexto de la virtualización de servidores, podemos definir el concepto de servidor como un ordenador tanto físico como virtual; el cual proporciona servicios a uno o varios clientes mediante una arquitectura cliente-servidor; dentro de este ordenador se pueden ejecutar diferentes aplicaciones donde se reciben y procesan peticiones de los clientes y se devuelve una respuesta adecuada (Breton, 2024). En este modelo de virtualización, las máquinas virtuales emulan servidores independientes, emulación que permite ejecutar una amplia variedad de sistemas operativos y aplicaciones dentro de un solo servidor físico, optimizando los recursos físicos y permitiendo una mayor flexibilidad operativa.

Figura 3

Ilustración de virtualización



Fuente: Adaptado de *An Overview of Virtual Machines*, por Madeline Farina, 2021, Medium (<https://medium.com/reverse-engineering-for-dummies/an-overview-of-virtual-machines-d409fc89ad8f>)

VMware vCenter

Conforme evoluciono el concepto de virtualización de servidores cada vez se volvía más complicado la administración de estos; donde una compañía utiliza una pequeña cantidad de servidores virtualizados la administración es controlable y se pueden monitorear con alguna herramienta de monitoreo básica ya que no requiere una mayor atención sin embargo cuando el número de estos aumenta a cientos la administración se vuelve un proceso mucho más complejo que requiere de herramientas mucho más avanzadas para lograr una infraestructura virtualizada controlada (Marchionni & Formoso, 2013).

Bajo este contexto es donde VMware vCenter surge como una de las herramientas más utilizadas para la administración y gestión de grandes cantidades de servidores y máquinas virtuales en la actualidad. Según la página oficial de VMware by Broadcom (2024) podemos definir VMware vCenter como: “un software de administración de servidores avanzado que proporciona una plataforma centralizada para controlar sus entornos VMware vSphere, lo que le permite automatizar y ofrecer una infraestructura virtual en toda la nube híbrida con confianza” (Broadcom Inc, 2005-2024).

Es importante aclarar que la implementación de VMware vCenter no solo simplifica la administración centralizada de la gran cantidad de máquinas virtuales, sino que adicionalmente promueve una mejora significativa en la eficiencia operativa, con la automatización de tareas repetitivas, monitoreo del rendimiento de cada máquina virtual en tiempo real y optimización de los recursos disponibles; sin embargo al ser una plataforma centralizada la misma genera una

gran cantidad de datos sobre el rendimiento y estado de los diferentes servidores y máquinas virtuales que residen en la plataforma; estos datos al ser utilizados de forma adecuada se vuelven fundamentales para el análisis y la toma de decisiones informadas, permitiendo optimizar aún más los procesos y los recursos del infraestructura tecnológica.

2.1.6 Interfaz de programación de aplicaciones (API, por sus siglas en inglés).

Historia de las APIs (Application programming interfaces)

Según la página oficial de Red Hat (2023) las APIs nacieron en los primeros días de la informática, mucho antes de que las computadoras personales fueran parte de la normalidad como en la actualidad. En sus inicios, las APIs se diseñaron como bibliotecas específicas para sistemas operativos; donde su función principal residía en facilitar a los desarrolladores el acceso a ciertas funcionalidades del sistema; en aquel entonces estas APIs en general operaban localmente, algunas de ellas permitían la comunicación entre computadoras centrales, ampliando su utilidad en entornos aún más complejos.

Con el pasar del tiempo, estas interfaces dejaron de estar dedicadas exclusivamente a sistemas locales; alrededor de los años 2000, las APIs comenzaron a consolidarse como herramientas fundamentales para la integración remota de datos; en otras palabras, posibilitaban la conexión entre sistemas diferentes tanto en equipos locales como remotos, lo que ayudo a impulsar una nueva era en la interacción entre plataformas, mejorando significativamente la interoperabilidad tecnológica (Red Hat, 2023).

Definición de las APIs (Application programming interfaces)

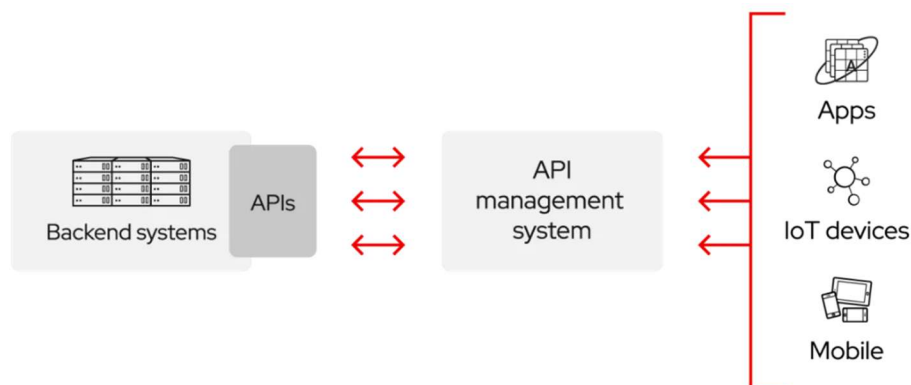
Basados en la definición proporcionada por la página oficial de Red Hat, podemos definir una API como: “un conjunto de definiciones y protocolos que se usa para diseñar e integrar el

software de las aplicaciones” (Red Hat, 2023). Basados en esta definición su principal objetivo es permitir que más de dos aplicaciones puedan intercambiar datos o utilizar funciones entre sí sin conocer sus implementaciones internas, básicamente las APIs actúan como una especie de "contrato" entre las partes involucradas: establecen qué solicitudes pueden hacerse, cómo deben hacerse y qué respuestas se esperan a cambio. Esta estandarización no solo simplifica el desarrollo de software, sino que también mejora la interoperabilidad entre diferentes sistemas, haciendo que los desarrolladores puedan integrar aplicaciones externas de manera más rápida y eficiente.

En la práctica, una API puede ser utilizada en múltiples ámbitos; desde acceder a un servidor para extraer datos o hasta integrar servicios de terceros como sistemas de pagos, mapas o redes sociales en una aplicación. Por ejemplo, al utilizar una API de mapas, una aplicación puede mostrar ubicaciones en tiempo real sin necesidad de desarrollar su propio sistema de geolocalización. Asimismo, las APIs son fundamentales en el mundo del desarrollo actual, ya que promueven la modularidad, la reutilización de componentes y la escalabilidad en proyectos de software; convirtiéndolas en esenciales para ámbitos como el desarrollo web, aplicaciones móviles, internet de las cosas y servicios en la nube, permitiendo la creación de sistemas aún más robustos.

Figura 4

¿Cómo funciona una API?



Fuente: Adaptado de *¿Cómo funciona una API?*, por Red Hat, 2023, Red Hat Inc (<https://www.redhat.com/es/topics/api/what-are-application-programming-interfaces>)

Ventajas de las APIs (Application programming interfaces)

Como se mencionó anteriormente el uso de las APIs ayudan a simplificar significativamente el proceso de diseño y programación de aplicaciones nuevas o bien de nuevos servicios, adicionalmente a que facilita la integración y la gestión de aplicaciones ya existentes. Sin embargo, ofrece además una lista de beneficios complementarios que cabe destacar porque van alineados a lo que se desea lograr con el desarrollo de este proyecto.

1. Mejora de la colaboración y la conectividad entre plataformas.

Las API como se mencionó anteriormente son fundamentales para integrar distintas aplicaciones y plataformas lo que permite que se comuniquen entre sí de una manera eficiente; ayudando a que en un entorno empresarial donde una organización promedio utiliza más de mil aplicaciones en la nube, las cuales en su mayoría se encuentran desconectadas entre sí, las API eliminan los silos de información que obstaculizan la productividad. Esta integración ayuda a automatizar

flujos de trabajo y fomenta la colaboración entre equipos, unos de los factores críticos que buscamos con el desarrollo de este proyecto (Goodwin, 2024).

2. Innovación y nuevas oportunidades de negocio.

Uno de los puntos más destacables de las API es la flexibilidad que brindan; permitiéndole a las empresas diversificar sus servicios, crear nuevas alianzas comerciales y acceder a mercados emergentes, factor crítico para el éxito de las compañías modernas. Esto no solo acelera la transformación digital, sino que también amplía las oportunidades de ingresos. Ejemplos como Stripe, caso de éxito que inició con una API sencilla que poseía apenas 7 líneas de código y ahora es líder en servicios financieros globales, demuestra el potencial innovador al implementar este tipo de tecnología. Las API les permiten a las empresas adaptarse rápidamente a las demandas cambiantes del mercado y responder a estas demandas con soluciones que le ayuden a mejorar su competitividad en el mercado (Goodwin, 2024).

3. Monetización de datos.

Muchas empresas han transformado sus APIs en fuentes de ingresos significativas, ofreciendo acceso a activos digitales valiosos. En algunos casos, se ofrecen inicialmente de forma gratuita para atraer desarrolladores y generar interés en su uso; más adelante, estas APIs se monetizan mediante modelos de suscripción o tarifas por acceso. Por ejemplo, AccuWeather logró atraer miles de desarrolladores y vender miles de claves de API en menos de un año, generando una comunidad sólida que sostiene su negocio actualmente. Este enfoque convierte a las API en herramientas no solo para la integración, sino también para la expansión de las compañías en el mercado (Goodwin, 2024).

Automatización de infraestructuras virtualizadas mediante PyVmomi

Para el desarrollo de este proyecto se pretende proponer la utilización de la API llamada PyVmomi; la cual es una herramienta de código abierto desarrollada en el lenguaje de programación Python, misma que sirve como un enlace para interactuar con varias APIs basadas en SOAP (simple object access protocol) de VMware vSphere, plataforma de virtualización en la que se encuentra basada este proyecto. PyVmomi se centrará en facilitar la conexión, autenticación y gestión de servicios asociados con infraestructuras virtuales, ofreciendo una capa de conveniencia sobre las API fundamentales de VMware Infrastructure Management (VIM, por sus siglas en inglés).

Otra ventaja adicional de PyVmomi se basa en que permite interactuar de forma simplificada con diferentes sistemas provenientes de VMware, tal como el sistema de autenticación de vCenter Single Sign-On, como el ESX Agent Manager y el Storage Policy API, convirtiéndola en una herramienta ideal para automatizar tareas de administración en entornos virtualizados, optimizando tiempos operativos y reduciendo al máximo los errores en procesos complejos relacionados a la administración de la infraestructura. Su diseño flexible y la gran cantidad de documentación extensiva hacen que PyVmomi sea especialmente útil para el presente proyecto de automatización e integración de servicios de entornos VMware vCenter (Broadcom Inc, 2023).

2.1.7 Automatización de procesos.

Historia de la automatización

La automatización de procesos remonta sus inicios durante la revolución industrial un periodo en el cual se transformó por completo la producción de manufactura y en el cual se presentó una gran influencia en el desarrollo de tecnologías que reemplazaron muchas tareas repetitivas y manuales. Uno de los hitos más importantes de este periodo fue la invención de la máquina de

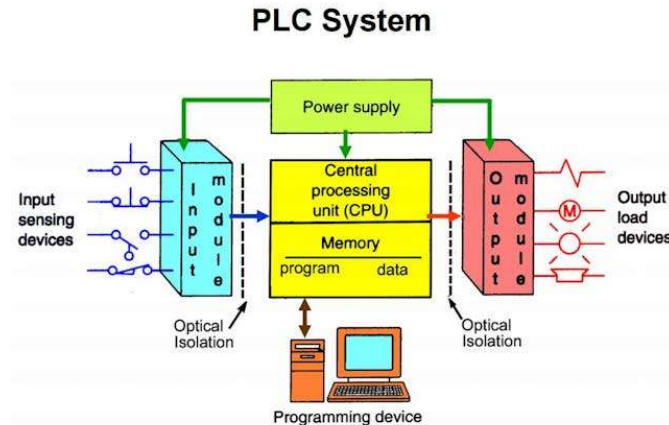
vapor de James Watt en 1775, la cual les permitió a las fábricas mejorar la eficiencia de sus operaciones. Este avance marcó el inicio de una era donde las maquinas mecánicas eran las encargadas de ejecutar tareas manuales como el ensamblaje y la manufactura con un nivel de precisión y rapidez mayor en comparación con los seres humanos (Groover, 2014).

El siguiente gran avance en cuanto a la automatización de procesos se presenta a mediados del siglo XX, en el transcurso de la tercera revolución industrial, donde el proceso de automatización experimentó un cambio radical con la incorporación de nuevas tecnologías digitales. La invención de los controladores lógicos programables (PLC) permitió que los sistemas de automatización se convirtieran en sistemas más adaptables y eficientes al integrar programación digital para el control sobre máquinas y procesos denominándose estos como sistemas computarizados (Groover, 2014).

La intervención de los PLC permitió el control automático de líneas de producción completas, ya que permitían personalizar y configurar diferentes procesos sin modificar el hardware de los sistemas adyacentes; representando un gran avance frente a los sistemas mecánicos y electrónicos y sentando las bases para lo que hoy conocemos como automatización industrial, donde los robots y sistemas programables manejan las operaciones de manufactura y ensamblaje.

Figura 5

PLC System



Fuente: Adaptado de *PLC System*, por Machine Design, 2015, Machine Design

<https://www.machinedesign.com/learning-resources/engineering-essentials/article/21834250/engineering-essentials-what-is-a-programmable-logic-controller>

En la actualidad nos encontramos inmersos en la cuarta revolución industrial, también conocida como la Industria 4.0 en la que la automatización ha presentado pasos agigantados con la introducción de tecnologías avanzadas como la inteligencia artificial y el big data.

Automatización

La evolución de la automatización, desde las primeras máquinas mecánicas de la revolución industrial hasta los sistemas digitales y tecnológicos actuales, ha dado lugar a un concepto de automatización más amplio y complejo que abarca no solo la ejecución de tareas repetitivas, sino también la optimización y el control de procesos complejos. Según la página oficial de IBM (2023) podemos definir la automatización como: “la aplicación de tecnología, programas, robótica o procesos para lograr resultados con una intervención humana mínima” (IBM, 2023).

La definición refleja la esencia de automatizar, cuyo objetivo no solo es reemplazar las tareas manuales y repetitivas, sino también mejorar significativamente la eficiencia operativa, la calidad del servicio o producto y la agilidad de la operación en diversos sectores, ya que la automatización puede aplicarse a muchos ámbitos; desde el sector financiero, y hasta la medicina, servicios públicos, fabricación automotriz, ámbito industrial, gestión de datos, logística, entre otros.

Además de su amplia aplicabilidad; uno de los principales beneficios de la automatización es la reducción de errores humanos, lo que se traduce en una mayor precisión y confiabilidad de los procesos; aumentando la calidad de los servicios o productos que se ofrezcan. La implementación de soluciones automatizadas también permite una mejor recopilación y análisis de datos, lo que resulta en una toma de decisiones más informada y basadas en evidencia permitiéndoles a las compañías dedicar más tiempo a actividades estratégicas que requieran un mayor nivel de pensamiento crítico; IBM (2023) afirma que todas aquellas organizaciones que empleen con mayor eficiencia el proceso de automatización podrán obtener beneficios competitivos significativos (IBM, 2023).

Industria 4.0

Cuando hablamos sobre el proceso de automatización en la actualidad es vital hacer referencia a la cuarta revolución industrial también llamada como la Industria 4.0. Este término, acuñado por primera vez en Alemania, describe la integración de diferentes tecnologías digitales avanzadas en los sistemas productivos y operacionales de todas las industrias, creando redes de valor más inteligentes y conectadas entre sí (Garrell & Guilera, 2019). Este periodo es de suma importancia abarcarlo en este proyecto debido a que la Industria 4.0 está marcada por la automatización inteligente, el uso de la inteligencia artificial (IA), y el big data, elementos que

permiten una gestión más eficiente y optimizada de los recursos, al tiempo que minimizan la intervención humana.

Retrocediendo en el tiempo durante la tercera revolución industrial las computadoras se integraron en las industrias y centros de producción lo que permitía en aquel entonces la automatización de tareas, más sin embargo las computadoras operaban de forma aislada es decir no tenían conexión entre sí y tampoco poseían conexiones con otras áreas de la organización; donde los datos se procesaban y almacenaban en sistemas locales; limitando de gran manera la capacidad de análisis de estos. Con la llegada de la Industria 4.0 se superan estas limitaciones al conectar los sistemas mediante redes avanzadas de comunicación y plataformas de big data, facilitando la recopilación, visualización y análisis de datos en tiempo real de toda la organización (Richardson, 2022).

Esta conectividad también hace posible integrar inteligencia artificial (IA) y análisis predictivo permitiendo generar insights precisos y anticipando posibles fallos o cuellos de botella en los procesos. Como resultado, la Industria 4.0 no solo elimina el aislamiento de los sistemas informáticos, sino que transforma los datos en un recurso estratégico, permitiendo niveles de eficiencia y personalización mucho más altos, una mejora en la toma de decisiones y una capacidad de respuesta mucho más ágil a las demandas del mercado.

La Industria 4.0 es posible gracias a la integración de diferentes tecnologías avanzadas, que integradas entre sí componen un entorno conectado, inteligente y adaptable. Entre estas tecnologías destacan el Edge computing, la robótica y los sistemas ciberfísicos, que permiten una sincronización precisa entre los procesos físicos y digitales. Además, tecnologías como el Internet de las Cosas (IOT) y la computación en la nube híbrida facilitan la recopilación y el procesamiento de datos en tiempo real, mientras que el Big Data y la visualización de datos

aportan análisis profundos que optimizan la toma de decisiones de las gerencias (Richardson, 2022).

Figura 6

Tecnologías Industria 4.0



Fuente: Adaptado de *Industria 4.0*, por Machine Monitoring, 2023, Machine Monitoring (<https://www.machinemonitoring.cl/web/>)

2.1.8 Bases de conocimiento como sistema de soporte a toma de decisiones.

Actualmente, las organizaciones que tienen sistemas de automatización que les permitan mejorar su eficiencia operativa, deben contar con sistemas aliados que les permitan contar con información confiable y en tiempo real para una toma de decisiones ágil y efectiva; porque los datos evolucionaron constantemente pasando de ser simples números o letras a activos estratégicos que permiten disminuir tiempos, reducir costos, aumentar la productividad y realizar análisis predictivos complejos.

Sistemas de apoyo a la toma de decisiones (DSS)

En relación con la necesidad de las compañías de contar con sistemas que permitan la automatización de procesos, como la gestión eficiente de la información surgen los denominados sistemas de información (SI); estos SI están compuestos por una estructura organizacional sólida donde la información puede fluir entre distintas áreas, mejorando la toma de decisiones integrales; cada SI se basa en cuatro actividades básicas: entrada, almacenamiento, procesamiento y salida de información. Es aquí donde surgen los sistemas de apoyo a decisiones (DSS) los cuales son fundamentales para las gerencias y responsables de la toma de decisiones, ya que les brindan información precisa, resumida y visualmente accesible en tiempo real (Bonilla Botia & Briceño Díaz, 2006).

Podemos definir los sistemas de apoyo a la toma de decisiones como una herramienta de inteligencia de negocio cuyo enfoque principal es la analítica de datos de una compañía; a diferencia de un sistema de planificación de recursos empresariales (ERP, por sus siglas en inglés), estos permiten informes más dinámicos y personalizados que permitan a los usuarios ajustar los SI según sus necesidades e inquietudes. Según Thor Olavsrud (2024) podemos definir un sistema de apoyo a la toma de decisiones como: “un sistema de información interactivo que analiza grandes volúmenes de datos para fundamentar las decisiones empresariales” (Olavsrud, 2024).

Bases de conocimiento.

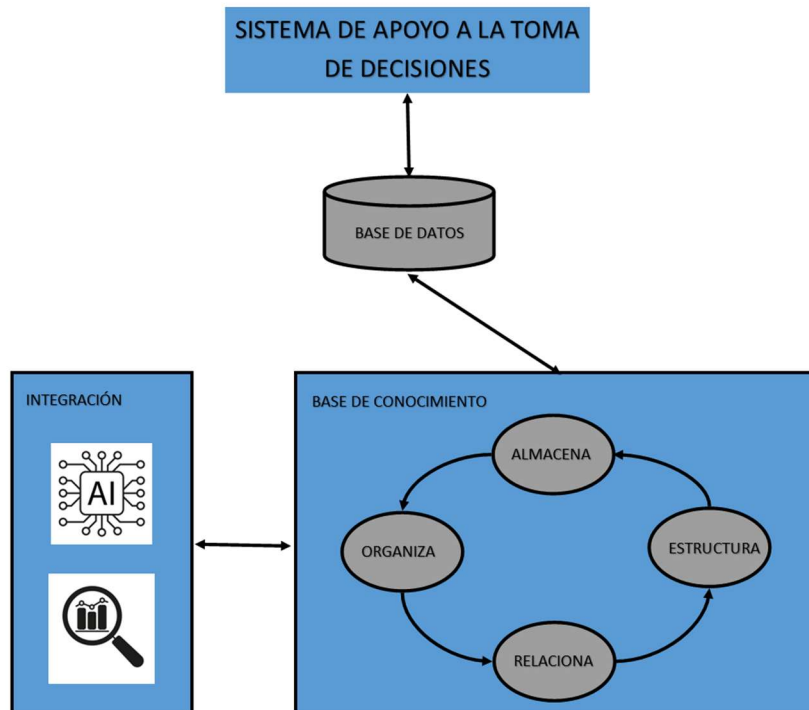
Una de las partes fundamentales de los sistemas de apoyo a la toma de decisiones son las bases de conocimiento; incluso estas pueden actuar como un sistema independiente al DSS, que puede integrarse con otros subsistemas o funcionar de manera independiente, permitiendo a las organizaciones mejorar la precisión y rapidez en sus decisiones. Este subsistema se conecta con los repositorios de conocimientos, también conocidos como sistemas de gestión del

conocimiento (KMS por sus siglas en ingles), para proporcionar un flujo continuo de información y análisis en tiempo real, disponible a través de servidores y tecnologías web (SHARDA, DELEN, & TURBAN, 2013).

En base a lo mencionado las bases de conocimientos son elementos cruciales para la toma de decisiones de cualquier compañía dado que; estas permiten que conforme las empresas enfrentan un crecimiento exponencial en la cantidad y diversidad de datos provenientes de diferentes fuentes que son generados por su operativa diaria, la base de conocimiento actúa como un repositorio centralizado que almacena, organiza, estructura y relaciona esta información en un único sistema, lo cual es especialmente importante en contextos donde se necesita responder a condiciones de mercado en constante cambio y tomar decisiones precisas con rapidez pero fundamentadas en datos reales.

Figura 7

Base de conocimiento



Fuente: Elaboración propia, 2024.

Al centralizar la información en una base de conocimiento esto garantiza que los datos relevantes estén disponibles para quienes los necesitan, adicionalmente facilita la integración con herramientas de inteligencia artificial y analítica de datos; las cuales facilita el análisis predictivo identificando patrones y comportamientos de los datos, otorgando una ventaja competitiva para la organización. Por otro lado, al gestionar grandes volúmenes de datos y la posibilidad de integración con inteligencia artificial, la base de conocimiento permite automatizar ciertos procesos de toma de decisiones, reduciendo la intervención humana en tareas repetitivas y aumentando la precisión en el análisis de datos complejos. En conjunto, estas

funciones de la base de conocimiento son esenciales para una gestión de datos y toma de decisiones eficiente en entornos empresariales modernos (SHARDA, DELEN, & TURBAN, 2013).

2.1.9 Inteligencia artificial en la automatización de procesos.

Como se ha mencionado a lo largo de esta investigación y principalmente en el periodo actual de la industria 4.0; la inteligencia artificial (IA) representa una herramienta clave en la automatización de procesos, debido a que permite no solo automatizar tareas repetitivas eliminando en gran parte la intervención humana, sino que también permite mejorar la toma de decisiones y optimizar flujos de trabajo aportando eficiencia operativa a nivel estratégico. Bajo este contexto la inteligencia artificial va más allá de la simple programación de tareas automatizadas; implica la implementación de algoritmos de aprendizaje automático, procesamiento de grandes volúmenes de datos y análisis predictivo para que los sistemas puedan adaptarse, aprender y mejorar continuamente.

Definición de inteligencia artificial.

En la actualidad definir el concepto de inteligencia artificial o IA se vuelve un poco complejo debido a que podemos encontrar diferentes conceptos de la misma dependiendo del entorno en el que se aplique, sin embargo Lasse Rouhiainen (2018) define la inteligencia artificial como: “la capacidad de las máquinas para usar algoritmos, aprender de los datos y utilizar lo aprendido en la toma de decisiones tal y como lo haría un ser humano” (Rouhiainen, 2018); por lo cual se puede concluir que la IA representa una ventaja en comparación con los humanos, dado que los sistemas de IA no requieren descanso, permitiéndoles realizar análisis complejos de datos en tiempo real y a gran escala, reduciendo significativamente la cantidad de errores. Adicionalmente la capacidad de aprendizaje y toma de decisiones de la IA ha transformado su

aplicación en múltiples sectores. Al adquirir la habilidad de mejorar sus procesos a través de técnicas de aprendizaje automático, la IA ahora desempeña roles anteriormente reservados solo para seres humanos, permitiendo a las organizaciones optimizar tareas de diagnóstico, predicción de tendencias y atención al cliente, entre otras. Esto ha llevado a una mayor eficiencia y precisión en áreas críticas, pero también plantea la necesidad de un enfoque cuidadoso y regulado para mitigar posibles desventajas de su crecimiento, como sesgos en los algoritmos o riesgos para la privacidad (Rouhiainen, 2018).

La inteligencia artificial tiene una aplicabilidad amplia y diversa que abarca desde áreas industriales hasta sectores sociales, revolucionando múltiples aspectos de la vida cotidiana y de las operaciones empresariales. En el sector de la salud, por ejemplo, la IA permite analizar grandes cantidades de datos médicos, detectar patrones y ayudar en diagnósticos tempranos, lo cual mejora la precisión en la atención y optimiza los tratamientos personalizados. Por otro lado; en el ámbito financiero, la IA facilita la detección de fraudes mediante algoritmos que examinan transacciones en tiempo real y analizan conductas inusuales, fortaleciendo así la seguridad y confianza en los sistemas bancarios; estos son únicamente algunos ejemplos de los diferentes sectores en los que la IA puede tener interacción.

Figura 8

Ejemplos de donde podría utilizar inteligencia artificial.



Fuente: Adaptado de Inteligencia artificial: 101 cosas que debes saber hoy sobre nuestro futuro (p.16), por L. Rouhiainen, 2018, Editorial Planeta S.A.

Inteligencia artificial integrada a bases de conocimientos.

A lo largo de la historia la IA ha sido fundamental para la construcción de las bases de conocimiento; recordando que las bases de conocimiento son el núcleo de los sistemas de apoyo a la toma de decisiones, ya que permiten almacenar, organizar y recuperar información específica de un dominio. Uno de los primeros sistemas expertos que demostró el potencial de estas bases de conocimiento fue Dendral, un sistema desarrollado en la Universidad de Stanford en 1965 por Edward Feigenbaum; en el cual se podían deducir estructuras químicas complejas, siendo pionero en la aplicación de IA para resolver problemas que requerían un conocimiento especializado en química.

Este sistema ejemplifica la importancia de las bases de conocimiento, ya que su éxito radicaba en su habilidad para almacenar y acceder a una amplia base de datos sobre estructuras

moleculares, lo que le permitía realizar deducciones precisas y relevantes para los expertos en la materia. Otro ejemplo destacado es Mycin, un sistema experto desarrollado en Stanford en los años 70 el cual fue iniciado por Edward Feigenbaum y finalizado por Edward Shortliffe; donde a través de una base de conocimiento estructurada y el uso de reglas de inferencia, Mycin podía sugerir tratamientos para enfermedades de la sangre basados en la sintomatología ingresada por el usuario (Ponce Gallegos, y otros, 2014).

Estos sistemas muestran cómo la inteligencia artificial, en conjunto con bases de conocimiento robustas, transforma el manejo de información especializada en un proceso eficiente, estructurado y accesible para los diferentes tipos de usuarios.

ChatGPT como herramienta de inteligencia artificial.

Chat Generative Pre-Trained Transformer (ChatGPT por sus siglas), es un modelo de lenguaje basado en inteligencia artificial desarrollado por OpenAI en noviembre del 2022, el mismo fue diseñado para interactuar con los usuarios que la utilicen de una forma conversacional; es decir este modelo de IA ha sido entrenado utilizando un enfoque que le permite participar en diálogos fluidos con los usuarios, comprender preguntas, reconocer sus propios errores, cuestionar premisas incorrectas e incluso rechazar solicitudes inapropiadas. La capacidad de mantener un diálogo natural y de aprendizaje constante lo convierte en una herramienta poderosa para una amplia gama de aplicaciones, desde la atención al cliente hasta la asistencia en la toma de decisiones empresariales (OpenAI, 2022).

ChatGPT es una herramienta de inteligencia artificial que se diferencia de otras debido a su enfoque interactivo, que le permite no solo generar respuestas, sino a la vez mantener conversaciones dinámicas enriqueciendo la experiencia del usuario; esta flexibilidad le permite adaptarse a una diversificación de contextos, convirtiéndola en una herramienta ideal para

integrar en diferentes sistemas tanto como de soporte, automatización de procesos, análisis de datos, y como parte de interfaces inteligentes que requieren una interacción constante con los usuarios (OpenAI, 2022).

Por su versatilidad y eficacia mencionada, desde su lanzamiento ChatGPT ha sido integrado en diversas plataformas y sectores de servicio desde su lanzamiento, ofreciendo una forma más natural e intuitiva de interactuar con los sistemas de inteligencia artificial y a la vez alcanzando el potencial de transformar la manera en que las empresas y organizaciones gestionan su información y brindan asistencia a sus diferentes clientes, mejorando tanto la eficiencia como la calidad del servicio al cliente y la toma de decisiones basadas en datos.

ChatGPT API para la integración con aplicaciones.

Según ChatGPT.es (2023) la API de ChatGPT es “una herramienta poderosa que ofrece a desarrolladores y usuarios la posibilidad de integrar IA en sus aplicaciones para entender y generar lenguaje natural” (ChatGPT.es, 2023). En otras palabras, la API de ChatGPT ofrece a los usuarios la capacidad de incorporar las funciones avanzadas de este modelo de inteligencia artificial a sus aplicaciones, permitiendo que estas interactúen de manera fluida con los usuarios en un formato conversacional. Esta tecnología facilita que las máquinas comprendan, procesen y generen respuestas coherentes relevantes a partir de entradas en lenguaje humano por el usuario, lo que optimiza significativamente la interacción entre el usuario y la aplicación.

La principal característica de la API de ChatGPT es que es extremadamente versátil, ofreciendo múltiples funcionalidades para mejorar las interacciones y procesos en una amplia variedad de aplicaciones. Una de ellas es la generación de contenido automatizada, permitiendo la creación de textos en varios idiomas para artículos, descripciones, correos electrónicos, entre otros.

Además, se puede utilizar en la traducción automática, facilitando la conversión de textos entre

diferentes idiomas, lo cual es ideal para crear plataformas en diferentes idiomas. En el ámbito de atención al cliente, la API permite la creación de chatbots inteligentes que entienden y responden en lenguaje natural, adaptándose a diversas plataformas como sitios web y aplicaciones móviles. También es capaz de realizar resúmenes automáticos, extrayendo las partes más relevantes de textos largos, lo cual resulta útil para aplicaciones de investigación o agregadores de noticias.

Para el beneficio de este proyecto ChatGPT es altamente eficaz en el análisis de contenido, permite analizar grandes volúmenes de datos para extraer patrones y tendencias; otra de sus potentes funcionalidades es la creación de recomendaciones personalizadas basadas en el comportamiento de los datos, mejorando la experiencia en plataformas de analítica de datos. Finalmente, gracias a su capacidad de adaptarse y crear modelos de lenguaje personalizados, la API de ChatGPT permite a los desarrolladores entrenar modelos específicos para tareas concretas, mejorando la eficiencia de las aplicaciones en diversos sectores (ChatGPT.es, 2023).

2.1.10 Analítica de datos.

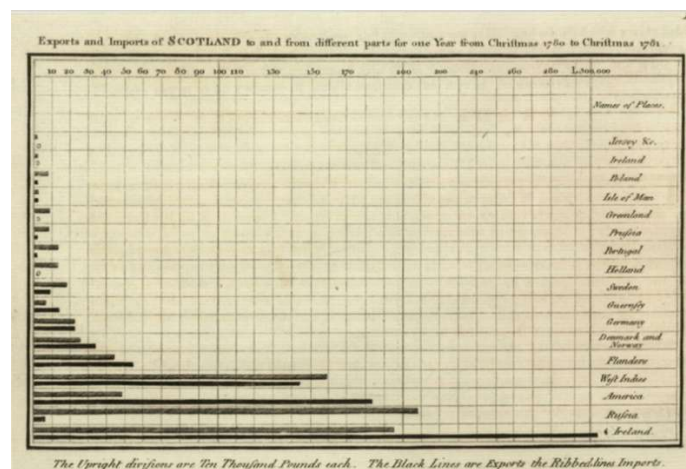
La analítica de datos es una disciplina que se centra en extraer información valiosa de los datos mediante el uso de técnicas matemáticas, estadísticas y computacionales. Este campo permite identificar patrones, tendencias y relaciones dentro de grandes volúmenes de datos también conocido en la actualidad como el big data, lo cual resulta esencial para la toma de decisiones informadas en diversas áreas, como los negocios, la salud, la educación y la tecnología. La capacidad de analizar y comprender datos en tiempo real ha dado lugar a una transformación en cómo las organizaciones interpretan y usan la información, haciendo de la analítica de datos una herramienta esencial en la era digital.

Historia de la analítica de datos.

La analítica de datos ha evolucionado a lo largo de siglos, donde algunos avances en herramientas utilizadas para manejar la información han marcado hitos importantes en la analítica de datos. En el año 1785, William Playfair introdujo el gráfico de barras, una de las técnicas más utilizadas en la visualización de datos; años después, Charles Joseph Minard creó un mapa que representaba visualmente las pérdidas del ejército de Napoleón en Rusia, relacionando datos sobre el clima y el tiempo con el tamaño del ejército. Posteriormente en el año 1890, Herman Hollerith desarrolló una máquina de tabulación que usaba tarjetas perforadas para procesar datos de manera más eficiente. Marcando un hito en la automatización y rapidez del análisis de datos.

Figura 9

Gráfico de barras de William Playfair

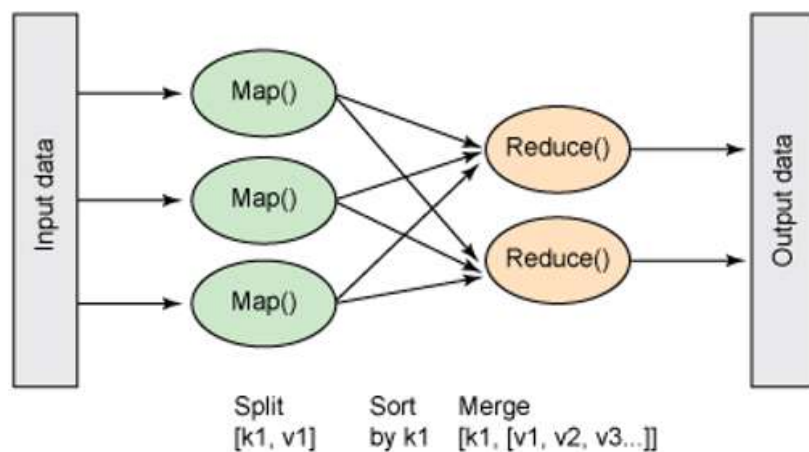


Fuente: Adaptado de *Gráfico de barras de William Playfair*, por Abel Méndez, 2021, Plan de Mejora (<https://www.plandemejora.com/william-playfair-biografia-y-aportaciones-a-la-estadistica/>)

Durante los años 70 y 80, la aparición de las bases de datos relacionales y el lenguaje SQL revolucionaron el análisis de datos en las compañías; en 1988, William H. Inmon popularizó el concepto de almacenes de datos. La minería de datos en los 90 permitió descubrir patrones en conjuntos de datos masivos, y los analistas comenzaron a usar lenguajes de programación como R y Python para explorar algoritmos de aprendizaje automático para que seguidamente en los 2000, la búsqueda en la web trajera innovaciones como MapReduce, Hadoop y Cassandra, herramientas clave en la era de big data. Avances que convirtieron el análisis de datos en una herramienta esencial para la toma de decisiones informadas en diversos sectores (Oracle Corporation, 2024).

Figura 10

Modelo MapReduce



Fuente: Adaptado de *MapReduce: el origen de la era Big Data*, por Mikel Niño, 2015, El blog de Mikel Niño (<https://www.mikelnino.com/2015/02/map-reduce-origen-era-big-data.html>)

Definición de analítica.

Según Oracle Corporation (2024) podemos definir el análisis como: “el proceso de descubrir, interpretar y transmitir patrones importantes presentes en los datos” (Oracle Corporation, 2024). Este proceso les facilita a las compañías y personas ver los datos importantes presentados en estadísticas de una forma más sencilla y entendible. Derivado a este proceso, surge lo que se conoce actualmente como análisis empresarial el cual es un estudio que implica la utilización de las estadísticas de los datos para la toma de decisiones informadas; permitiéndole a las empresas incrementar sus ingresos, reducir gastos operativos y mejoras empresariales.

Actualmente el análisis empresarial se aplica en la mayoría de los sectores industriales modernos; dado a que los datos en la actualidad es el activo más valioso para las compañías, donde estas pretenden con mayor frecuencia aprovechar más el análisis, accediendo a mayores volúmenes de datos para obtener estadísticas detalladas de manera rápida y accesible para más personas y departamentos dentro de la organización, a la vez que buscan reducir costos operativos. Adicionalmente, con el auge de las plataformas en la nube muchas empresas utilizan soluciones de autoservicio que permiten a los empleados sin conocimientos técnicos profundos realizar análisis por sí mismos, lo que facilita la toma de decisiones basada en datos en todos los niveles de la organización.

2.1.11 Aplicación de Power BI en la analítica de datos.

Bajo este contexto expuesto sobre la importancia de la analítica de datos, es fundamental abordar herramientas que facilitan el proceso de recopilación, análisis y visualización de información. En este sentido, Power BI se destaca como una de las soluciones más potentes y útiles del mercado. Esta herramienta desarrollada por Microsoft no solo permite a las organizaciones transformar datos en información útil, sino que también ofrece una interfaz intuitiva que facilita la creación de informes interactivos y paneles de control. A medida que las empresas buscan

optimizar su toma de decisiones mediante el análisis de datos, Power BI se presenta como una opción indispensable para aquellos que desean convertir datos complejos en informes visuales comprensibles.

Según la página oficial de Microsoft Learn (2024) Microsoft Power BI se puede definir como un conjunto integrado de herramientas de software y aplicaciones que facilitan la transformación de diferentes fuentes de datos en información coherente, visualmente atractiva y de fácil interacción. Power BI permite una conexión sencilla a las fuentes de datos, ayuda a visualizar y descubrir los aspectos clave y posibilita compartir esos insights con las personas o grupos que se desee (Microsoft, 2024). Esta herramienta se ha vuelto indispensable en el análisis de datos y en la toma de decisiones informadas, ya que permite a las organizaciones obtener una comprensión más profunda y ágil de sus datos, impulsando estrategias más precisas y alineadas con sus objetivos.

Este enfoque innovador y accesible le ha permitido a Microsoft posicionarse, una vez más, entre los líderes en el Cuadrante Mágico de Gartner 2024 en la categoría de plataformas de análisis e inteligencia empresarial; una categoría que busca destacar a las plataformas que les permita a las organizaciones comprender en profundidad sus datos y permitiéndoles la toma de decisiones informadas al responder preguntas claves sobre el negocio. Gracias a estas herramientas como Power BI, los usuarios pueden identificar patrones y tendencias de los datos permitiéndoles crear indicadores relevantes de rendimiento. La capacidad de combinar datos estructurados y no estructurados dentro de estas plataformas potencia aún más el análisis y la detección de anomalías, proporcionando a las organizaciones los conocimientos necesarios para optimizar sus decisiones y estrategias (Schlegel , y otros, 2024).

Figura 11

Cuadrante mágico de Gartner para plataformas de análisis e inteligencia empresarial



Fuente: Adaptado de *Cuadrante mágico para plataformas de análisis e inteligencia empresarial*, 2024, Gartner (<https://gartner.com/doc/reprints?id=1-2HVUGEM6&ct=240620&st=sb>)

El uso de Power BI se ha convertido en un pilar fundamental para las organizaciones que buscan mantenerse competitivas en un entorno moderno impulsado por datos. La capacidad de analizar, visualizar y compartir información clave; facilitando la toma de decisiones informadas y habilitando a las empresas para anticiparse a tendencias, optimizar sus operaciones y adaptarse rápidamente a las necesidades del mercado. Al posicionarse como líder en el Cuadrante

Mágico de Gartner, Power BI reafirma su relevancia como una plataforma estratégica que ayuda a convertir el potencial de los datos en una ventaja tangible y sostenible.

CAPÍTULO III: MARCO METODOLÓGICO

3.1 TIPO DE INVESTIGACIÓN

El tipo de investigación del presente proyecto es aplicada por que su objetivo principal es diseñar y desarrollar una propuesta de solución tecnológica para responder a una necesidad específica de la empresa Artificial Skills (ASK); mediante la construcción de una base de conocimiento integrada con herramientas de inteligencia artificial y analítica de datos. Esta investigación busca mejorar los procesos de integración y gestión de datos de las máquinas virtuales en la plataforma VMware vCenter, buscando ampliar el conocimiento teórico sobre el uso de la inteligencia artificial y el análisis de datos y adicionalmente proponer una herramienta funcional que permita a ASK mejorar su capacidad de toma de decisiones informadas, incrementando la eficiencia y control sobre sus activos tecnológicos.

3.1.1 Enfoque de la investigación.

Según lo establecido por Sampieri Hernández R. (2014) la presente investigación se caracteriza por un enfoque mixto, debido a que integra métodos cualitativos y cuantitativos para desarrollar una propuesta de automatización de procesos basada en inteligencia artificial y analítica de datos con Power BI (Sampieri, 2014).

La parte cualitativa de la investigación permitirá conocer las necesidades específicas de la empresa Artificial Skills en cuanto a la gestión de datos en su plataforma VMware vCenter, obteniendo información detallada sobre los procesos para optimizar la integración de datos de las máquinas virtuales. En cuanto al enfoque cuantitativo, se emplearán herramientas de análisis de datos para evaluar y medir el impacto de la automatización en la eficiencia operativa y en la toma de decisiones dentro ASK. Al combinar ambos enfoques, la investigación no solo describe y diagnostica el estado actual de la gestión de datos en ASK, sino que también proporciona datos

medibles y objetivos que respaldan el desarrollo de una posible solución para mejorar la gestión y visualización de datos.

3.2 FUENTES Y SUJETOS DE INFORMACIÓN

En este apartado se presentan las fuentes y sujetos de información esenciales para el desarrollo de la investigación. La identificación y definición de estos elementos es fundamental para garantizar que los datos recopilados sean precisos, relevantes y suficientes para responder a los objetivos del proyecto. En particular, se detallan las fuentes documentales y digitales consultadas, incluyendo bases de datos, informes técnicos y literatura académica, así como los sujetos involucrados en el proceso, como especialistas, usuarios clave y personal técnico de la empresa Artificial Skills (ASK).

3.2.1 Fuentes primarias.

Las fuentes primarias de este proyecto representan aquellos recursos que ofrecen información precisa y directa sobre los resultados originales de la investigación, proporcionando una base sólida para el desarrollo de la propuesta de automatización de procesos para la gestión de datos en VMware vCenter en la empresa Artificial Skills (ASK). Estas fuentes incluyen resultados obtenidos en primera instancia mediante encuestas, observaciones, análisis de datos específicos y revisión documental.

1. Observación directa: en primer lugar, se realizará una observación directa de los procesos actuales de gestión de datos en VMware vCenter en ASK, lo que permitirá registrar, de manera precisa, el flujo de trabajo actual, los desafíos operativos y las áreas de oportunidad de mejora en el entorno de virtualización.

2. Revisión documental interna: se realizará una revisión documental de los procesos y flujos internos definidos por Artificial Skills para una mayor comprensión del proceso de extracción y análisis de datos actual.
3. Cuestionario: se realizará un cuestionario (tipo encuesta) dirigido al personal técnico de Artificial Skills para recopilar información sobre sus experiencias, necesidades, problemas actuales y expectativas con relación a su operativa diaria y cargas de trabajo en el proceso de análisis de datos.
4. Datos de la plataforma de VMware vCenter: se analizarán los datos generados por la plataforma de VMware vCenter, como las métricas de rendimiento de las máquinas virtuales; los cuales serán fundamentales para evaluar el estado actual y definir los cambios propuestos en la gestión de estos.

3.2.2 Fuentes secundarias.

Las fuentes secundarias en este proyecto ofrecen información previamente recopilada y analizada por otros autores, lo cual es fundamental para comprender el contexto teórico y técnico en el que se desarrollará la propuesta de automatización. Estas fuentes secundarias incluyen literatura académica, informes de la industria y publicaciones técnicas, que proporcionan un marco de referencia para sustentar la investigación y el diseño de soluciones.

1. Literatura académica y artículos de investigación: se consultarán publicaciones académicas que aborden temas relacionados con inteligencia artificial y analítica de datos para identificar las mejores prácticas y enfoques relevantes en entornos similares, ayudando a fundamentar teóricamente la propuesta de automatización.
2. Informes de la industria: se revisarán informes técnicos y estudios de caso publicados por organizaciones de tecnología y consultoras que ofrezcan análisis sobre la

implementación de tecnologías de automatización. Esta información será útil para entender las tendencias actuales y los desafíos comunes en el sector de automatización y análisis de datos.

3. Documentación de proveedores de tecnología: se analizarán manuales, guías de uso y documentos técnicos de proveedores de herramientas como VMware y Microsoft (Power BI) para obtener un conocimiento detallado de las funcionalidades y limitaciones de estas plataformas que puedan ser aplicadas de manera eficiente en ASK.

3.2.3 Sujetos de información.

Los sujetos de información desempeñan un papel crucial al proporcionar el acceso a la información necesaria, ya sea a través de su experiencia en la plataforma tecnológica, en los procesos internos o en el análisis y gestión de los datos involucrados. La colaboración de los distintos participantes dentro de la organización será esencial para entender los flujos actuales, identificar áreas de mejora y garantizar que la propuesta final se ajuste a las necesidades reales de Artificial Skills.

A continuación, se presenta una tabla que detalla a los sujetos que participarán en este proceso, sus roles y el tipo de información que aportarán para el desarrollo del proyecto.

Tabla 2 *Sujetos de información*

Puesto laboral	Profesión u oficio	Experiencia	Relación con el tema
Director de datos e Inteligencia Artificial	Ingeniero en informática Máster en análisis de datos.	10 años	Fuente principal de los datos técnicos a utilizar durante la propuesta, adicionalmente es el punto de contacto para las reuniones con el equipo de trabajo.
Programador Sr	Licenciado en Ingeniería Informática	5 años	Fuente de información de las herramientas y métodos utilizados actualmente en el proceso.
Analista de datos I	Bachiller en ingeniería Informática	3 años en el sector de análisis de datos.	Fuente de información de los procesos internos en el análisis de datos.
Operador I	Técnico medio en informática.	1 año.	Fuente de información de los procesos internos en la recolección de los datos.

Fuente: Elaboración propia, 2024.

3.3 TÉCNICAS Y HERRAMIENTAS DE RECOLECCIÓN DE DATOS

Para la obtención de información precisa en el desarrollo de este proyecto, es crucial seleccionar adecuadamente las técnicas y herramientas de recolección de datos. Se utilizarán herramientas de recolección mixtas, que combinan instrumentos cuantitativos y cualitativos, lo cual permite una visión más completa y equilibrada del objeto de estudio (MEDINA ROMERO, y otros, 2023). Esta combinación ofrece mediciones numéricas precisas y a la vez una comprensión profunda y detallada de los datos. Además, cada técnica se ha elegido por su capacidad de adaptarse a las necesidades del proyecto, garantizando así que la información obtenida sea fiable y representativa. En el siguiente apartado, se describen las técnicas y herramientas implementadas, las cuales contribuirán al análisis y toma de decisiones el presente proyecto de investigación.

3.3.1 Cuestionarios o encuestas.

Según Medina Romero y otros (2023) las encuestas o cuestionarios son métodos de investigación diseñados para reunir datos de un amplio grupo de personas. Esta herramienta es flexible y fácil de aplicar, permitiendo a los investigadores acceder a información relevante sobre los comportamientos, actitudes, opiniones y características demográficas de una población específica (MEDINA ROMERO, y otros, 2023). Adicionalmente estas herramientas son sumamente versátiles ya que pueden ser aplicadas vía diferentes métodos en línea, como por teléfono, correo o formularios web permitiendo una flexibilidad en su aplicabilidad.

Debido a las ventajas mencionadas; los cuestionarios y encuestas se emplearán como herramientas iniciales para recopilar datos cuantitativos y cualitativos sobre las experiencias, necesidades y opiniones del personal técnico de la empresa Artificial Skills (ASK). Estas

herramientas permiten obtener información estructurada de una muestra significativa de participantes, lo cual facilita el análisis de patrones y tendencias en las respuestas. Las preguntas estarán diseñadas para explorar aspectos específicos de la gestión de datos en VMware vCenter, así como las expectativas y desafíos actuales del equipo en el manejo de esta plataforma.

Para esto; se llevará a cabo un cuestionario en formato de encuesta mediante Google forms, que incluirá tanto preguntas abiertas como cerradas; las preguntas abiertas permitirán a los participantes expresar sus opiniones y experiencias de forma detallada, mientras que las preguntas cerradas facilitarán la recopilación de datos específicos y cuantificables. La información obtenida será fundamental para identificar áreas clave de mejora y guiar el desarrollo de la propuesta de automatización de procesos, ajustándose a las necesidades operativas de ASK.

3.3.2 Observación.

Según Medina Romero y otros (2023) la técnica de la observación es: “un método de investigación en el que se registra y analiza el comportamiento y las acciones de individuos, grupos o fenómenos en su entorno natural” (MEDINA ROMERO, y otros, 2023). En el contexto de este proyecto, la observación directa permitirá obtener una visión detallada del flujo de trabajo actual y del proceso de gestión de datos de la plataforma de VMware vCenter dentro de la empresa Artificial Skills (ASK). Esta técnica contribuirá a identificar de manera precisa los procedimientos operativos, así como los desafíos que enfrentan los usuarios al interactuar con la plataforma.

A través de la observación, se busca capturar información clave sobre los métodos de trabajo y las interacciones del equipo técnico con la recolección y almacenamiento de datos de la plataforma de VMware vCenter, proporcionando una comprensión realista de la dinámica de

trabajo actual. Esta técnica complementará los datos obtenidos en las encuestas y cuestionarios, permitiendo una triangulación de información que refuerza la validez de los hallazgos y ayuda a establecer una base sólida para la propuesta de automatización de procesos.

Como se mencionó anteriormente, se empleará una técnica de observación no participante, la cual se caracteriza por la recopilación de información de manera directa, sin involucrarse activamente en el proceso observado. En este enfoque, el investigador se limita a observar y registrar los datos de los sujetos o procesos de estudio sin intervenir ni modificar su curso, lo que permite obtener una visión objetiva y precisa de las dinámicas en juego (MEDINA ROMERO, y otros, 2023). En el presente estudio, la observación no participante se aplicará de dos formas, las cuales se explican a continuación:

1. Observación directa: En este enfoque, se recopilará información directamente de los procesos operativos en la gestión de los datos de la plataforma de VMware vCenter en Artificial Skills (ASK) sin intervenir en el desarrollo de las actividades. Esta modalidad permitirá observar el flujo de trabajo en tiempo real y registrar los procedimientos, áreas de mejora y desafíos en el sistema.
2. Observación indirecta: Para complementar la observación directa, se analizarán documentos internos y registros históricos, como manuales de procesos y reportes previos de rendimiento. Este método permitirá obtener información adicional y contextualizar mejor los hallazgos obtenidos en la primera forma de observación.

3.4 VARIABLES DE INVESTIGACIÓN

En todo proyecto de investigación, las variables juegan un papel fundamental, ya que permiten medir, analizar y comparar los diferentes elementos que influyen en el proceso estudiado. En

este proyecto, se identificarán y definirán las variables clave que guiarán la propuesta de automatización de procesos para la gestión de datos en VMware vCenter en la empresa Artificial Skills (ASK). Estas variables estarán relacionadas con los aspectos técnicos, operativos y organizacionales involucrados en el proceso de gestión de datos, así como con los resultados esperados de la implementación de la solución propuesta en base a los objetivos plasmados. La identificación y el análisis de estas variables facilitarán la comprensión profunda del impacto de la automatización propuesta.

Tabla 3 *Tabla de operacionalización de variables*

Objetivo Especifico	Variable	Definición Conceptual
1. Analizar los procesos actuales de integración y gestión de máquinas virtuales en la plataforma VMware vCenter identificando áreas que puedan ser automatizadas o mejoradas mediante inteligencia artificial y analítica de datos.	Procesos actuales en VMware vCenter.	Evaluación del desempeño de los procesos actuales para identificar cuellos de botella y oportunidades de mejora.
2. Identificar las herramientas de inteligencia artificial y analítica de datos que pueden ser propuestas para la automatización de procesos en la gestión de los datos de las máquinas virtuales en la plataforma de VMware vCenter.	Herramientas de IA y analítica de datos.	Conjunto de tecnologías y herramientas específicas de IA y análisis de datos que son viables para automatizar la gestión en VMware vCenter.

<p>3. Proponer una base de conocimiento estructurada que permita centralizar y gestionar la información relacionada con la gestión de los datos de las máquinas virtuales en VMware vCenter.</p>	<p>Base de conocimiento estructurada.</p>	<p>Diseño de una estructura de base de conocimiento que permita la centralización de información relevante y facilite la toma de decisiones automatizada.</p>
<p>4. Desarrollar una propuesta de integración de información aplicando herramientas de inteligencia artificial (IA) y analítica de datos (Power BI) para la automatización de procesos en la gestión de los datos de máquinas virtuales en la plataforma VMware vCenter.</p>	<p>Propuesta de integración de IA y analítica de datos.</p>	<p>Propuesta detallada de integración de IA y Power BI en VMware vCenter, abordando la automatización de procesos clave en la gestión de datos.</p>

Fuente: Elaboración propia, 2024.

3.5 DISEÑO DE LA INVESTIGACIÓN

Según Sampieri Hernández R. (2014) el diseño de investigación es el marco estratégico que guía el proceso de recolección y análisis de datos, con el fin de responder las preguntas planteadas en el estudio y cumplir los objetivos propuestos (Sampieri, 2014). En investigaciones de enfoque mixto, como la presente, se combinan elementos tanto del diseño cuantitativo como cualitativo; este enfoque permite integrar datos numéricos y cualitativos para analizar de forma más profunda las necesidades, opiniones y comportamientos de los participantes. Este diseño será la

columna vertebral del estudio, orientando las fases, las herramientas y los métodos empleados en cada etapa, con el fin de garantizar que los resultados sean relevantes y coherentes con las metas de la investigación.

Este diseño establecido servirá como la guía fundamental para abordar el problema planteado en esta propuesta, proporcionando una posible solución de automatización para mitigar los desafíos identificados en la gestión de datos de máquinas virtuales de la plataforma de VMware vCenter. A continuación, en la siguiente figura, se presenta el diagrama de flujo que ilustra el proceso que se seguirá para desarrollar la propuesta del proyecto. Este diagrama visualiza las etapas clave que se llevarán a cabo para alcanzar los objetivos establecidos.

Figura 12

Etapas del diseño de investigación



Fuente: Elaboración propia, 2024.

Fase 1: Planificación del proyecto.

En esta primera fase se establecerá el marco estratégico del proyecto, en la cual se definirán claramente los objetivos, el alcance y los recursos necesarios para el desarrollo de la propuesta

de automatización. Además, se diseña un cronograma detallado, especificando tiempos y recursos para cada fase del proyecto. Esta etapa es de suma importancia debido a que ayuda a establecer una base sólida para el proyecto, alineando las expectativas con el alcance establecido.

Fase 2: Diagnostico de la situación actual.

La segunda etapa se centra en evaluar y documentar el estado actual del proceso de la gestión de datos de las máquinas virtuales en la plataforma de VMware vCenter en la empresa Artificial Skills (ASK). Se identificarán los puntos críticos, limitaciones, y las necesidades específicas del entorno actual para determinar las áreas clave en la que automatización pueda representar una ventaja competitiva. Este análisis incluye la observación directa sobre el proceso, la revisión de documentos internos y la realización de cuestionarios con el personal clave. Esta etapa permitirá desarrollar una comprensión profunda de las operaciones actuales, sentando las bases para que la propuesta de automatización esté fundamentada en las necesidades reales de la organización.

Fase 3: Investigación y recolección de información.

En esta fase, se recopilarán y analizarán los datos de diversas fuentes internas y externas, incluyendo estudios de caso, documentación técnica de VMware vCenter, documentaciones técnicas de Microsoft Power BI y tendencias en el uso de herramientas de inteligencia artificial para la gestión de datos. También se examinarán mejores prácticas de la industria y se realiza una revisión de los requisitos tecnológicos necesarios. Con esta investigación se pretende definir las características de las herramientas y metodologías más adecuadas para el proyecto, que sirva para fundamentar teórica y técnicamente; la construcción de la propuesta de automatización.

Fase 4: Desarrollo de la propuesta de automatización.

Seguidamente se construirá la propuesta en base a los datos y la investigación realizada. La cual incluirá una base de conocimiento que almacene los datos de una forma más estructurada; integrándola con herramientas de inteligencia artificial y analíticas de datos como Power BI, para la creación de un flujo de trabajo optimizado teniendo en cuenta los requisitos y características identificados en etapas previas. La propuesta final es un prototipo detallado que incluirá diagramas de flujo y especificaciones técnicas.

Fase 5: Presentación y validación de la propuesta.

Como última fase, se entregará el proyecto completo en un informe detallado, documentando la propuesta de automatización de la gestión de datos de las máquinas virtuales en la plataforma de VMware vCenter. El informe incluirá un análisis exhaustivo de la reducción de tiempos, mejoras en la eficiencia operativa y otros beneficios proyectados, basado en los resultados de las fases anteriores. Esta documentación final busca validar que la solución responde adecuadamente a los objetivos establecidos al inicio del proyecto y proporciona a la empresa Artificial Skills (ASK) una base sólida para considerar una futura implementación que le permita generar una ventaja competitiva en el mercado.

3.6 MATRIZ DE COHERENCIA

En esta sección, se presenta la matriz de coherencia que vincula los objetivos, entregables, instrumentos de recolección de datos y temas del marco teórico permitiendo comprender de forma integral el desarrollo y coherencia del proceso investigativo. La siguiente matriz pretende proporcionar un esquema detallado que asegura que todos los componentes del proyecto estén alineados con los objetivos propuestos, apoyando así una ejecución lógica y eficiente del estudio.

Tabla 4 *Matriz de coherencia*

Objetivo	Entregable	Fase o etapa que permite realizar el entregable.	Técnicas de recolección de la información	Instrumentos	Temas relacionados para marco teórico
1. Analizar los procesos actuales de integración y gestión de máquinas virtuales en la plataforma VMware vCenter identificando áreas que puedan ser automatizadas o mejoradas mediante inteligencia artificial y analítica de datos.	Informe de diagnóstico de los procesos actuales	Fase 2	Cuestionario tipo encuesta. Observación.	Encuestas. Análisis documental. Observación	Virtualización. VMware vCenter y la gestión de máquinas virtuales.
2. Identificar las herramientas de inteligencia artificial y analítica de datos que pueden ser propuestas para la automatización de procesos en la gestión de los datos de las máquinas virtuales en la plataforma de VMware vCenter.	Lista de herramientas viables para automatización	Fase 3	Informes de industria. Documentación de proveedores de tecnología.	Análisis de mercado de los softwares actuales en el campo.	Automatización de procesos. Interfaz de programación de aplicaciones. Inteligencia artificial en automatización de procesos. Analítica de datos. Aplicación de Power BI en la analítica de datos.

<p>3. Proponer una base de conocimiento estructurada que permita centralizar y gestionar la información relacionada con la gestión de los datos de las máquinas virtuales en VMware vCenter.</p>	<p>Documento con la estructura y modelo de la base de conocimiento</p>	<p>Fase 4</p>	<p>Literatura académica y artículos de investigación. Informes de industria. Documentación de proveedores de tecnología.</p>	<p>Revisión de literatura y software disponible</p>	<p>Sistemas de apoyo a la toma de decisiones. Bases de conocimiento. Interfaz de programación de aplicaciones</p>
<p>4. Desarrollar una propuesta de integración de información aplicando herramientas de inteligencia artificial (IA) y analítica de datos (Power BI) para la automatización de procesos en la gestión de los datos de máquinas virtuales en la plataforma VMware vCenter.</p>	<p>Propuesta de integración y automatización</p>	<p>Fase 4 y 5</p>	<p>Literatura académica y artículos de investigación. Informes de industria. Documentación de proveedores de tecnología.</p>	<p>Análisis técnico</p>	<p>Automatización de procesos. Interfaz de programación de aplicaciones Inteligencia artificial en automatización de procesos. Analítica de datos. Aplicación de Power BI en la analítica de datos.</p>

Fuente: Elaboración propia, 2024.

CAPÍTULO IV: DIAGNÓSTICO

4.1 DESCRIPCIÓN DE LA SITUACIÓN ACTUAL

Introducción.

En este capítulo se pretende llevar a cabo la fase 2 del presente proyecto; expuesta en el diseño de la investigación; teniendo como objetivo evaluar el estado actual del proceso de gestión de datos de las máquinas virtuales de la plataforma de VMware vCenter en Artificial Skills. Mediante este análisis se pretende identificar los desafíos y las oportunidades de mejora que enfrenta la organización en la gestión de los datos dentro de su infraestructura virtualizada; el cual nos permitirá fundamentar el desarrollo de la propuesta de automatización, permitiendo que el diseño de esta sea en base a las necesidades y desafíos reales de Artificial Skills.

Situación actual.

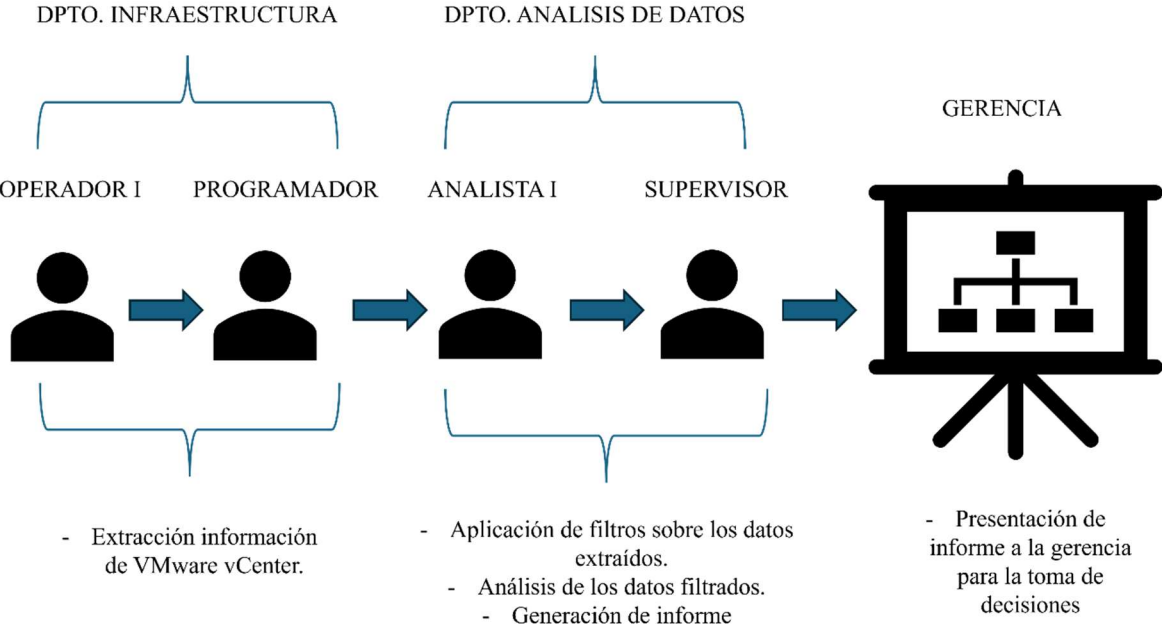
Actualmente el proceso de gestión de datos de la infraestructura virtualizada en VMware vCenter en Artificial Skills presenta una serie de desafíos significativos que le impiden sacar un máximo provecho de sus datos, lo cual afecta negativamente su eficiencia operativa y la toma de decisiones gerenciales informadas. El proceso actual se ejecuta de forma manual que si bien es cierto ha sido funcional a lo largo del tiempo, pero poco práctico; adicionalmente en la actualidad no existe una base de conocimiento donde se centralice toda la información generada por la plataforma lo que ocasiona que con el pasar de los meses dicha información se pierda y no pueda ser consultada en caso de ser necesario.

Al ser un proceso ejecutado manualmente genera una sobrecarga de trabajo, donde para generar un reporte de consumo de la infraestructura virtual utilizado para la gestión de cobros y tomas de decisiones sobre las infraestructuras como servicio tanto de Ask como de sus clientes; existe una dependencia de muchas áreas y personas involucradas dentro de la organización; generando

atrasos operativos (ocasionando tardías en las entregas de los informes a las diferentes gerencias), la posibilidad de cometer errores derivados por el factor humano y la omisión de información que pueda ser vital; que sumado genera un riesgo operativo – administrativo con consecuencias graves en la toma de decisiones y cobros administrativos.

Figura 13

Personal involucrado proceso actual



Fuente: Elaboración propia, 2024.

4.1.1 Diagnostico operativo.

En esta etapa se pretende evaluar el funcionamiento del proceso actual de gestión de datos de las máquinas virtuales en la plataforma de VMware vCenter desde el punto de vista operativo, dicho análisis incluye la revisión de activos empresariales como procesos y modelos operativos. El objetivo principal de este diagnóstico operativo es plasmar el diagrama de flujo del proceso actual para brindar una mayor claridad sobre el flujo de trabajo; sus problemáticas o áreas de

mejora que sustenten el desarrollo de este proyecto. La recolección de la información fue ejecutada mediante el método de observación, tanto directa como indirecta, donde se visualizó de primera mano la ejecución del proceso actual, registrando los hallazgos relevantes.

El proceso de gestión de datos actualmente da inicio cuando la gerencia de TI ya sea de forma mensual o bajo demanda solicita un reporte de consumo de la infraestructura virtualizada, con el fin de conocer los consumos de las diferentes máquinas virtuales que se ejecutan en VMware vCenter, esto para la generación de costos o bien para la toma de decisiones gerenciales. El operador de la infraestructura virtualizada es la primera persona encargada de generar un reporte manualmente desde la aplicación de RVTools; siendo esta una herramienta gratuita desarrollada para entornos de virtualización basados en VMware; proporcionando una interfaz simple para recopilar, visualizar y exportar información detallada sobre las máquinas virtuales, hosts y otros componentes de una infraestructura virtualizada.

Figura 14

Vista RVTools

VM	Powerstate	Template	SRM Placeholder	Config status	DNS Name	Connection state	Guest state	Heartbeat	Consolidation Needed	PowerOn	Suspend time	Creation
DEBIAN	poweredOff	False	False	green		connected	notRunning	gray	False			2020/11/
Debian_template	poweredOff	True	False	green		connected	notRunning	gray	False			2020/11/
MARS	poweredOff	False	False	green		connected	notRunning	gray	False			2020/11/
PLUTO	poweredOff	False	False	green		connected	notRunning	gray	False			2020/11/
SATURNUS	poweredOff	False	False	green		connected	notRunning	gray	False			2020/11/
Saturnus-IC	poweredOff	False	False	green		connected	notRunning	gray	False			2020/11/
vCLS (4)	poweredOn	False	False	green	None.None	connected	running	green	False	2021/04/02 16:00:56		2021/04/
vCLS (5)	poweredOn	False	False	green	None.None	connected	running	green	False	2021/04/02 15:55:12		2021/04/
vCLS (6)	poweredOn	False	False	green	None.None	connected	running	green	False	2021/04/02 16:09:46		2021/04/
VENUS	poweredOff	False	False	green		connected	notRunning	gray	False			2020/11/

Fuente: Adaptado de RVTools, 2024, Rvtools (<https://www.robware.net/home>)

Una vez que el operador de la infraestructura genera el reporte a través de la herramienta RVTtools, se descarga un archivo en formato .csv (valores separados por comas, por sus siglas en inglés). Este archivo contiene información detallada sobre las máquinas virtuales alojadas en la plataforma VMware vCenter. Entre los datos que se incluyen se encuentran el nombre asignado a cada máquina virtual, el sistema operativo residente, el estado actual, el clúster y el host en el que se encuentra alojada, así como todas las configuraciones asociadas. Además, el reporte proporciona información clave sobre los consumos de recursos, tales como CPU, memoria RAM, y almacenamiento en disco; extrayendo en su totalidad una lista de 95 datos diferentes por máquina virtual.

Posteriormente generado el archivo .csv por parte del operador de infraestructura este lo carga a una carpeta compartida de la nube de One Drive, inmediatamente el mismo debe notificarle al analista para que este pueda darle continuidad al proceso. El analista se encarga de convertir el archivo .csv a un archivo Excel identificando las columnas de consumos de VCPU (unidad central de procesamiento), RAM (memoria de acceso aleatorio), STORAGE (Almacenamiento), adicionalmente en el proceso se necesita excluir registros que no son cobrables o representativos. Todo este proceso de filtrado de datos puede significar un gran margen de error al no limpiar correctamente los datos o bien al cometer un error en el copiado de los datos.

Tabla 5 *Lista de datos incluidos en el informe.*

DATOS INCLUIDOS	
DATO	DESCRIPCIÓN
Nombre máquina virtual	Nombre único que permite identificar la máquina virtual dentro de la infraestructura virtualizada.
Nombre clúster	Nombre que permite identificar el clúster dentro de vCenter donde habita la máquina virtual.
Nombre host	Nombre identificador del host en el que habita la máquina virtual.
Sistema operativo	Dato utilizador para conocer el sistema operativo alojado en la máquina virtual.
Estado máquina virtual	Permite conocer el estado de energía de la máquina virtual (encendido o apagado).
Consumo CPU	Cantidad de unidades CPU consumidas por la máquina virtual, que permite posteriormente calcular un costo.
Consumo RAM	Cantidad de memoria RAM consumida por la máquina virtual, que permite posteriormente calcular un costo.
Consumo STORAGE	Cantidad de almacenamiento en disco consumido por la máquina virtual, que permite posteriormente calcular un costo.

Fuente: Elaboración propia, 2024.

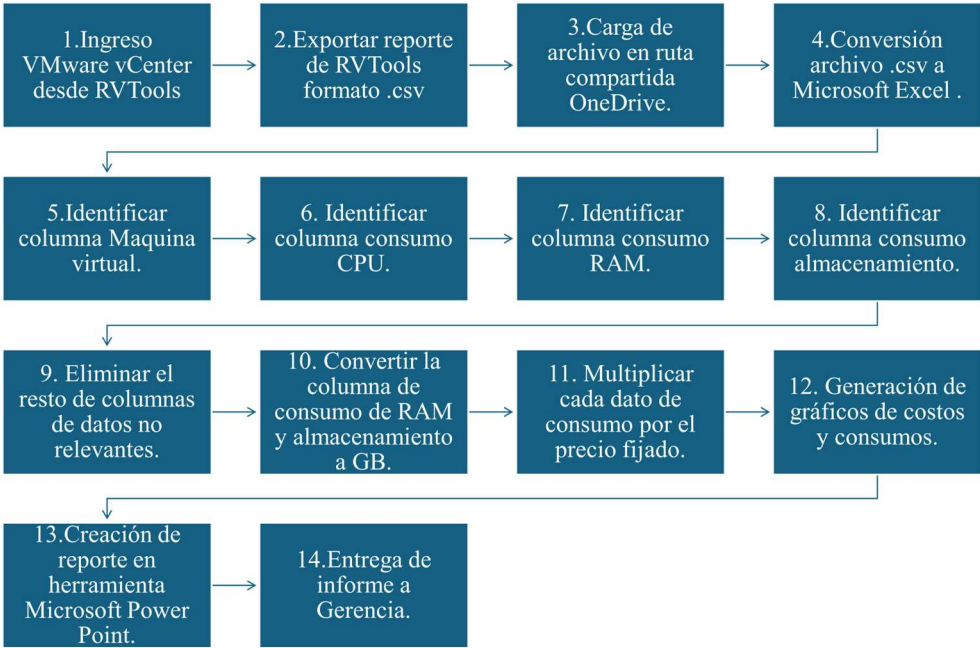
Posteriormente, una vez que el analista complete el proceso de filtrado de los datos relevantes del archivo, deberá identificar los valores correspondientes al consumo de CPU, RAM y almacenamiento y convertir dichos datos a la unidad de medida de GB (Gigabyte) de cada una de las máquinas virtuales. Esta información será utilizada para calcular el costo unitario por máquina, mediante la multiplicación de los consumos registrados por un costo fijo previamente establecido. Con estos datos, se generará un informe que permitirá identificar los costos

individuales por máquina virtual y proceder a la emisión de las facturas correspondientes a cada cliente.

Adicionalmente, esta información será empleada para elaborar un informe gerencial que detalla los consumos de la infraestructura virtualizada. Este informe destacará las máquinas con mayor consumo de recursos, identificándolas como las más costosas, y servirá como insumo clave para la toma de decisiones estratégicas y el análisis preventivo de los recursos empleados en la infraestructura virtual. Dicho informe se elabora a partir de gráficos generados con los datos procesados en Excel, los cuales se integran posteriormente en una presentación corporativa formalizada en un documento de Microsoft PowerPoint la cual será presentada ante la gerencia para la toma de decisiones.

Figura 15

Diagrama de flujo proceso actual



Fuente: Elaboración propia, 2024.

El análisis del flujo de trabajo identificado nos demuestra la dependencia existente de múltiples pasos manuales para la rutina de recolección y presentación de datos. Esta metodología, al estar mayoritariamente sujeta a la ejecución por parte de una persona, incrementa significativamente el riesgo de errores involuntarios. Adicionalmente, se observa que el formato actual de almacenamiento de estos datos no está diseñado para un acceso estructurado y sostenible en el tiempo, convirtiendo esta información en recursos de uso muerto que desaparecen mes tras mes. Este análisis marca la necesidad urgente de optimizar el flujo de trabajo mediante la reducción de procesos manuales y de implementar un sistema de almacenamiento persistente que garantice la disponibilidad y utilidad de los datos a largo plazo, maximizando su valor estratégico para la toma de decisiones.

Durante la observación del proceso se realiza un registro del tiempo requerido para la ejecución del proceso en su totalidad donde un operador de infraestructura y un analista de datos se dedicaron de manera exclusiva a esta actividad; arrojando como resultado un tiempo aproximado de 5.8 horas (348 minutos de forma dedicada). Esta estimación incluye todas las etapas del proceso, desde la recolección de los datos mediante la herramienta de RVTools hasta la elaboración final del reporte con sus correspondientes indicadores. Este cálculo comprende las 14 actividades o pasos necesarios identificados en el flujo de trabajo, cada uno de los cuales contribuye al tiempo total requerido destacando la complejidad y el esfuerzo necesario para completar cada fase de manera manual.

Tabla 6 *Cálculo de tiempo por actividad*

ACTIVIDAD	TIEMPO EN MINUTOS
Ingreso VMware vCenter desde RVTools.	10
Exportar reporte de RVTools formato .csv	5
Carga de archivo en ruta compartida OneDrive.	5
Conversión archivo .csv a Microsoft Excel.	10
Identificar columna Máquina virtual.	10
Identificar columna consumo CPU.	10
Identificar columna consumo RAM.	10
Identificar columna consumo almacenamiento.	10
Eliminación datos no relevantes.	60
Convertir la columna de consumo de RAM y almacenamiento a GB.	8
Multiplicar cada dato de consumo por el precio fijado.	10
Generación de gráficos de costos y consumos.	60
Creación de reporte en herramienta Microsoft Power Point.	130
Entrega de informe a Gerencia.	10
Total	348

Fuente: Elaboración propia, 2024.

4.1.2 Diagnóstico técnico.

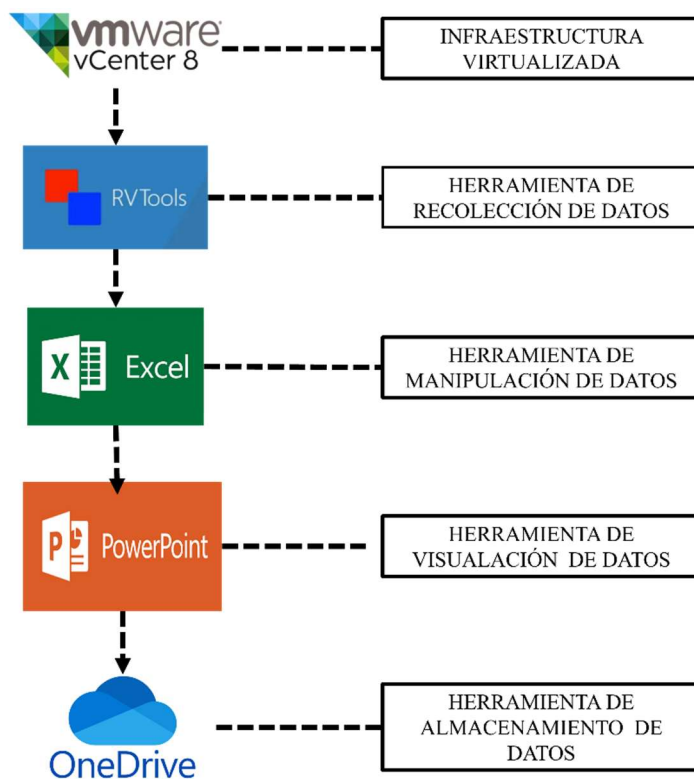
El presente diagnóstico técnico tiene como propósito analizar la infraestructura lógica del departamento de tecnologías de la información en Artificial Skills. Este análisis busca identificar las herramientas principales utilizadas en el proceso de gestión de datos de las máquinas virtuales de la plataforma de VMware vCenter; así como evaluar su impacto en la eficiencia operativa y determinar las brechas técnicas existentes en la actualidad. Los hallazgos permitirán establecer una idea clara para comprender las limitaciones técnicas y fundamentar las recomendaciones necesarias para optimizar el sistema de gestión de datos de la infraestructura virtualizada.

En la actualidad las diferentes actividades involucradas en el proceso son realizadas manualmente, donde los diferentes participantes intervienen con una serie de tecnologías que implican una considerable inversión de tiempo y esfuerzo. Los colaboradores dependen de una serie de tecnologías como Microsoft Excel para el procesamiento de datos y PowerPoint para la presentación de informes. Sin embargo, estas herramientas a pesar de que han sido útiles hasta la actualidad; presentan limitaciones significativas en cuanto a una eficiencia operativa real; adicionalmente la manipulación manual de los datos incrementa el riesgo de errores, como inconsistencias en la información o pérdidas de datos críticos, lo que puede derivar en impactos negativos sobre la confiabilidad de los informes generados al trabajar sobre archivos planos.

El segundo hallazgo de suma importancia del proceso actual es que no existe un sistema estructurado de almacenamiento que permita la consulta y análisis histórico de los datos de los diferentes informes generados a lo largo del tiempo. Actualmente Ask utiliza la nube de OneDrive para el almacenamiento de los diferentes archivos, desde el archivo de Excel, hasta el informe final de PowerPoint, lo que genera un uso limitado y temporal de la información recopilada. Este sistema no solo incrementa la probabilidad de errores humanos ya que todos los involucrados tienen acceso a estos archivos, sino que también restringe la capacidad de la organización para responder de manera ágil a las necesidades operativas y gerenciales al perder la información generada mes tras mes.

Figura 16

Herramientas tecnológicas utilizadas.



Fuente: Elaboración propia, 2024.

Cabe destacar que Ask dispone de un servidor virtual de desarrollo, el cual opera con el sistema operativo Microsoft Windows Server 2022. Este servidor cuenta con una configuración técnica, que incluye 4 núcleos de CPU, 24 GB de memoria RAM y 500 GB de almacenamiento en disco. Estas especificaciones técnicas permiten al servidor ofrecer un rendimiento estable y eficiente, diseñado especialmente para desarrollo, prueba e implementación de aplicaciones o soluciones tecnológicas. Este servidor tiene la capacidad de ejecutar procesos complejos y el mismo puede soportar la solución propuesta en este proyecto.

4.1.3 Diagnóstico de percepción.

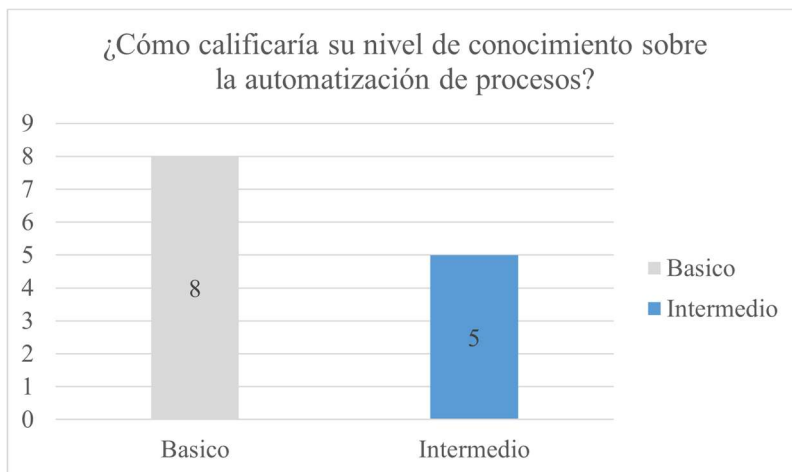
En el siguiente diagnóstico de percepción se presentan y analizan los resultados obtenidos a través del cuestionario aplicado a 13 colaboradores de Artificial Skills; mismo que incluía 11 preguntas tanto de tipo abiertas como cerradas, centrandose su fin en recopilar información clave sobre su percepción respecto al proceso actual de gestión de datos de la infraestructura virtualizada en VMware vCenter. A través de esta herramienta, se identificaron factores en común y opiniones relacionadas con el nivel de conocimiento sobre la automatización de procesos, las herramientas utilizadas actualmente, los tiempos empleados en actividades específicas y las expectativas hacia una posible propuesta de solución.

Los datos recopilados a través del cuestionario han sido analizados y representados mediante gráficos de barras con el objetivo de facilitar su interpretación; en el caso de las preguntas de carácter abierto, se realizó un análisis cualitativo para identificar factores en común entre las opiniones de los 13 colaboradores encuestados. La intención principal es identificar las principales áreas de mejora, los desafíos operativos más destacados y las perspectivas sobre la aceptación de las herramientas propuestas. El análisis obtenido permite obtener una perspectiva clara sobre las percepciones del colaborador, señalando sus inquietudes, expectativas y aún más importante sus necesidades.

Análisis de la percepción sobre el nivel de conocimiento en automatización de procesos.

Figura 17

Pregunta 1



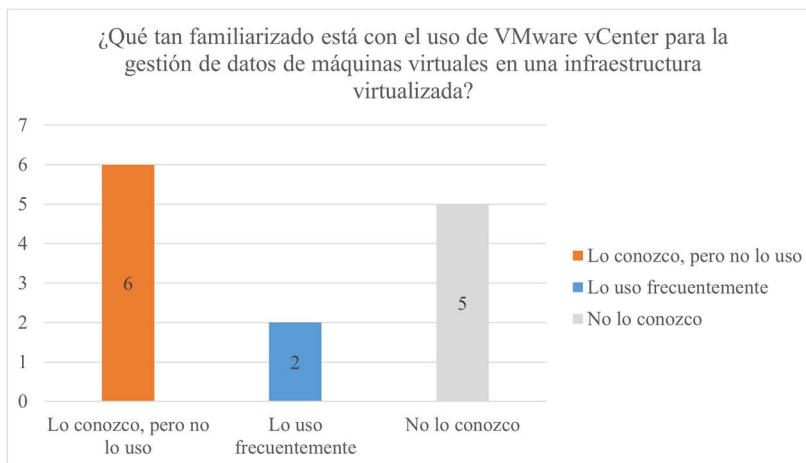
Fuente: Elaboración propia, 2024.

Según los resultados obtenidos sobre la primera pregunta (ver figura 16 – Pregunta 1) que tenía como objetivo conocer el nivel de conocimiento de los colaboradores sobre la automatización de procesos, el 61.5% de los colaboradores (8 encuestados de 13) calificaron su nivel de conocimiento en automatización de procesos como “básico”. El 38.5% (5 colaboradores) indicaron tener un nivel intermedio, mientras que ningún participante se identificó con un nivel avanzado. Estos datos reflejan que existe una brecha significativa en el conocimiento técnico sobre la automatización de procesos dentro del equipo de ASK.

Análisis de la percepción sobre el conocimiento del uso de VMware vCenter

Figura 18

Pregunta 2



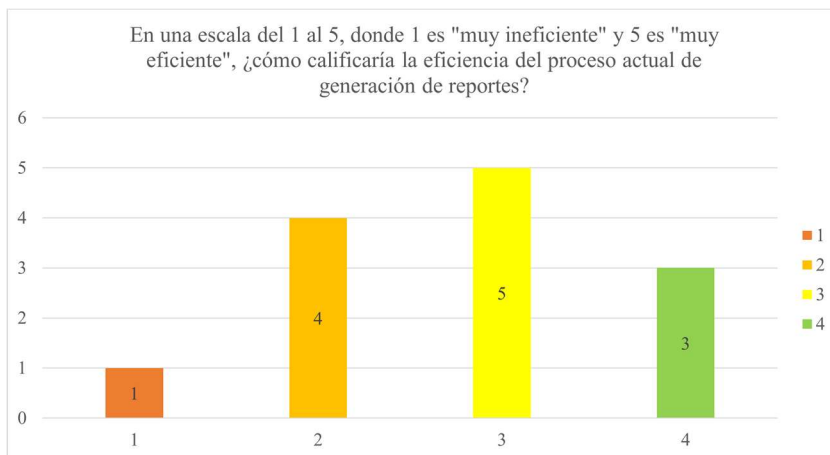
Fuente: *Elaboración propia, 2024.*

La segunda pregunta realizada al personal de Ask (ver figura 17 – Pregunta 2) pretendía evaluar el grado de experiencia del personal técnico con la plataforma VMware vCenter; donde los datos obtenidos señalan una distribución heterogénea en el nivel de familiaridad de los colaboradores con la plataforma VMware vCenter para la gestión de datos en una infraestructura virtualizada, un 46.2% (6 encuestados de 13) afirmaron conocer la herramienta, pero no utilizarla en su flujo de trabajo habitual, el 38.4% de los encuestados indican no tener ningún conocimiento sobre la herramienta y finalmente únicamente el 15.4% (2 colaboradores) reportaron usar VMware vCenter de forma frecuente, estos datos tan distribuidos nos muestran que existe una brecha significativa sobre la plataforma de VMware, ya que no todos los involucrados poseen el mismo nivel de conocimiento.

Análisis de la percepción sobre la eficiencia del proceso actual de generación de reportes

Figura 19

Pregunta 3



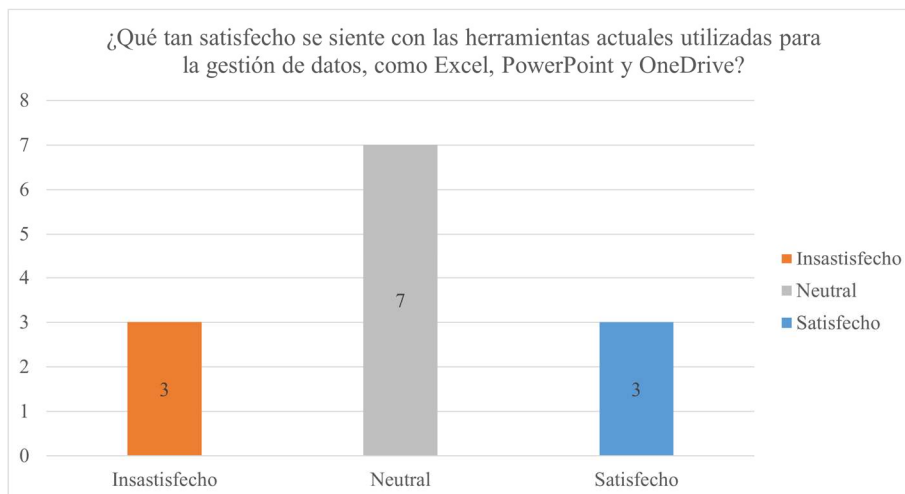
Fuente: Elaboración propia, 2024.

En la tercera pregunta del cuestionario (ver figura 18 – Pregunta 3) se buscaba cuantificar la percepción del personal sobre la eficacia del proceso actual en una escala del 1 al 5, donde 1 representa "muy ineficiente" y 5 "muy eficiente". Los resultados muestran principalmente que un total de 4 colaboradores (30.8%) lo valoraron con un 2 reflejando una percepción de eficiencia bastante baja; en su mayoría 5 colaboradores (38.5%) lo evaluaron con un 3, lo que indica una percepción intermedia, sin una inclinación hacia la eficiencia o ineficiencia; por último, únicamente 3 colaboradores (23.1%) otorgaron una calificación de 4, mostrando una opinión relativamente favorable. Cabe destacar que no se registraron calificaciones de 5, lo que sugiere que ningún encuestado percibe el proceso actual como "muy eficiente".

Análisis de percepción sobre la satisfacción con las herramientas actuales utilizadas para la gestión de datos.

Figura 20

Pregunta 4



Fuente: Elaboración propia, 2024.

La cuarta pregunta (ver figura 19 – Pregunta 4) apuntaba a evaluar el nivel de satisfacción del personal con las herramientas existentes y su efectividad para la gestión de datos. Donde el 23.1% de los encuestados se sienten insatisfechos con las herramientas actuales, señalando posibles limitaciones o inconvenientes en su uso, el 53.8% manifestó una percepción neutral, lo que podría indicar que aunque las herramientas cumplen con ciertas necesidades no destacan en términos de funcionalidad, facilidad de uso o eficiencia; finalmente el 23.1% indicó estar satisfechos con estas herramientas, lo que refleja que para un segmento reducido (únicamente 3 colaboradores de 13) las herramientas actuales cumplen con sus expectativas.

Análisis principales dificultades en el proceso actual de gestión de datos en VMware vCenter.

La quinta pregunta (ver tabla 7 – Listado de respuestas pregunta 5) de carácter abierta buscaba identificar los puntos críticos y áreas problemáticas en el flujo operativo actual desde su punto

de vista, donde se identifican varios factores que representan desafíos significativos en el proceso actual de gestión de datos. Analizando los factores en común de las respuestas podríamos dividir los resultados en 4 importantes aspectos.

1. Lentitud y tiempo prolongado en el proceso:

- a. Los colaboradores mencionan reiteradamente la lentitud al momento de recoger, filtrar y generar reportes.
- b. Otra preocupación de los encuestados del proceso manual reside en que para transformar y manipular los datos extraídos requiere una considerable inversión de tiempo.

2. Dependencia de múltiples áreas:

- a. Algunos de los colaboradores encuestados identifican la dependencia de otras áreas como un obstáculo, lo que consideran que genera demoras y reduce la eficiencia del flujo operativo.

3. Complejidad técnica y herramientas limitadas:

- a. Varios colaboradores destacan desafíos en el manejo de Excel, desde la falta de conocimiento sobre términos específicos; hasta la dificultad en la creación de reportes y gráficos personalizados.

4. Manejo del gran volumen de datos:

- a. Desde el punto de vista de los encuestados el gran volumen de datos generados por VMware vCenter es percibido como un desafío; complicando la recolección, filtrado y análisis de estos.
- b. La manipulación en la generación de un reporte por diferentes personas es percibida por los encuestados como una causa de confusión y posibles errores.

Tabla 7 Listado de respuestas pregunta 5

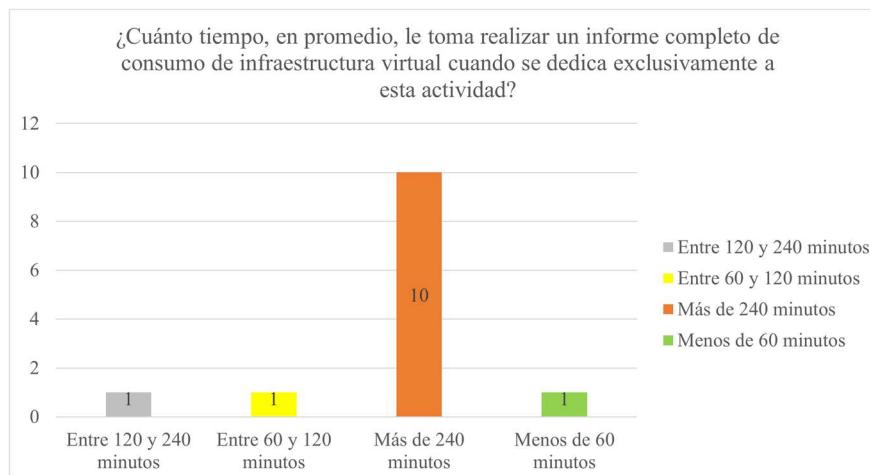
ID	Desde su perspectiva, ¿cuáles son las principales dificultades en el proceso actual de gestión de datos en VMware vCenter?
1	Transformación de data
2	Generación de reportería con la herramienta excel ya que del archivo general extraído existen muchos términos que desconozco
3	Lentitud a la hora de recoger los datos.
4	Creación de reporte basado en necesidades específicas
5	Las herramientas
6	El proceso de recolección y filtrado de datos el cual me toma bastante tiempo
7	La cantidad de datos generados por la plataforma.
8	Existe una confusión muy grande derivado a que un sólo reporte es manipulado por diferentes personas debido a las cargas de trabajo.
9	Generar los gráficos de los reportes en un proceso que toma mucho tiempo debido a que se me complica trabajar en Excel.
10	Lentitud derivada a la dependencia de otras áreas.
11	La dependencia.
12	La dependencia de diferentes áreas y la pérdida de información.
13	Lentitud a la hora de filtrar los datos debido a la gran cantidad de estos.

Fuente: Elaboración propia, 2024.

Análisis de percepción tiempo promedio dedicado a la elaboración de informes de consumo de infraestructura virtual.

Figura 21

Pregunta 6



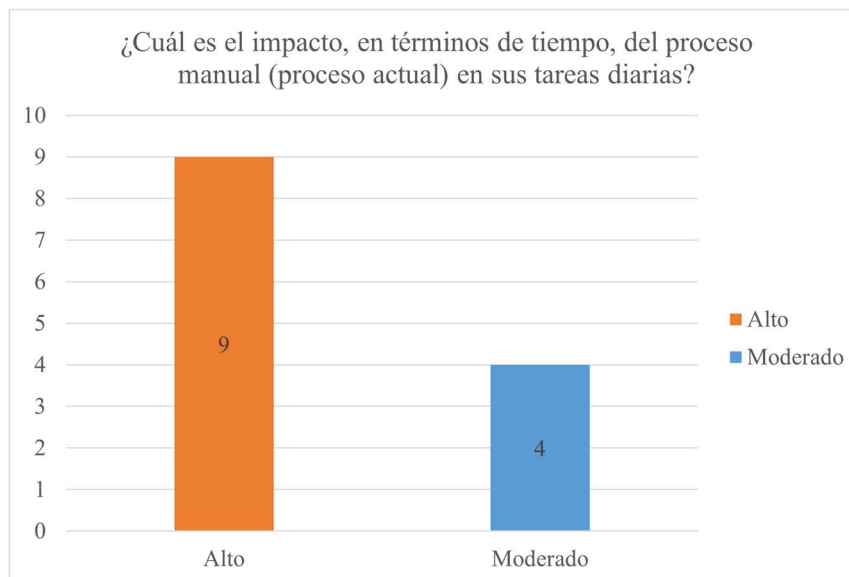
Fuente: Elaboración propia, 2024.

La intención principal de la pregunta número 6 (ver figura 20 – Pregunta 6) pretendía obtener un rango de tiempo del tiempo que tardan los colaboradores en la ejecución del proceso actual. Según las respuestas recopiladas, el tiempo que los colaboradores dedican a elaborar un informe completo varía significativamente. Sin embargo, se observa una inclinación hacia tiempos prolongados; la mayoría de los colaboradores (10 de 13) indicó que les toma más de 240 minutos realizar un informe completo cuando se dedican exclusivamente a esta actividad, representando un 77% de los encuestados, lo que evidencia que el proceso actual es altamente demandante en términos de tiempo. Únicamente tres colaboradores reportaron tiempos menores a este umbral, lo que podría estar relacionado con su rol específico o con una mayor experiencia del colaborador.

Análisis de percepción impacto operativo del proceso actual en términos de tiempo.

Figura 22

Pregunta 7



Fuente: Elaboración propia, 2024.

La pregunta 7 (ver figura 21 – Pregunta 7) se encuentra relacionada con la anterior, ya que en esta ocasión se pretendía evaluar el impacto operativo del proceso manual en la productividad del personal de ASK. Se identifica un impacto significativo del proceso manual en las tareas diarias de los colaboradores; el 69% de los encuestados indicaron que el impacto del proceso manual es alto, demostrando que esta actividad interfiere de manera considerable en sus rutinas diarias. Únicamente 4 colaboradores; el 31% calificaron el impacto como moderado indicando una menor interferencia en comparación con el grupo mayoritario, pero a la vez reflejan que el proceso sigue siendo una carga significativa para ellos. De igual manera cabe destacar que para ningún colaborador estas actividades son de impacto bajo en sus operativa diaria.

Análisis de percepción pérdida de información.

Con el desarrollo de la octava pregunta (ver tabla 8 – Listado de respuestas pregunta 8) se buscaba recopilar información de los colaboradores respecto a la pérdida de información y su impacto en las operaciones actuales. Los resultados de esta pregunta muestran un factor en común claro; 12 de los 13 colaboradores han experimentado problemas relacionados con la pérdida de información, mientras que solo uno indicó no haber enfrentado esta dificultad. Entre las respuestas destacan situaciones como la necesidad de repetir informes eliminados, la falta de precisión en los reportes debido a información perdida, y la creación de trabajos duplicados por la ausencia de un almacenamiento centralizado.

Se señala que los procesos manuales actuales agravan este problema al no garantizar una trazabilidad adecuada de los datos. Algunos colaboradores mencionaron que la pérdida de información afecta no solo la operatividad interna, sino también la relación con los clientes, al generar retrasos significativos en la entrega de reportes.

Tabla 8 *Listado de respuestas pregunta 8.*

ID	¿Ha experimentado problemas relacionados con la pérdida de información debido a la falta de una herramienta centralizada para el almacenamiento y gestión de datos?
1	Si, debido a la perdida de información hay vacíos en tiempos de recolección de datos
2	Si en ocasiones se debe repetir ciertos informes ya que los mismos fueron eliminados y en cierto momento se vuelven a requerir
3	Sí, la información no se encuentra centralizada y los procesos que actualmente se ejecutan es de forma muy manual, lo que ha dejado espacio a varios problemas a la hora de manipular la data.
4	Si. La falta de almacén de datos hace que mis reportes no sean lo más preciso posible. Además de presentar información que puede ser de mucha relevancia para la vida de un equipo o servidor.
5	Perdida de datos
6	No
7	Si ya que los informes desaparecen con el tiempo y en ocasiones al ser solicitados de nuevo nos toca recrearlos
8	Si, perdida de información
9	Si perdida de información

10	Realización de un doble trabajo por pérdida de información
11	Si, pérdida de información
12	Si existe una gran pérdida de reportes a lo largo del tiempo que genera un doble trabajo
13	Si pérdida de información y problemas en la entrega de informes con el cliente debido a la pérdida de estos.

Fuente: Elaboración propia, 2024.

Análisis de percepción sugerencias de mejora.

Seguidamente con la siguiente pregunta (ver tabla 9 – Listado de respuestas pregunta 9) se pretendía identificar las sugerencias de los colaboradores sobre posibles mejoras en la gestión de datos para optimizar tiempos y reducir esfuerzos manuales. Las respuestas recopiladas muestran un consenso claro entre los colaboradores entorno a la necesidad de automatizar procesos y mejorar las herramientas actuales. La mayoría menciona la automatización o herramientas "automáticas" para abordar las ineficiencias operativas y reducir los tiempos invertidos en tareas manuales actuales. Algunas respuestas adicionales sugieren implementar otro tipo de mejoras, tales como:

- Herramientas para la extracción de datos simplificada.
- Filtrado automatizado y centralización de la información.
- Reducción de personal involucrado.
- Mejoras en el análisis y presentación de datos.

Tabla 9 Listado de respuestas pregunta 9.

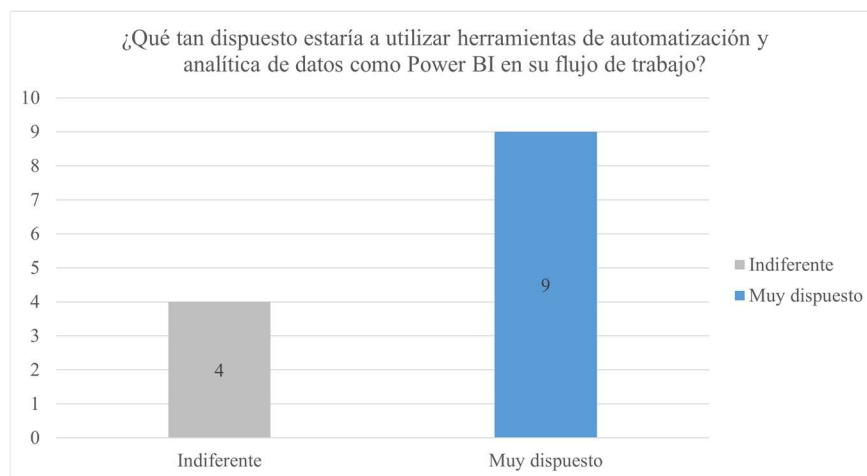
ID	¿Qué mejoras considera que podrían implementarse en la gestión actual para reducir tiempos operativos?
1	Automatización.
2	Algún procedimiento para la extracción de datos que sea más simplificado para nosotros.
3	Automatizar los procesos.
4	Automatización de procesos manuales.
5	Aplicas nuevas herramientas.
6	Alguna herramienta que filtre de manera automática los datos provenientes de VMware.
7	Algún método más formal para el almacenamiento de los reportes y también una herramienta que facilite la extracción de datos, pero con un nivel mayor de filtro.
8	Una herramienta que permita reducir la cantidad de personal que participa en la generación de los informes.
9	Herramientas más ágiles.
10	Mejores herramientas para el análisis de datos.
11	Una centralización de la información.
12	Una forma de almacenar los reportes de tal forma que sean más duraderos a lo largo del tiempo.
13	Una forma automática de extraer datos ya filtrados desde VMware.

Fuente: Elaboración propia, 2024.

Análisis de percepción nivel de aceptación a la incorporación de herramientas de automatización al proceso de gestión de datos.

Figura 23

Pregunta 10



Fuente: Elaboración propia, 2024.

Posteriormente con la pregunta número 9 (ver figura 22 – Pregunta 10) se procuraba medir el nivel de aceptación potencial de herramientas automatizadas por parte del personal. Donde una mayoría del 69% de los encuestados se muestran muy dispuestos a utilizar herramientas de automatización y analítica de datos como Power BI. El restante 31% se declararon indiferentes, indicando una neutralidad, desconocimiento o falta de experiencia previa con estas tecnologías y finalmente ningún colaborador indico no estar de acuerdo con la posibilidad de herramientas de automatización, indicando a nivel general un nivel de aceptación muy alto.

Análisis de percepción expectativas de una posible solución automatizada para el proceso de gestión de datos.

Por último, con la pregunta final (ver Tabla 10 – Listado de respuestas pregunta 11) se buscaba entender qué resultados o mejoras esperan los colaboradores en caso de que existiera la posibilidad de implementar una herramienta de automatización para la gestión de datos de las

máquinas virtuales. Los resultados de esta pregunta muestran un claro interés por la optimización de tiempos y el aumento de la eficiencia operativa. La mayoría de las respuestas destacan una expectativa de reducción significativa en el tiempo requerido para realizar tareas diarias, lo cual refleja una necesidad de agilizar el proceso de generación de reportes y análisis. Las respuestas también sugieren un deseo de mayor autonomía en las labores, reduciendo la dependencia de otras áreas y permitiendo un enfoque más profundo en el análisis de datos.

Tabla 10 *Listado respuestas pregunta 11.*

ID	Si existiera la posibilidad de implementar una solución automatizada para la gestión de los datos de las máquinas virtuales ¿Cuáles serían sus expectativas sobre la solución automatizada?
1	Generación de reportaría por chat, análisis y entendimiento completo de los datos en base.
2	Disminuir el tiempo que toma la generación de un solo reporte
3	Disminución en el tiempo de ejecución de tareas diarias.
4	Que pueda ser predictivo.
5	Reducción de tiempo
6	Reducción de tiempo que mejore mi eficiencia operativa
7	Una reducción de tiempo en mis labores diarias
8	Algo que me permita dedicarle mayor tiempo al análisis de los datos
9	Reducción de tiempo operativo
10	Reducción de tiempos operativos y reducción de dependía de áreas
11	Menos dependencia de otras áreas
12	Menos dependencia
13	Mayor rapidez y eficiencia operativa

Fuente: Elaboración propia, 2024.

4.1.4 Determinación de brechas.

En la siguiente sección se expondrán las brechas y conclusiones encontradas durante la realización del diagnóstico de situación actual expuesto en el capítulo 4 del presente proyecto.

Gracias a este análisis, se busca destacar los principales hallazgos y oportunidades de mejora

detectados que nos servirán como una guía de comparación respecto al modelo futuro de automatización en el que se basa esta investigación.

Brechas operativas.

Dentro del análisis realizado sobre el diagnóstico operativo del proceso de gestión de datos de la infraestructura virtual actualmente se logra destacar una serie de brechas operativas relevantes las cuales marcan las principales debilidades a nivel operativo del proceso actual con respecto a lo planteado con el desarrollo de este proyecto. Estas brechas no solo impactan negativamente la eficiencia operativa, sino que a la vez limita la capacidad de ASK para tomar decisiones estratégicas basadas en datos confiables y oportunos.

1. Ausencia de un proceso automatizado.

Durante el análisis realizado se encuentra que actualmente la gestión realizada sobre los datos extraídos de la infraestructura virtual es un proceso de suma intervención manual, en la cual en la mayoría de las etapas del flujo de trabajo se necesita la intervención manual tanto de un operador de infraestructura como de un analista de datos. Esto ocasiona que el tiempo invertido por cada colaborador tenga un gran impacto sobre su operativa diaria; alcanzando incluso tiempos de 5,8 horas para la realización de un reporte de consumo, resaltando que este tiempo implica la dedicación exclusiva de este colaborador escenario que es poco común ya que los diferentes operadores y analistas en su operativa normal no se encuentran dedicados a esta actividad.

Además del tiempo requerido, que según el 69% de los encuestados tiene un impacto operativo alto y afecta significativamente sus tareas diarias, la alta dependencia de la intervención manual de múltiples colaboradores incrementa el riesgo de errores en la recopilación y procesamiento

de datos, lo cual puede comprometer la precisión de los informes. En respuesta a este problema la propuesta de automatización planteada en este proyecto busca reducir significativamente la intervención manual durante el proceso, optimizando los tiempos operativos y minimizando el riesgo de errores, lo que mejorará la eficacia y eficiencia general del proceso.

2. Dependencia de múltiples áreas y colaboradores.

Durante el proceso operativo actual se determina que el mismo depende de la participación de múltiples áreas y colaboradores de Ask, para completar cada una de las etapas del flujo de trabajo; desde la recolección de datos hasta la transformación de los datos en un reporte final. Esta dependencia genera retrasos operativos significativos más, aun considerando las cargas de trabajo de los colaboradores que dificultan su dedicación exclusiva a estas tareas. Adicionalmente, la falta de autonomía en el proceso puede ocasionar confusiones y duplicación de esfuerzos, como señalaron varios encuestados en el diagnóstico de percepción.

Este problema no solo retrasa la ejecución del proceso, sino que incrementa el riesgo de omitir información crítica, afectando la confiabilidad de los reportes generados. La solución propuesta en este proyecto pretende realizar una centralización de las diferentes tareas clave y proponer la implementación de herramientas que permitan reducir la dependencia entre áreas, promoviendo un flujo de trabajo más ágil.

3. Ausencia de un almacenamiento de información estructurado.

Un hallazgo de suma importancia para el desarrollo del presente proyecto fue la ausencia de una base de conocimiento estructurada que permita almacenar la información generada por la infraestructura virtual; enfocándonos en el aspecto operativo, actualmente la información se almacena temporalmente en plataformas como OneDrive, lo que incrementa la probabilidad de

pérdida de datos y dificulta el acceso a registros pasados. Un 92% de los encuestados reportaron haber experimentado problemas relacionados con la pérdida de información, lo cual genera la necesidad de repetir informes eliminados o de realizar trabajos de forma duplicada. Este problema no solo afecta la operativa interna, sino también la relación con los clientes, debido a retrasos en la entrega de reportes.

La propuesta de implementación de una solución de almacenamiento centralizado busca permitirle a Ask mantener un registro histórico de los datos, garantizando su trazabilidad y accesibilidad para solicitudes futuras, buscando reducir la duplicación de esfuerzos y mejorar significativamente la continuidad operativa.

4. Desconocimiento técnico.

Gracias al diagnóstico de percepción realizado se muestra que el 61.5% de los colaboradores califican su nivel de conocimiento en automatización como “básico” y el 38.5% tienen un conocimiento intermedio. Esto refleja una brecha técnica importante que limita la capacidad del equipo para implementar mejoras tecnológicas y optimizar los procesos actuales (incluso ningún colaborador en la actualidad registra un nivel de conocimiento “alto”); debido a esto es que no existe un proceso automatizado en la actualidad debido a que los involucrados por razones de desconocimiento sobre automatización y desconfianza prefieren seguir realizando el proceso de forma manual.

Adicional, la falta de familiaridad con herramientas clave como VMware vCenter también se destacó como un problema, ya que solo el 15.4% de los colaboradores reportaron usarla de forma frecuente. Esta disparidad en el conocimiento técnico dificulta la estandarización de procesos y aumenta la dependencia de unos pocos expertos, especialmente en la etapa de filtrado de datos, donde los analistas cuando se encuentran trabajando con los datos recolectados

mediante la herramienta RVTools manejan una gran cantidad de datos a los cuales no se encuentran familiarizados complicándoles de forma significativa este proceso de filtrado de datos y arriesgándolos a cometer un error involuntario debido a la falta de conocimiento.

Como parte de este proyecto, se propondrá la construcción de una base de conocimiento estructurada que reduzca significativamente la interacción manual entre los analistas y el filtrado de datos. La idea es que los datos sean recolectados de manera automática mediante el uso de una API integrada, lo que permitirá agilizar el proceso, minimizar errores y liberar a los analistas para tareas más estratégicas.

Brechas técnicas.

Adicionalmente de las brechas operativas identificadas; gracias al diagnóstico técnico y de percepción se identifican una serie de brechas técnicas derivadas de las diferentes herramientas y tecnologías involucradas en el proceso actual. Estas herramientas si bien es cierto actualmente son útiles para lograr el objetivo final de generación de reportes, más sin embargo estas no son las más eficientes para un adecuado proceso de gestión de los datos de las máquinas virtuales en una infraestructura virtual. Estas brechas técnicas representan una serie de atrasos operativos y de riesgos que actualmente afectan negativamente el proceso; se exponen a continuación dichas brechas técnicas encontradas.

1. Uso de herramientas limitas.

El diagnóstico técnico evidencia que las herramientas utilizadas actualmente, como RVTools para la extracción de datos, Excel para el procesamiento de datos y PowerPoint para la presentación de informes son funcionales tal como se mencionó anteriormente, sin embargo, presentan limitaciones significativas para manejar el gran volumen de los datos generados por

la plataforma de VMware vCenter. Estas herramientas no están diseñadas para soportar procesos de análisis avanzados o automatización, lo que genera una mayor dependencia del trabajo manual de los colaboradores; tanto para la recopilación de datos como para la transformación de dichos datos en data útil para la toma de decisiones.

Esta problemática nuevamente al igual que las anteriores incrementa la probabilidad de errores involuntarios y de generar inconsistencias en la información o peor aún pérdida de datos críticos que afecten negativamente la confiabilidad de los informes y por ende la operativa y credibilidad de la empresa. Para abordar esta brecha, el proyecto propone la implementación de herramientas avanzadas como APIs automatizadas y plataformas de análisis de datos como PowerBi que le permitan a Ask gestionar los grandes volúmenes de datos generados por la plataforma de VMware vCenter reduciendo de forma significativa los tiempos invertidos por el personal en la generación de los informes y convirtiendo esta información mucho más confiable.

2. Ausencia de un almacenamiento de información estructurado.

Al igual que la brecha identificada en el apartado de brechas operativas, existe una limitante técnica relacionada a la ausencia de un almacenamiento de información centralizado en el cual se pueda guardar de forma recurrente los datos de la infraestructura, esta limitante técnica nace debido a que, en la actualidad, los datos generados por los procesos se almacenan temporalmente en plataformas como OneDrive; sistema que no garantiza la trazabilidad, el análisis histórico y la disponibilidad inmediata de los datos. Esta brecha técnica dificulta de gran manera la continuidad operativa y aumenta el riesgo de pérdida de información clave; sumado a que a debido a la ausencia de un sistema robusto de almacenamiento impide la integración de datos para análisis mucho más avanzados.

Adicionalmente en base al análisis de percepción la mayoría de los encuestado (el 53.8%) manifestó una percepción neutral con respecto a su nivel de satisfacción con el uso de las herramientas actuales involucradas en el proceso, lo que indica que, aunque las herramientas cumplen con ciertas necesidades no destacan en términos de funcionalidad, facilidad de uso o eficiencia, desde el punto de vista de los colaboradores. Lo que se espera con la iniciativa de este proyecto es proponer la creación de una base de conocimiento centralizada y estructurada que permita un almacenamiento seguro, accesible y escalable para los datos generados en la infraestructura virtual; que adicionalmente permita la integración con herramientas de inteligencia artificial para evitar esfuerzos dobles y la pérdida de información crítica.

3. Infraestructura subutilizada.

La siguiente brecha técnica identifica está relacionada con la subutilización de la infraestructura que posee actualmente Ask la cual posee un servidor virtual de desarrollo con especificaciones técnicas adecuadas (4 núcleos de CPU, 24 GB de RAM y 500 GB de almacenamiento), esta infraestructura actualmente no está siendo aprovechada en su totalidad, ya que la misma no se utiliza para soportar soluciones automatizadas ni para un adecuado almacenamiento y gestión de los datos.

Esta brecha representa una oportunidad desaprovechada para mejorar la eficiencia y optimización del sistema actual, el proyecto propone aprovechar este servidor para implementar soluciones automatizadas, con la integración de APIs y herramientas de analítica de datos, maximizando la utilización del servidor de forma que le ayude a Ask a sacar un verdadero provecho de su infraestructura al mismo tiempo que pueda contar con herramientas adecuadas que le permita mejorar el proceso de gestión de datos de las máquinas virtuales .

CAPÍTULO V: DISEÑO Y DESARROLLO DEL PROYECTO

5.1 DESARROLLO DE LA PROPUESTA

En el presente capítulo número V de este proyecto, se pretende desarrollar la propuesta final del proyecto la cual se enfoca en optimizar los procesos de gestión de los datos de las máquinas virtuales en la plataforma de VMware vCenter; mediante la implementación de una base de conocimiento estructurada integrada con herramientas de analíticas de datos e inteligencia artificial. El desarrollo de esta propuesta se realizará basándose en los objetivos específicos del proyecto y en el análisis realizado en el capítulo IV de este trabajo de investigación, con el fin de desarrollar una propuesta que se fundamente en base a las necesidades y oportunidades de mejoras reales de la compañía Artificial Skills.

5.1.1 Contexto de la situación actual.

Como se mencionó anteriormente el desarrollo de esta propuesta pretende ir alineada respecto a las necesidades y oportunidades de mejora identificadas en el capítulo IV del presente proyecto. Gracias a este análisis podemos identificar que una de las necesidades más notables del proceso actual de gestión de datos de la infraestructura virtual en Artificial Skills es la amplia cantidad de procesos manuales que se ejecutan diariamente; lo cual se traduce en atrasos operativos significativos y un gran desgaste del personal involucrado. Esta situación ocasiona que a Ask se enfrente a una serie de riesgos relevantes de incurrir en errores humanos derivados de la alta intervención de diferentes áreas y colaboradores durante el proceso, afectando negativamente la calidad de los reportes de la infraestructura virtual (ver figura 14 - Diagrama de flujo proceso actual).

Adicionalmente otro de los hallazgos importantes durante el diagnóstico de la situación actual está directamente relacionado con el procesamiento completo de los datos generados por la

plataforma de VMware vCenter, donde los colaboradores actualmente presentan una gran dolencia en el proceso de recolección de los datos, ocasionado por el uso de las herramientas altamente limitadas para esta necesidad (ver figura 15 - Herramientas tecnológicas utilizadas). Esto sumado a que la información recolectada durante el proceso no está siendo almacenada adecuadamente ya que la misma al momento de utilizar herramientas pocas prácticas para dicho fin, se convierte en información de un solo uso no perdurable en el tiempo, ocasionando pérdida de información crítica y dobles esfuerzos operativos.

Mencionado lo anterior, la necesidad de implementar una solución que aborde estas problemáticas de manera integral es de suma importancia para Artificial Skills; por lo cual se formulará una propuesta de automatización que busca reducir significativamente la intervención manual en los procesos operativos, optimizar el uso de las herramientas tecnológicas y garantizar el almacenamiento estructurado de la información, adicionado a esto que dicha solución pueda integrarse con herramientas de inteligencia artificial tanto para la automatización de procesos como para un aprovechamiento máximo de los datos generados. Seguidamente, se analizarán las herramientas más idóneas para el diseño de esta propuesta, asegurando su alineación con los objetivos específicos planteados en este proyecto.

5.1.2 Actores.

A continuación, en el siguiente apartado, se describen los actores principales involucrados en el sistema; en la tabla adjunta se detalla el código, nombre y descripción del rol del actor dentro del proceso automatizado.

Tabla 11 *Tabla de actores del sistema.*

Código de autor	Nombre del autor	Descripción
AC-1	Gerente de TI	Responsable de la supervisión y toma de decisiones estratégicas basadas en la información proporcionada por el sistema.
AC-2	Analista de datos	Encargado de generar reportes automatizados y analizar tendencias y consumos de la infraestructura virtualizada.
AC-3	Administrador de infraestructura	Responsable de la gestión de los recursos virtualizados y la identificación de anomalías en el sistema.
AC-4	Usuario	Persona que consulta información sobre la infraestructura a través del asistente virtual para la toma de decisiones operativas.

Fuente: Elaboración propia, 2024.

5.1.3 Historias de usuario.

En este apartado se exponen las historias de usuario, las cuales fueron obtenidas gracias al diagnóstico de la situación actual y de la retroalimentación del diagnóstico de percepción realizado. A continuación, se expone la tabla con cada una de las correspondientes historias de usuario.

Tabla 12 *Historias de usuario.*

ID	Actividad	Actor	Funcionalidad	Criterios de aceptación
HU-1	Autenticación	AC-1	El sistema debe validar el acceso del usuario mediante el sistema de autenticación seguro vinculado a VMware vCenter.	El usuario debe ser autenticado con credenciales válidas.
HU-2	Integración con PyVmomi	AC-4	El sistema debe utilizar la API PyVmomi para extraer información de VMware vCenter en tiempo real.	Los datos deben obtenerse sin errores y en el tiempo esperado.

HU-3	Extracción de datos automatizada	AC-4	El sistema debe ejecutar un script de Python programado para extraer datos periódicamente de VMware vCenter.	El script debe ejecutarse automáticamente en la frecuencia definida y registrar logs de ejecución.
HU-4	Filtrado de datos	AC-3	El sistema debe procesar los datos extraídos y filtrar información relevante para su análisis.	Los datos filtrados deben cumplir con los criterios predefinidos y almacenarse correctamente.
HU-5	Almacenamiento en Base de Conocimiento	AC-2, AC-3	El sistema debe almacenar la información filtrada en una base de conocimiento MySQL.	Los datos deben guardarse correctamente y estar disponibles para consultas posteriores.
HU-6	Análisis de datos	AC-2, AC-3	El sistema debe aplicar reglas de inferencia en la base de conocimiento para analizar tendencias y generar recomendaciones.	Los análisis deben ser precisos y generar reportes claros.
HU-7	Integración con Asistente Virtual	AC-1, AC-4	El sistema debe permitir consultas en lenguaje natural mediante un asistente virtual sobre el estado de los recursos virtualizados.	El asistente debe responder correctamente y en un tiempo razonable.
HU-8	Integración con Power BI	AC-2, AC-3	El sistema debe generar reportes gráficos en Power BI a partir de la información almacenada en la base de conocimiento.	Los reportes deben visualizarse correctamente con datos actualizados en tiempo real.

Fuente: Elaboración propia, 2024.

5.1.4 Requerimientos funcionales.

Los requerimientos funcionales propuestos describen las características y capacidades específicas que debe poseer el proyecto para cumplir con los objetivos plasmados en el capítulo I del presente documento. Estos requisitos establecen cómo debe operar el sistema, qué funcionalidades debe proporcionar a los usuarios y cómo debe interactuar con todos los

componentes tecnológicos sugeridos; dado que el sistema busca optimizar la gestión de la infraestructura virtualizada en VMware vCenter mediante la aplicación de automatización e inteligencia artificial, los requerimientos funcionales incluyen la extracción y análisis de datos, el almacenamiento estructurado de información, la integración con herramientas de visualización y la interacción con los usuarios mediante asistentes virtuales.

A continuación, se detallan los requerimientos funcionales que garantizarán el correcto desempeño y utilidad del sistema.

Tabla 13 *Requerimientos funcionales.*

ID	Requerimiento	Descripción	Prioridad	Clasificación
RF-01	Integración con VMware vCenter	El sistema debe permitir la integración con la plataforma VMware vCenter para la extracción de datos en tiempo real mediante la API de PyVmomi.	Alta	Integración
RF-02	Base de conocimiento	El sistema debe contar con una base de conocimiento que pueda almacenar, organizar y procesar información estructurada para su uso prolongado en la toma de decisiones.	Alta	Almacenamiento
RF-03	Análisis de datos	El sistema debe incluir la integración de inteligencia artificial (API ChatGPT) para el análisis predictivo y automatización de procesos operativos.	Media	integración y Procesamiento
RF-04	Generación de reportes en Power BI	Integración con Power BI para la generación de reportes interactivos sobre la gestión de recursos virtualizados.	Alta	Visualización
RF-05	Acceso centralizado	Proporcionar acceso a los datos centralizados sin la intervención manual del personal de TI.	Media	Seguridad
RF-06	Pantalla de inicio de sesión	El sistema debe contar con una pantalla de inicio de sesión para la autenticación de usuarios.	Alta	Interfaz de usuario
RF-07	Pantalla de inicio	El sistema debe contar con una pantalla de inicio donde se muestren opciones principales de navegación.	Media	Interfaz de usuario
RF-08	Visualización del dashboard de Power BI	El sistema debe permitir la visualización de reportes y métricas en un dashboard de Power BI.	Alta	Visualización
RF-09	Pantalla del asistente virtual	El sistema debe contar con una pantalla donde los usuarios puedan interactuar con el asistente virtual para generar consultas.	Media	Interfaz de usuario

Fuente: Elaboración propia, 2024.

5.1.5 Requerimientos no funcionales.

Los requerimientos no funcionales como se mencionó anteriormente en el marco teórico, a diferencia de los requerimientos funcionales, estos establecen los criterios de calidad, rendimiento y seguridad que debe cumplir el sistema para garantizar su eficiencia y fiabilidad.

En el presente proyecto, los requerimientos no funcionales son fundamentales para asegurar la disponibilidad continua del sistema, la capacidad de escalabilidad, la seguridad de la información y la facilidad de su uso para los distintos tipos de usuarios tanto internos como externos. A continuación, la lista detallada de estos:

Tabla 14 *Requerimientos no funcionales.*

ID	Requerimiento	Descripción	Prioridad	Clasificación
RNF-01	Disponibilidad	El sistema debe garantizar acceso 24/7 a los datos para los usuarios autorizados.	Alta	Fiabilidad
RNF-02	Escalabilidad	La solución debe ser capaz de manejar un incremento progresivo en el volumen de datos sin afectar su rendimiento.	Media	Rendimiento
RNF-03	Seguridad	El sistema deberá tener todos los recursos asegurados, ya sea por medio de autorizadores o roles de acceso.	Alta	Seguridad
RNF-04	Usabilidad	La interfaz debe ser intuitiva y fácil de manejar, permitiendo que los usuarios de distintas áreas puedan interpretar los reportes sin necesidad de conocimientos técnicos avanzados.	Media	Experiencia de Usuario
RNF-05	Mantenimiento	El sistema debe permitir actualizaciones y mejoras sin afectar la operatividad de los procesos automatizados.	Media	Mantenibilidad

Fuente: Elaboración propia, 2024.

5.1.6 Especificación de requerimientos.

En esta sección, se aborda la fase de especificación de requerimientos, un proceso fundamental para el desarrollo exitoso del proyecto. Tras la identificación inicial de las necesidades del sistema, se procede a la elaboración detallada de cada requerimiento. Esta profundización se logra mediante la creación de historias de usuario, donde cada una de ellas desglosa de manera exhaustiva la necesidad específica y la solución propuesta. El objetivo es establecer una base

sólida y precisa que guíe las etapas posteriores del proyecto, asegurando que el producto final cumpla con las expectativas del sistema.

Tabla 15 *Requerimiento RF-01.*

ID del Requerimiento:	RF-01		
Creado por:	José Delgado Arce	Modificado por:	
Fecha creación:	19/03/2025	Ult. Actualización:	
Módulo:	Integración		
Fuentes:	Análisis operativo y diagnóstico de percepción ejecutados		
Actores:	AC-1, AC-2, AC-3		
Objetivo:	Permitir la integración con VMware vCenter para la automatización de la extracción de datos en tiempo real mediante la API PyVmomi.		
Descripción:	El sistema debe conectarse a la plataforma VMware vCenter utilizando la API PyVmomi para obtener información sobre los recursos virtualizados y el estado de la infraestructura en tiempo real y de forma automatizada.		
Importancia/Prioridad:	Alta		
Clasificación:	Integración		
Elementos de entrada de datos:	El sistema debe solicitar la siguiente información:		
	<i>Nombre</i>	<i>Descripción</i>	
	Credencial	Credencial usuario administrador de vCenter (String).	
	IP	Dirección IP del servidor vCenter (String).	
Elementos de resultados de datos:	Información sobre el estado de las máquinas virtuales		
Restricciones y supuestos:	<ul style="list-style-type: none"> El usuario debe contar con permisos adecuados en VMware vCenter para acceder a la API. La conexión con VMware vCenter debe cumplir con las políticas de seguridad establecidas por la empresa. 		
Validado por:	José Delgado Arce		
Comentarios:			

Fuente: Elaboración propia, 2024.

Tabla 16 *Requerimiento RF-02.*

ID del Requerimiento:	RF-02		
Creado por:	José Delgado Arce	Modificado por:	
Fecha Creación:	19/03/2025	Ult. Actualización:	
Módulo:	Almacenamiento		
Fuentes:	Análisis operativo y diagnóstico de percepción ejecutados		
Actores:	AC-1, AC-2, AC-3		
Objetivo:	Permitir que el sistema cuente con una base de conocimiento capaz de almacenar, organizar y procesar información estructurada para su uso en la toma de decisiones.		
Descripción:	El sistema debe contar con una base de conocimiento que permita almacenar, organizar y procesar información estructurada para su uso prolongado, favoreciendo la toma de decisiones en la organización.		
Importancia/Prioridad:	Alta		
Clasificación:	Almacenamiento		
Elementos de entrada de datos:	El sistema debe solicitar la siguiente información:		
	<i>Nombre</i>	<i>Descripción</i>	
	Información estructurada	Datos organizados que el sistema debe poder almacenar y procesar.	
Elementos de resultados de datos:	Información procesada y organizada de las máquinas virtuales para la toma de decisiones		
Restricciones y supuestos:	<ul style="list-style-type: none"> • La base de conocimiento debe ser accesible de manera eficiente para los usuarios autorizados. • El sistema debe asegurar la integridad de la información almacenada y la correcta organización de los datos. 		
Validado por:	José Delgado Arce		
Comentarios:			

Fuente: Elaboración propia, 2024.

Tabla 17 *Requerimiento RF-03.*

ID del Requerimiento:	RF-03		
Creado por:	José Delgado Arce	Modificado por:	
Fecha Creación:	19/03/2025	Ult. Actualización:	
Módulo:	Integración y Procesamiento		
Fuentes:	Análisis operativo y diagnóstico de percepción ejecutados		
Actores:	AC-1, AC-2, AC-3		
Objetivo:	Permitir la integración de inteligencia artificial para el análisis predictivo y la automatización de procesos operativos mediante la API ChatGPT.		
Descripción:	El sistema debe integrar la API ChatGPT para realizar análisis predictivo y automatizar procesos operativos, mejorando la eficiencia y precisión de las decisiones en tiempo real.		
Importancia/Prioridad:	Media		
Clasificación:	Integración y Procesamiento		
Elementos de entrada de datos:	El sistema debe solicitar la siguiente información:		
	<i>Nombre</i>	<i>Descripción</i>	
	Datos operativos	Información relevante para el análisis predictivo y automatización de procesos almacenada en la base de conocimiento.	
Elementos de resultados de datos:	<ul style="list-style-type: none"> • Resultados de análisis predictivo • Procesos operativos automatizados 		
Restricciones y supuestos:	<ul style="list-style-type: none"> • La integración con la API ChatGPT debe cumplir con los requisitos de seguridad y privacidad de la empresa. • El sistema debe poder gestionar las solicitudes de la API ChatGPT sin afectar el rendimiento general. 		
Validado por:	José Delgado Arce		
Comentarios:			

Fuente: Elaboración propia, 2024.

Tabla 18 *Requerimiento RF-04.*

ID del Requerimiento:	RF-04		
Creado por:	José Delgado Arce	Modificado por:	
Fecha Creación:	19/03/2025	Ult. Actualización:	
Módulo:	Visualización		
Fuentes:	Análisis operativo y diagnóstico de percepción ejecutados		
Actores:	AC-2		
Objetivo:	Integrar el sistema con Power BI para generar reportes interactivos sobre la gestión de recursos virtualizados.		
Descripción:	El sistema debe permitir la integración con Power BI, facilitando la creación de reportes interactivos que proporcionen información detallada sobre la gestión de los recursos virtualizados, mejorando la visualización y toma de decisiones.		
Importancia/Prioridad:	Alta		
Clasificación:	Visualización		
Elementos de entrada de datos:	El sistema debe solicitar la siguiente información:		
	<i>Nombre</i>	<i>Descripción</i>	
	Datos sobre recursos virtualizados	Información sobre el estado, uso y rendimiento de los recursos virtualizados para la creación de reportes, almacenados en la base de conocimiento	
Elementos de resultados de datos:	<ul style="list-style-type: none"> Reportes interactivos generados en Power BI sobre la gestión de recursos virtualizados. 		
Restricciones y supuestos:	<ul style="list-style-type: none"> La integración con Power BI debe cumplir con los estándares de seguridad y protección de datos establecidos por la empresa. El sistema debe garantizar la actualización en tiempo real de los datos mostrados en los reportes. 		
Validado por:	José Delgado Arce		
Comentarios:			

Fuente: Elaboración propia, 2024.

Tabla 19 *Requerimiento RF-05*

ID del Requerimiento:	RF-05		
Creado por:	José Delgado Arce	Modificado por:	
Fecha Creación:	19/03/2025	Ult. Actualización:	
Módulo:	Seguridad		
Fuentes:	Análisis operativo y diagnóstico de percepción ejecutados		
Actores:	AC-1, AC-2, AC-3		
Objetivo:	Permitir el acceso centralizado a los datos sin requerir la intervención manual del personal de TI.		
Descripción:	El sistema debe proporcionar un acceso centralizado a los datos, eliminando la necesidad de intervención manual por parte del personal de TI, garantizando así una gestión eficiente, segura y automatizada de la información.		
Importancia/Prioridad:	Media		
Clasificación:	Seguridad		
Elementos de entrada de datos:	El sistema debe solicitar la siguiente información:		
	<i>Nombre</i>	<i>Descripción</i>	
	Credenciales de usuario	Datos de autenticación para acceder al sistema de forma segura	
Elementos de resultados de datos:	<ul style="list-style-type: none"> Acceso a los datos centralizados en la base de conocimiento sin necesidad de intervención manual del personal de TI 		
Restricciones y supuestos:	<ul style="list-style-type: none"> El sistema debe garantizar que solo los usuarios autorizados puedan acceder a los datos centralizados. Los permisos de acceso deben estar configurados de acuerdo con las políticas y roles de seguridad de la empresa. 		
Validado por:	José Delgado Arce		
Comentarios:			

Fuente: Elaboración propia, 2024.

Tabla 20 *Requerimiento RF-06.*

ID del Requerimiento:	RF-06		
Creado por:	José Delgado Arce	Modificado por:	
Fecha Creación:	19/03/2025	Ult. Actualización:	
Módulo:	Interfaz de usuario		
Fuentes:	Análisis operativo y diagnóstico de percepción ejecutados		
Actores:	AC-4		
Objetivo:	Garantizar la autenticación de usuarios mediante una pantalla de inicio de sesión.		
Descripción:	El sistema debe contar con una pantalla de inicio de sesión que permita la autenticación de los usuarios mediante credenciales, asegurando el acceso seguro a la plataforma.		
Importancia/Prioridad:	Alta		
Clasificación:	Interfaz de usuario		
Elementos de entrada de datos:	El sistema debe solicitar la siguiente información:		
	<i>Nombre</i>	<i>Descripción</i>	
	Credenciales de usuario	Nombre de usuario y contraseña para la autenticación	
Elementos de resultados de datos:	<ul style="list-style-type: none"> • Acceso autorizado al sistema según los permisos asignados al usuario. 		
Restricciones y supuestos:	<ul style="list-style-type: none"> • El sistema debe implementar mecanismos de seguridad para evitar accesos no autorizados. • Se deben aplicar protocolos de cifrado en el almacenamiento y transmisión de credenciales. 		
Validado por:	José Delgado Arce		
Comentarios:			

Fuente: Elaboración propia, 2024.

Tabla 21 *Requerimiento RF-07*

ID del Requerimiento:	RF-07		
Creado por:	José Delgado Arce	Modificado por:	
Fecha Creación:	19/03/2025	Ult. Actualización:	
Módulo:	Interfaz de usuario		
Fuentes:	Análisis operativo y diagnóstico de percepción ejecutados		
Actores:	AC-4		
Objetivo:	Proporcionar una pantalla de inicio con opciones principales de navegación para mejorar la experiencia del usuario.		
Descripción:	El sistema debe contar con una pantalla de inicio donde se muestren las opciones principales de navegación, facilitando el acceso rápido a las funciones más utilizadas por los usuarios.		
Importancia/Prioridad:	Media		
Clasificación:	Interfaz de usuario		
Elementos de entrada de datos:	El sistema debe solicitar la siguiente información:		
	<i>Nombre</i>	<i>Descripción</i>	
	Opciones de navegación	Menús y accesos directos a las funcionalidades principales del sistema	
Elementos de resultados de datos:	<ul style="list-style-type: none"> • Visualización de la pantalla de inicio con las opciones de navegación disponibles. 		
Restricciones y supuestos:	<ul style="list-style-type: none"> • La pantalla de inicio debe ser intuitiva y de fácil acceso para los usuarios. • El diseño debe seguir las buenas prácticas de usabilidad y experiencia de usuario. 		
Validado por:	José Delgado Arce		
Comentarios:			

Fuente: Elaboración propia, 2024.

Tabla 22 *Requerimiento RF-08*

ID del Requerimiento:	RF-08		
Creado por:	José Delgado Arce	Modificado por:	
Fecha Creación:	19/03/2025	Ult. Actualización:	
Módulo:	Visualización		
Fuentes:	Análisis operativo y diagnóstico de percepción ejecutados		
Actores:	AC-4		
Objetivo:	Permitir la visualización de reportes y métricas a través de un dashboard en Power BI.		
Descripción:	El sistema debe integrar un dashboard de Power BI que permita la visualización interactiva de reportes y métricas clave, facilitando el análisis y la toma de decisiones basadas en datos.		
Importancia/Prioridad:	Alta		
Clasificación:	Visualización		
Elementos de entrada de datos:	El sistema debe solicitar la siguiente información:		
	<i>Nombre</i>	<i>Descripción</i>	
	Datos analíticos	Información estructurada y procesada para la generación de reportes en Power BI	
Elementos de resultados de datos:	<ul style="list-style-type: none"> • Dashboard con reportes y métricas interactivas en Power BI 		
Restricciones y supuestos:	<ul style="list-style-type: none"> • La integración con Power BI debe cumplir con los estándares de seguridad y acceso establecidos por la empresa. • El sistema debe garantizar la actualización en tiempo real de los datos reflejados en el dashboard. 		
Validado por:	José Delgado Arce		
Comentarios:			

Fuente: Elaboración propia, 2024.

Tabla 23 *Requerimiento RF-09.*

ID del Requerimiento:	RF-09		
Creado por:	José Delgado Arce	Modificado por:	
Fecha Creación:	19/03/2025	Ult. Actualización:	
Módulo:	Interfaz de usuario		
Fuentes:	Análisis operativo y diagnóstico de percepción ejecutados		
Actores:	AC-4		
Objetivo:	Proporcionar una pantalla de interacción con el asistente virtual para la generación de consultas.		
Descripción:	El sistema debe contar con una pantalla dedicada donde los usuarios puedan interactuar con el asistente virtual, permitiéndoles generar consultas y recibir respuestas en tiempo real.		
Importancia/Prioridad:	Media		
Clasificación:	Interfaz de usuario		
Elementos de entrada de datos:	El sistema debe solicitar la siguiente información:		
	<i>Nombre</i>	<i>Descripción</i>	
	Consulta del usuario	Texto ingresado por el usuario para solicitar información o asistencia	
Elementos de resultados de datos:	<ul style="list-style-type: none"> • Respuesta de la asistente virtual basada en la consulta realizada basada en los datos analizados de la infraestructura virtual. 		
Restricciones y supuestos:	<ul style="list-style-type: none"> • El asistente virtual debe procesar las consultas en tiempo real y proporcionar respuestas precisas. • La interfaz debe ser intuitiva y accesible para todos los usuarios. 		
Validado por:	José Delgado Arce		
Comentarios:			

Fuente: Elaboración propia, 2024.

5.1.7 Casos de uso

A continuación, se presentan los casos de uso asociados a los requerimientos funcionales del sistema.

Tabla 24 *Listado casos de uso.*

Caso de Uso	Requerimiento	Descripción
CU-1	RF-06	Autenticación de usuario en el sistema para garantizar acceso seguro mediante credenciales y doble factor de autenticación.
CU-2	RF-05	Validación de permisos del usuario para asegurar que cuenta con los privilegios necesarios dentro del sistema.
CU-3	RF-07	Visualización de la página de inicio, permitiendo la selección de opciones principales de navegación.
CU-4	RF-01	Extracción de datos en tiempo real desde VMware vCenter utilizando la API de PyVmomi.
CU-5	RF-02	Procesamiento y almacenamiento de información en la base de conocimiento para su uso en toma de decisiones.
CU-6	RF-03	Análisis de datos mediante inteligencia artificial para la generación de reportes y automatización de procesos.
CU-7	RF-08	Visualización de reportes y métricas en un dashboard interactivo de Power BI.
CU-8	RF-04	Generación de reportes en Power BI con información de la base de conocimiento y datos analizados.
CU-9	RF-09	Interacción con el asistente virtual para realizar consultas sobre los datos almacenados y análisis generados.

Fuente: Elaboración propia, 2024.

5.1.8 Descripción casos de uso

A continuación, se detallan los casos de uso con su respectiva tabla con la información precisa de cada uno de los flujos que forman parte del sistema, asegurando su correcta integración y operación.

Tabla 25 Caso de uso CU-1.

Caso de uso	CU-1		
Creado por:	José Delgado Arce	Fecha de creación:	
Actores	AC-1 – 4	Prioridad	Alta
Objetivo	El usuario debe autenticarse con éxito cuando ingresa a la aplicación.	Requisitos	RF-1-9
Pre-condición	El usuario debe contar con credenciales válidas y estar registrado en el sistema.		
Postcondición			
Flujo de eventos			
1	El usuario accede a la pantalla de inicio de sesión.		
2	Ingresa sus credenciales (usuario y contraseña).		
3	El sistema verifica la autenticidad de los datos.		
4	Si la autenticación es correcta, el sistema concede acceso.		
Flujos alternos			
1	Si las credenciales son incorrectas, se muestra un mensaje de error y se solicita el reingreso.		
Extensiones e inclusiones			

Fuente: Elaboración propia, 2024.

Tabla 26 *Caso de uso CU-2.*

Caso de uso	CU-2		
Creado por:	José Delgado Arce	Fecha de creación:	
Actores	AC-1 – 4	Prioridad	Alta
Objetivo	Garantizar que el usuario cuenta con los privilegios necesarios dentro del sistema.	Requisitos	RF-1-9
Pre-condición	El usuario debe estar autenticado en el sistema.		
Postcondición	El usuario accede únicamente a las funciones autorizadas.		
Flujo de eventos			
1	El usuario accede al sistema tras la autenticación.		
2	El sistema verifica los permisos asociados al perfil.		
3	Si los permisos son válidos, el usuario puede acceder a sus funciones permitidas.		
4	Si el usuario intenta acceder a una función no autorizada, se muestra un mensaje de restricción.		
Flujos alternos			
1	Si no se encuentran permisos válidos, se redirige al usuario a una pantalla de error.		
Extensiones e inclusiones			

Fuente: Elaboración propia, 2024.

Tabla 27 Caso de uso CU-3.

Caso de uso	CU-3		
Creado por:	José Delgado Arce	Fecha de creación:	
Actores	AC-1 – 4	Prioridad	Alta
Objetivo	Permitir al usuario visualizar la página de inicio con las opciones principales de navegación.	Requisitos	RF-07
Pre-condición	El usuario debe estar autenticado en el sistema.		
Postcondición	El usuario visualiza correctamente las opciones de navegación.		
Flujo de eventos			
1	El usuario accede al sistema tras la autenticación.		
2	Se carga la pantalla de inicio con las opciones principales de navegación.		
3	El usuario selecciona una opción para continuar.		
Flujos alternos			
1	Si ocurre un error en la carga de la página, se muestra un mensaje de error y se intenta recargar la página.		
Extensiones e inclusiones			

Fuente: Elaboración propia, 2024.

Tabla 28 Caso de uso CU-4.

Caso de uso	CU-4		
Creado por:	José Delgado Arce	Fecha de creación:	
Actores	AC-1 – 4	Prioridad	Alta
Objetivo	Extraer datos en tiempo real desde VMware vCenter para su análisis.	Requisitos	RF-01
Pre-condición	El sistema debe estar configurado con acceso a VMware vCenter.		
Postcondición	Los datos extraídos están disponibles para análisis.		
Flujo de eventos			
1	El sistema se conecta a VMware vCenter mediante la API de PyVmomi.		
2	Se extraen los datos requeridos en tiempo real.		
3	Los datos se almacenan en la base de conocimiento.		
Flujos alternos			
1	Si la conexión falla, se muestra un mensaje de error y se envía correo de fallo en el proceso de extracción.		
Extensiones e inclusiones	CU-5 - Almacenamiento en la base de conocimiento.		

Fuente: Elaboración propia, 2024.

Tabla 29 Caso de uso CU-5.

Caso de uso	CU-5		
Creado por:	José Delgado Arce	Fecha de creación:	
Actores	AC-1 – 4	Prioridad	Alta
Objetivo	Garantizar que los datos extraídos sean procesados y almacenados correctamente en la base de conocimiento.	Requisitos	RF-02
Pre-condición	Los datos deben haber sido extraídos correctamente desde VMware vCenter.		
Postcondición	La información es accesible para su análisis y consulta posterior.		
Flujo de eventos			
1	El sistema recibe los datos extraídos desde VMware vCenter.		
2	Se procesan los datos para su estructuración y optimización.		
3	La información se almacena en la base de conocimiento.		
4	Se validan los datos almacenados para garantizar su integridad.		
Flujos alternos			
1	Si la validación de datos falla, se ejecuta un proceso de corrección automática o se notifica al administrador.		
Extensiones e inclusiones	<ul style="list-style-type: none"> • CU-4 - Extracción de Datos desde VMware vCenter. • CU-6 - Análisis de Datos mediante Inteligencia Artificial. 		

Fuente: Elaboración propia, 2024.

Tabla 30 Caso de uso CU-6.

Caso de uso	CU-6		
Creado por:	José Delgado Arce	Fecha de creación:	
Actores	AC-1 – 4	Prioridad	Media
Objetivo	Aplicar inteligencia artificial para el análisis predictivo y automatización	Requisitos	RF-03
Pre-condición	La información debe estar almacenada en la base de conocimiento.		
Postcondición	Se generan reportes y predicciones basadas en los datos analizados.		
Flujo de eventos			
1	El usuario selecciona la opción de análisis de datos.		
2	El sistema accede a la base de conocimiento y recupera los datos.		
3	Se aplican algoritmos de inteligencia artificial para detectar patrones y generar predicciones.		
4	Los resultados del análisis se almacenan y se presentan en reportes detallados.		
Flujos alternos			
1	Si la cantidad de datos es insuficiente para el análisis, se genera una advertencia al usuario.		
Extensiones e inclusiones	<ul style="list-style-type: none"> • CU-5 - Procesamiento y almacenamiento de información. • CU-8 - Generación de reportes en Power BI. 		

Fuente: Elaboración propia, 2024.

Tabla 31 Caso de uso CU-7.

Caso de uso	CU-7		
Creado por:	José Delgado Arce	Fecha de creación:	
Actores	AC-1 – 4	Prioridad	Alta
Objetivo	Permitir la visualización de reportes y métricas en Power BI.	Requisitos	RF-08
Pre-condición	Los datos deben estar almacenados y procesados en la base de conocimiento.		
Postcondición	Se muestran los reportes de forma gráfica e interactiva.		
Flujo de eventos			
1	El usuario accede al módulo de reportes.		
2	Se recuperan los datos procesados de la base de conocimiento.		
3	Los datos se presentan en un dashboard de Power BI con gráficos interactivos.		
Flujos alternos			
1	Si no hay datos disponibles, se muestra un mensaje de error.		
Extensiones e inclusiones	<ul style="list-style-type: none"> • CU-6 - Análisis de datos mediante inteligencia artificial. • CU-8 - Generación de reportes en Power BI. 		

Fuente: Elaboración propia, 2024.

Tabla 32 Caso de uso CU-8.

Caso de uso	CU-8		
Creado por:	José Delgado Arce	Fecha de creación:	
Actores	AC-1 – 4	Prioridad	Alta
Objetivo	Permitir la generación de reportes interactivos con Power BI.	Requisitos	RF-04
Pre-condición	Los datos deben estar disponibles y organizados en la base de conocimiento.		
Postcondición	Se generan reportes estructurados con métricas relevantes.		
Flujo de eventos			
1	El usuario selecciona la opción de generación de reportes.		
2	Se recuperan los datos procesados de la base de conocimiento.		
3	El sistema genera reportes en Power BI con métricas relevantes.		
4	Los reportes se almacenan para futuras consultas.		
Flujos alternos			
1	Si no hay datos suficientes para generar el reporte, se muestra un mensaje de error.		
Extensiones e inclusiones	<ul style="list-style-type: none"> • CU-5 - Procesamiento y almacenamiento de datos. • CU-7 - Visualización de reportes en Power BI. 		

Fuente: Elaboración propia, 2024.

Tabla 33 Caso de uso CU-9.

Caso de uso	CU-9		
Creado por:	José Delgado Arce	Fecha de creación:	
Actores	AC-1 – 4	Prioridad	Media
Objetivo	Permitir la interacción con el asistente virtual para realizar consultas sobre los datos almacenados.	Requisitos	RF-09
Pre-condición	La información debe estar disponible en la base de conocimiento.		
Postcondición	Se obtiene una respuesta precisa basada en los datos almacenados.		
Flujo de eventos			
1	El usuario accede a la interfaz del asistente virtual.		
2	Ingresa una consulta en lenguaje natural.		
3	El asistente analiza la solicitud y recupera los datos relevantes.		
4	Se presenta la respuesta al usuario en la interfaz.		
Flujos alternos			
1	Si no hay datos suficientes para responder, se genera una respuesta predeterminada con las diferentes opciones.		
Extensiones e inclusiones	<ul style="list-style-type: none"> • CU-5 - Procesamiento y almacenamiento de datos. • CU-6 - Análisis de datos mediante inteligencia artificial. 		

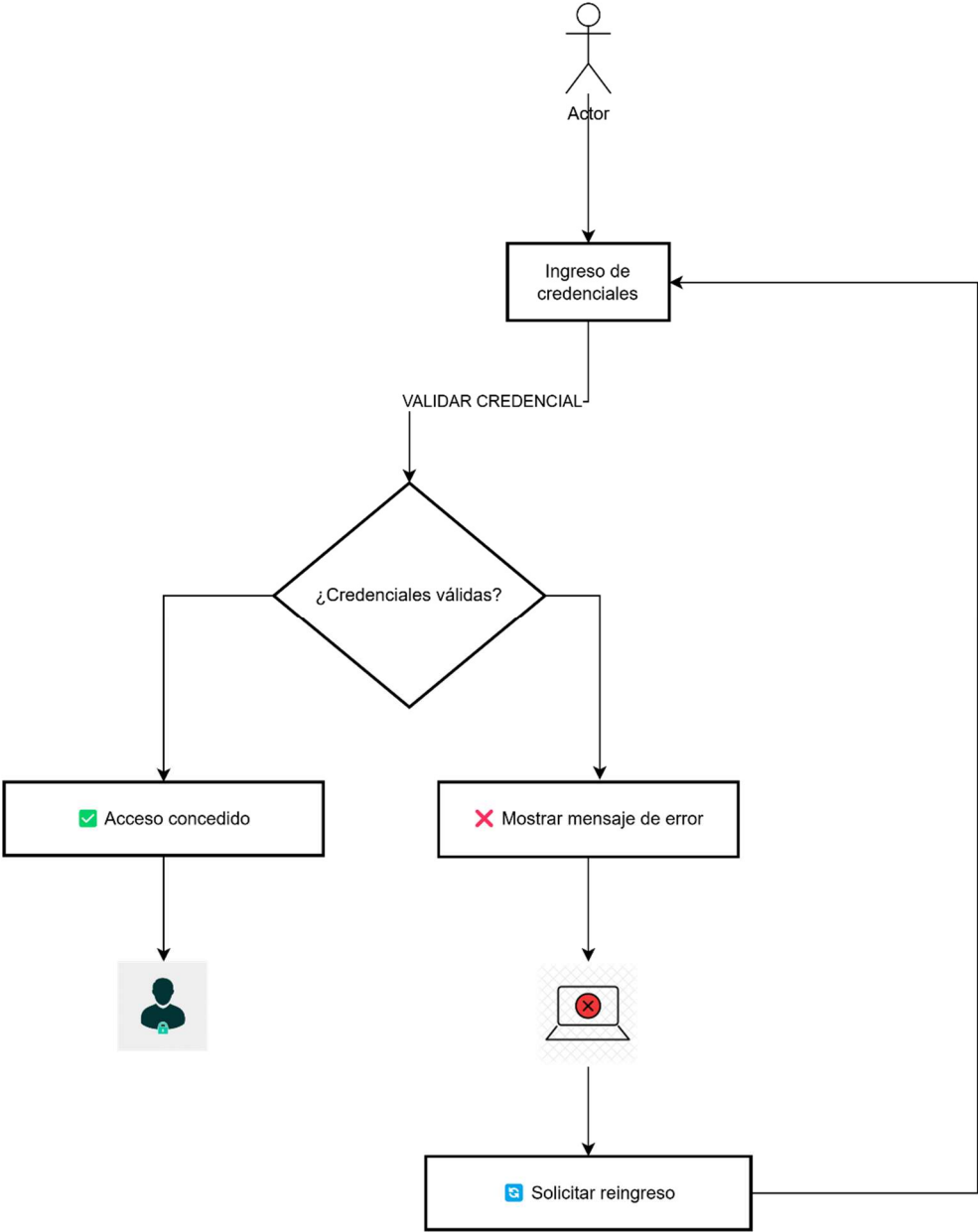
Fuente: Elaboración propia, 2024.

5.1.9 Diagramas de casos de uso.

A continuación, se muestran los diagramas de los casos de uso, los cuales tienen como objetivo simplificar de una manera gráfica la representación de los actores y sus relaciones con la funcionalidad del sistema.

Figura 24

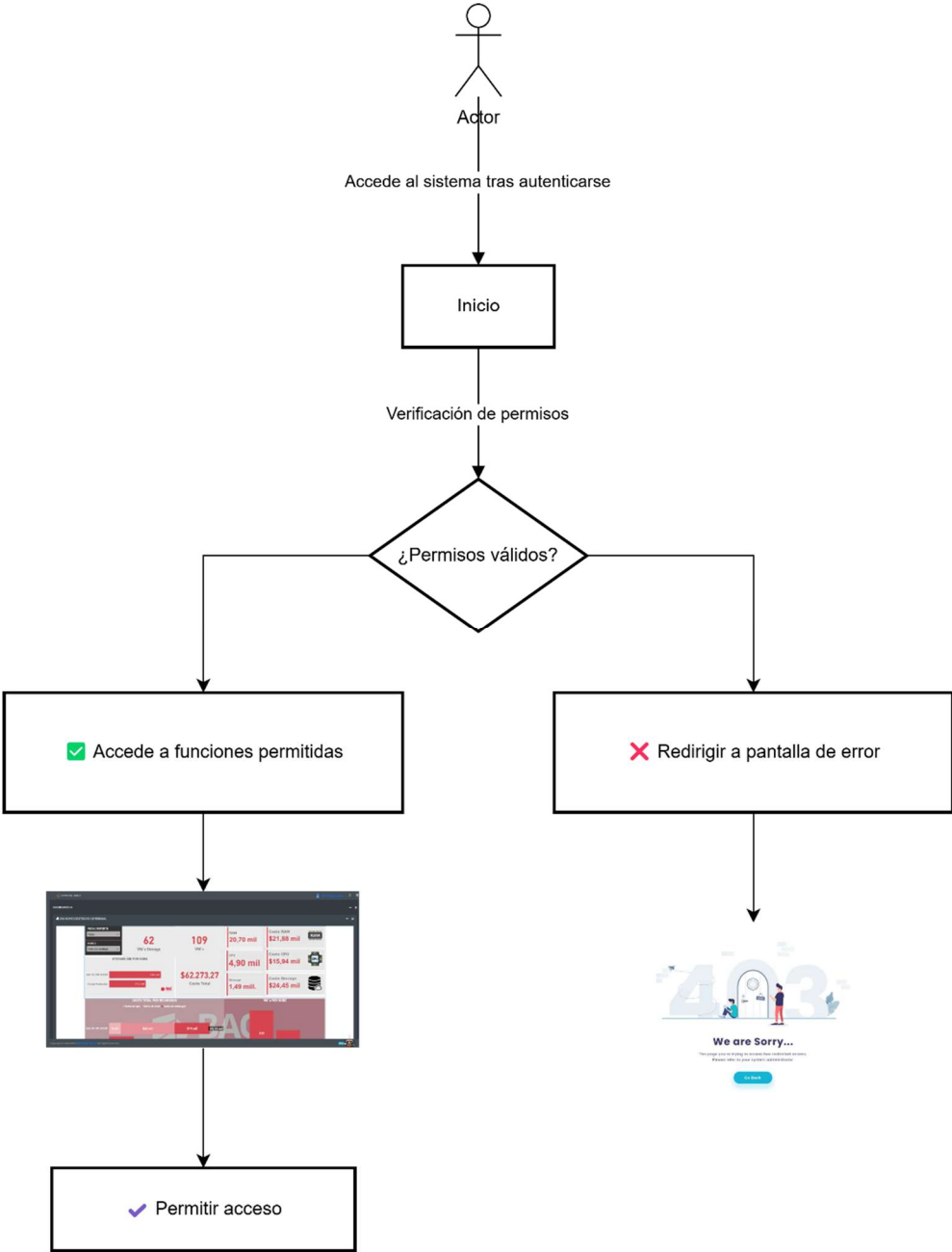
Diagrama de caso de uso CU-1



Fuente: Elaboración propia, 2024.

Figura 25

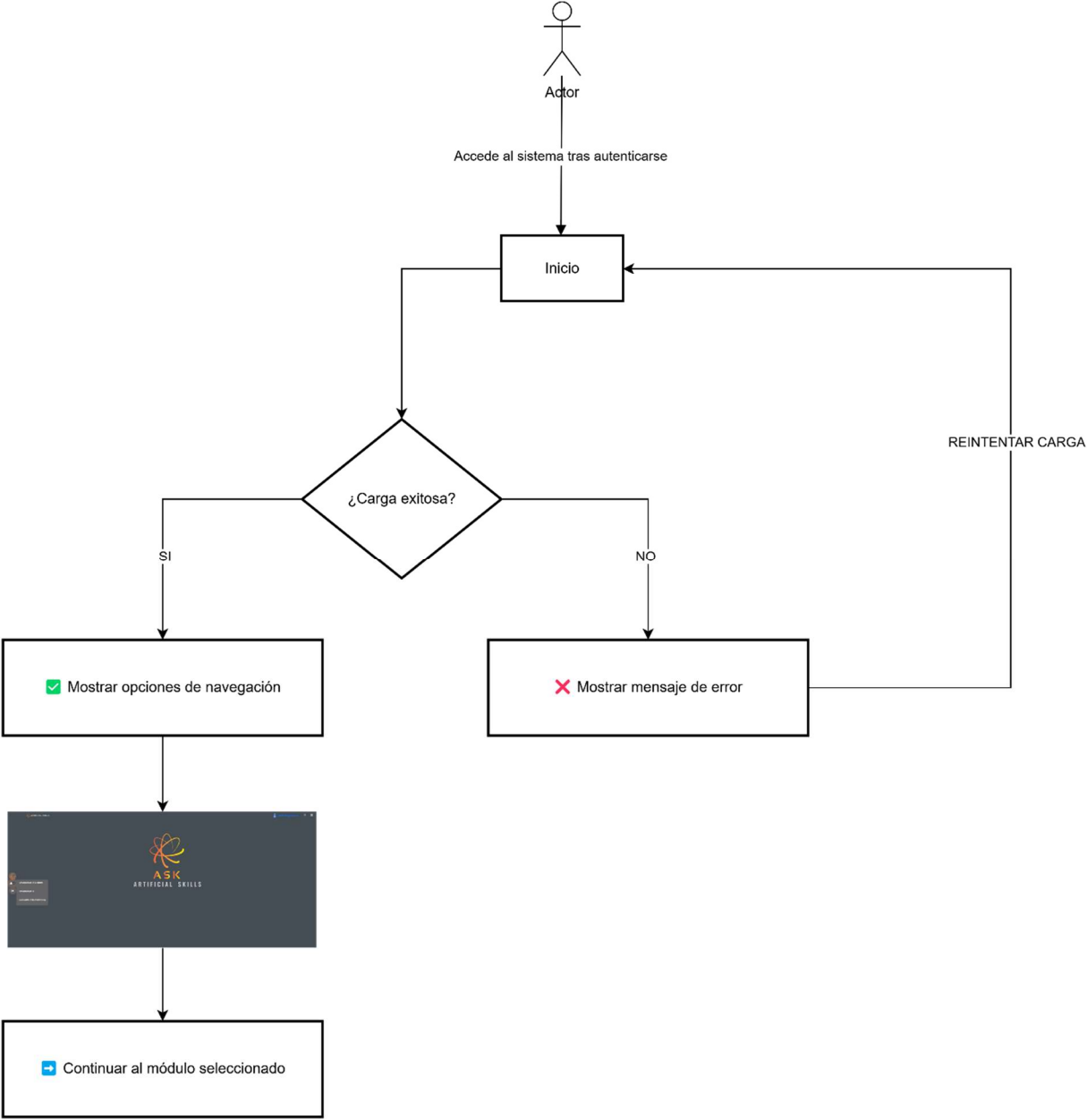
Diagrama de caso de uso CU-2



Fuente: Elaboración propia, 2024.

Figura 26

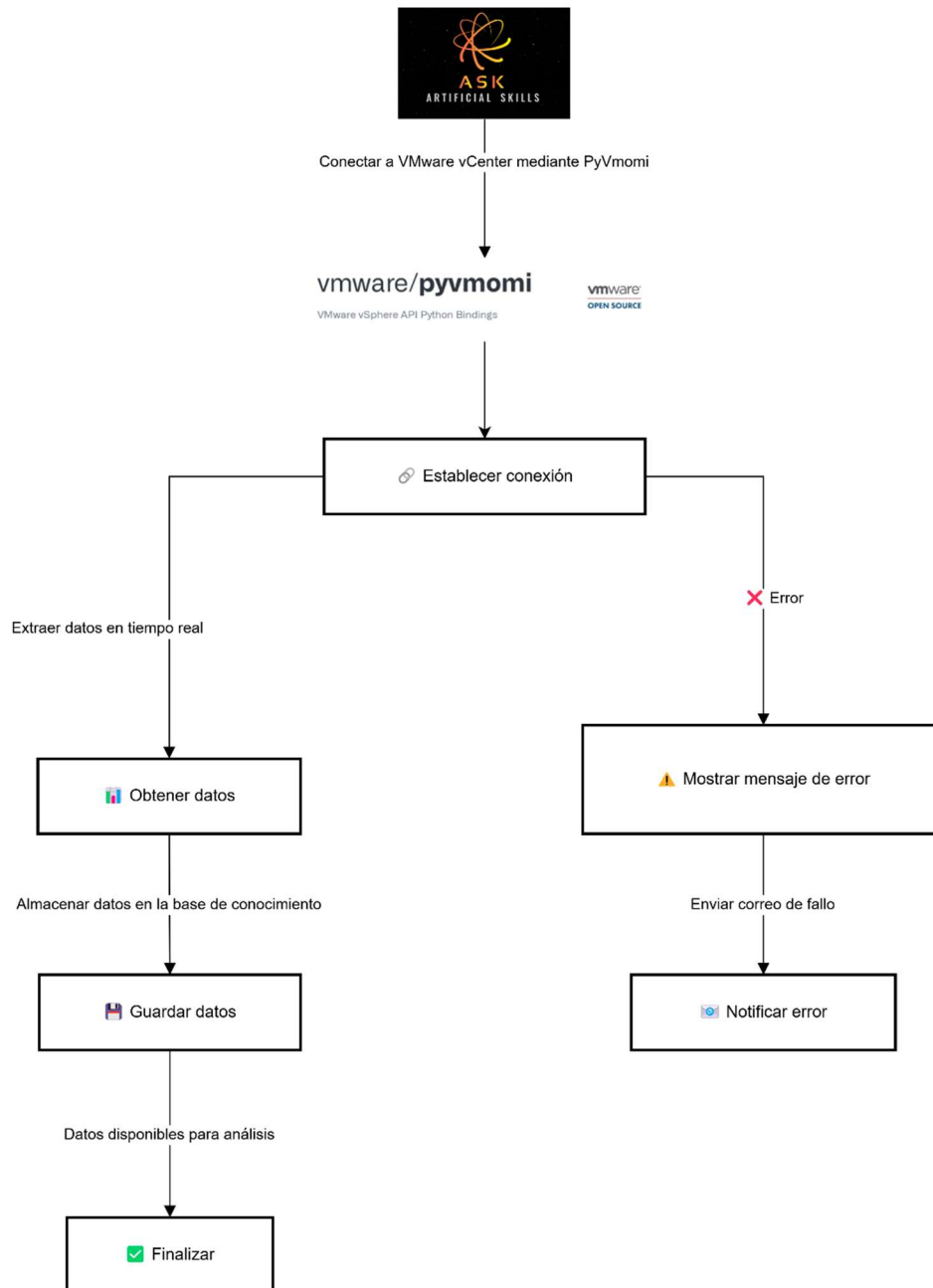
Diagrama de caso de uso CU-3



Fuente: Elaboración propia, 2024.

Figura 27

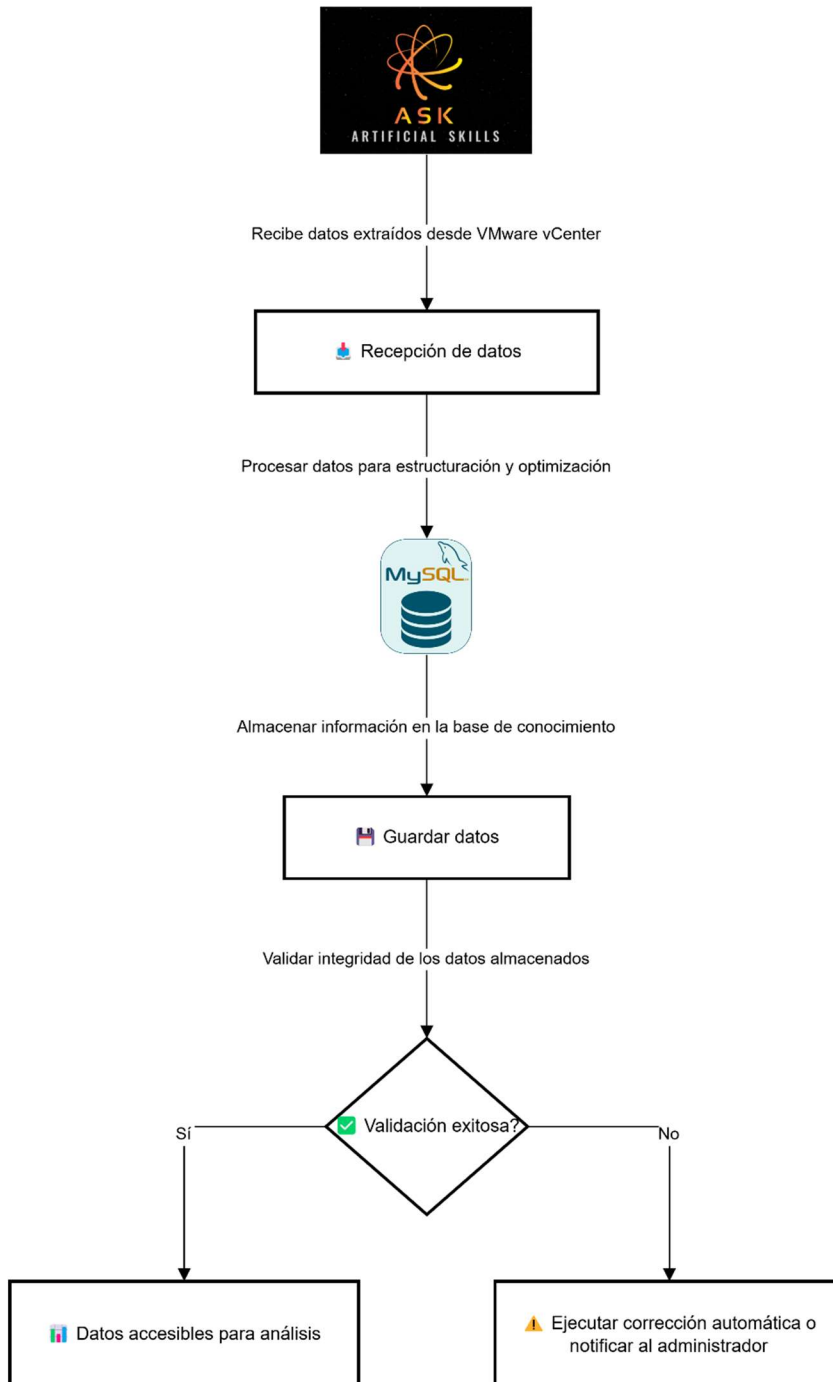
Diagrama de caso de uso CU-4



Fuente: Elaboración propia, 2024.

Figura 28

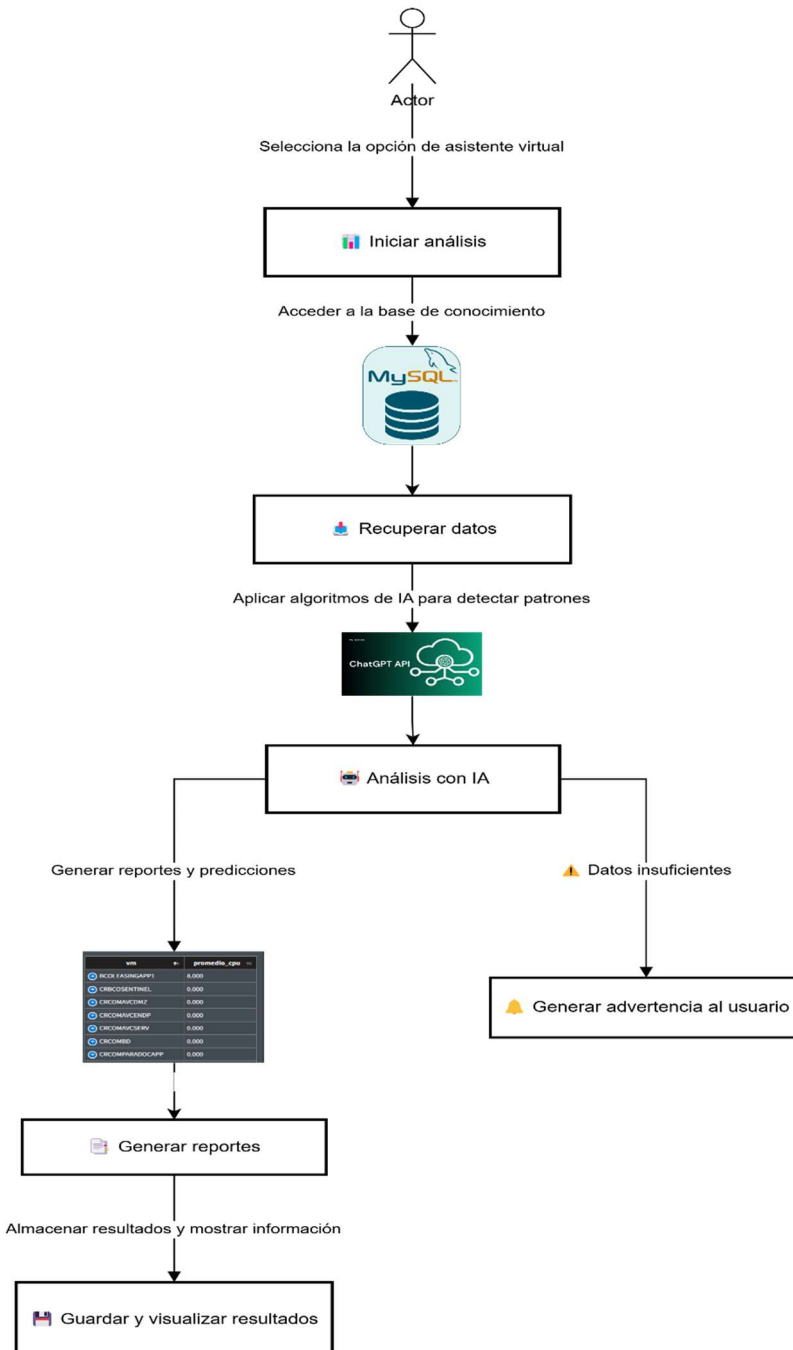
Diagrama de caso de uso CU-5



Fuente: Elaboración propia, 2024.

Figura 29

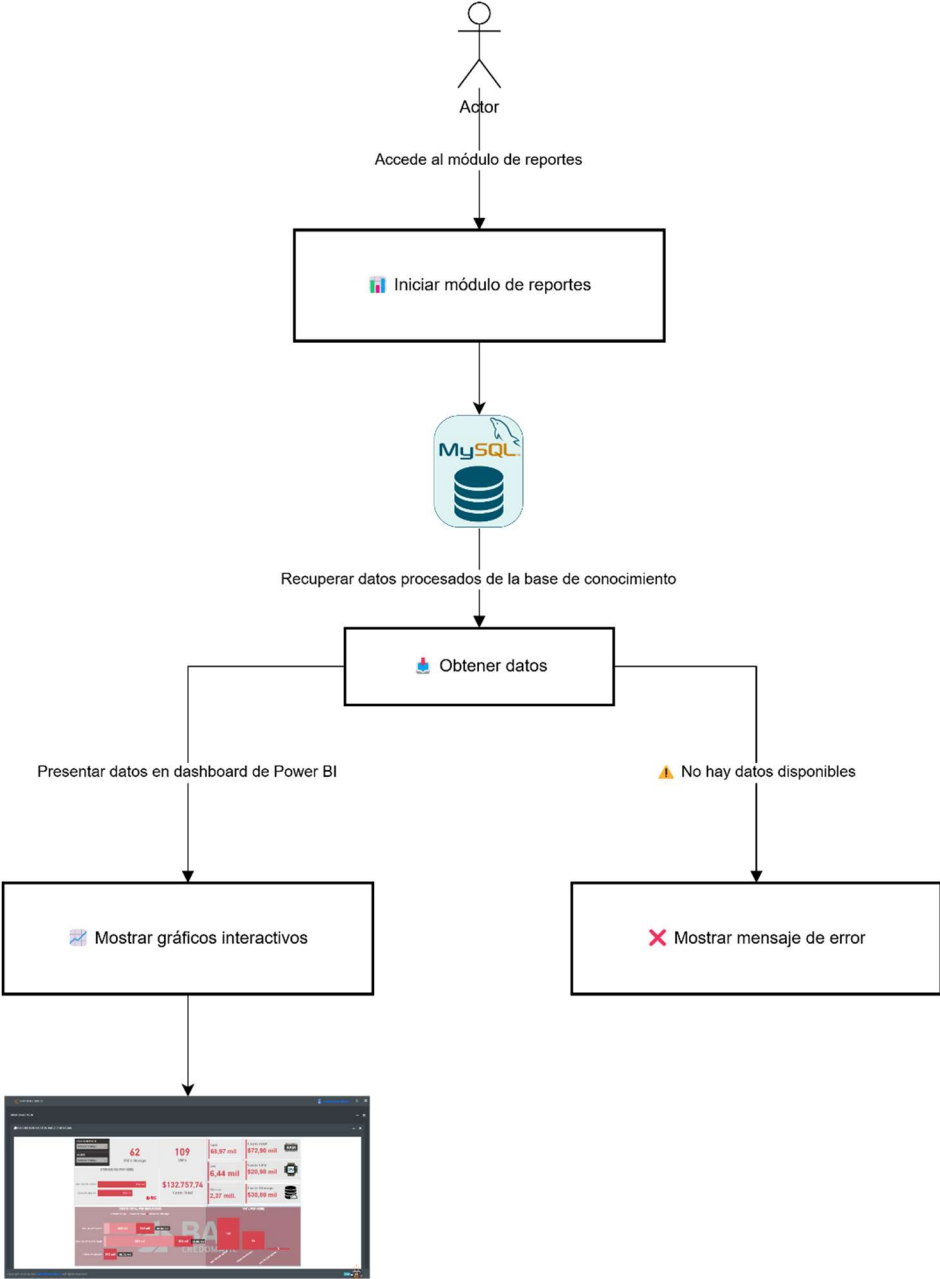
Diagrama de caso de uso CU-6



Fuente: Elaboración propia, 2024.

Figura 30

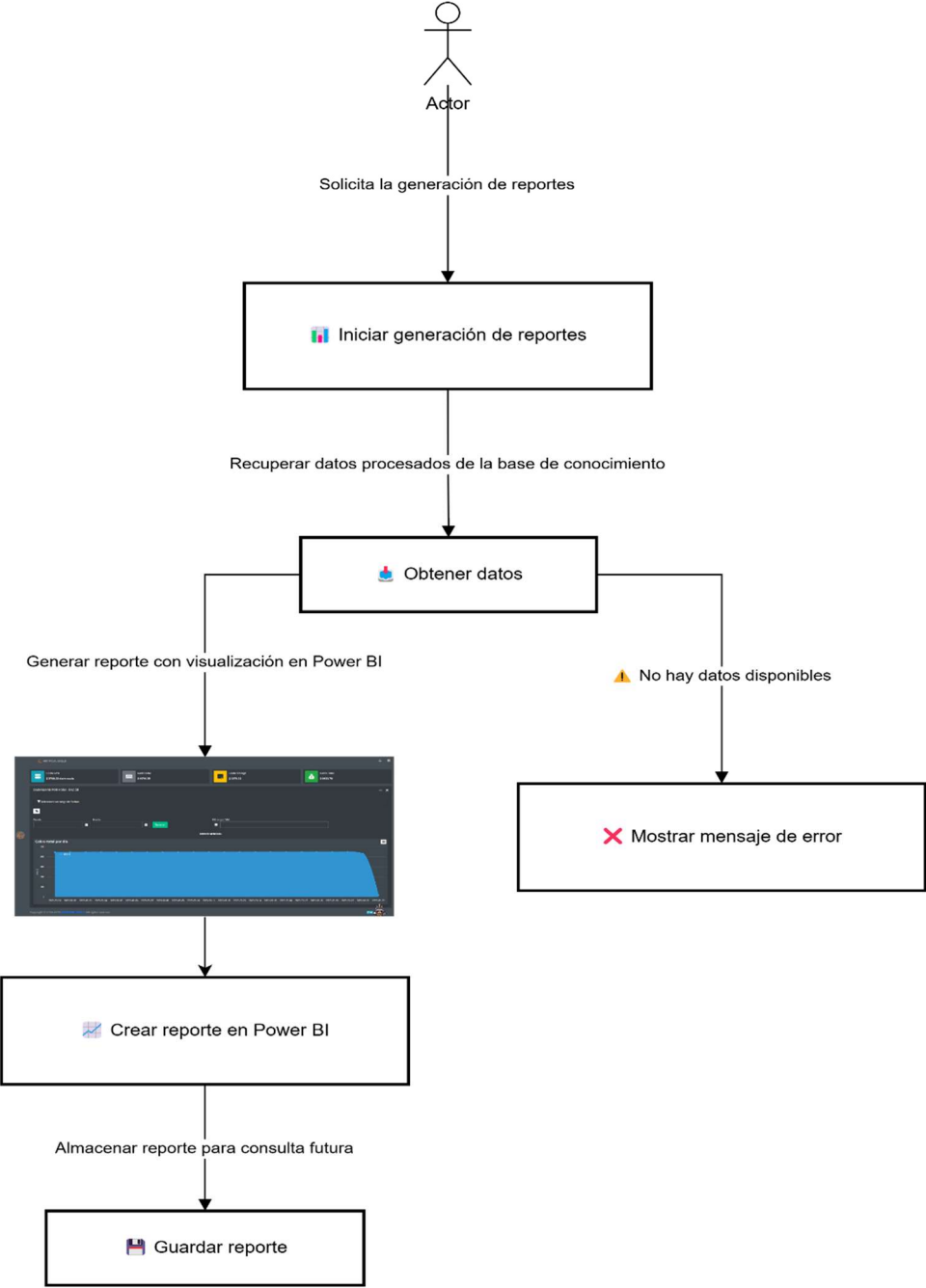
Diagrama de caso de uso CU-7



Fuente: Elaboración propia, 2024.

Figura 31

Diagrama de caso de uso CU-8



Fuente: Elaboración propia, 2024.

Figura 32

Diagrama caso de uso CU-9



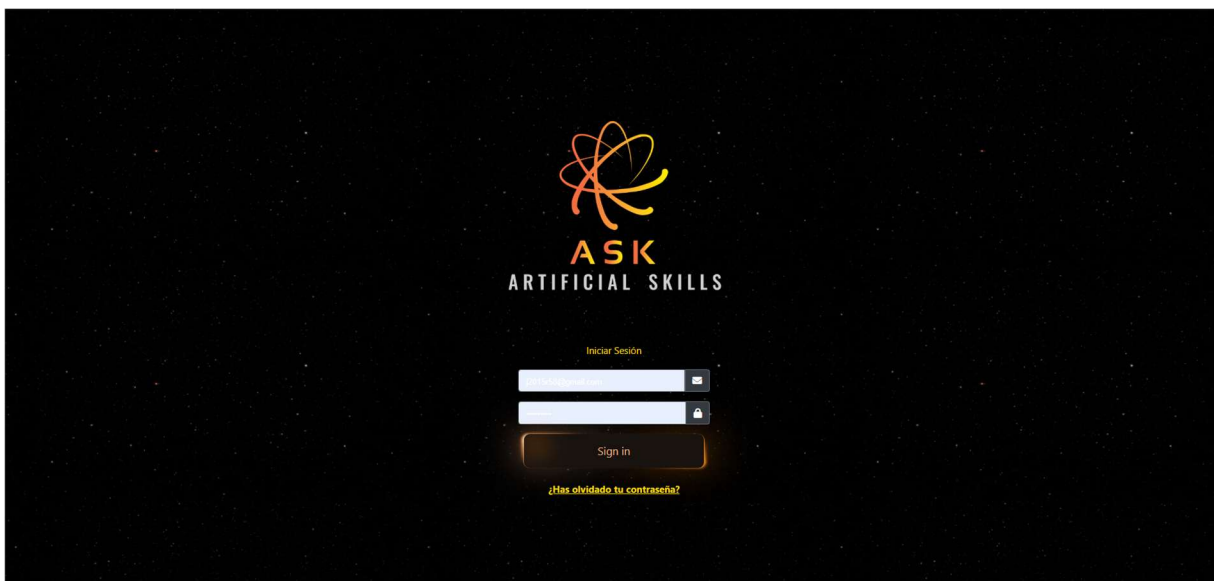
Fuente: Elaboración propia, 2024.

5.1.10 Diseño de la interfaz.

El diseño de la interfaz de usuario es un aspecto fundamental del sistema, ya que define la forma en que los usuarios interactúan con las funcionalidades disponibles. Una interfaz intuitiva, eficiente y bien estructurada permite mejorar la experiencia del usuario, optimizar la navegación y garantizar el acceso a las herramientas clave del sistema; por dicho motivo a continuación se muestra el diseño propuesto con los elementos clave del sistema.

Figura 33

Página de login



Fuente: Elaboración propia, 2024.

Figura 34

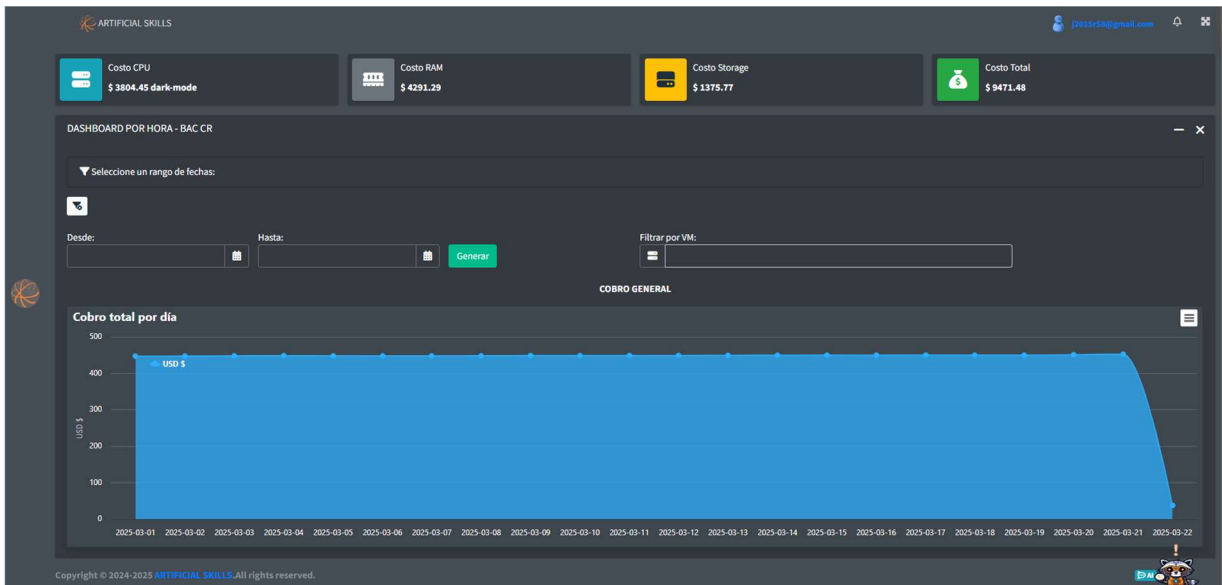
Página de inicio



Fuente: Elaboración propia, 2024.

Figura 35

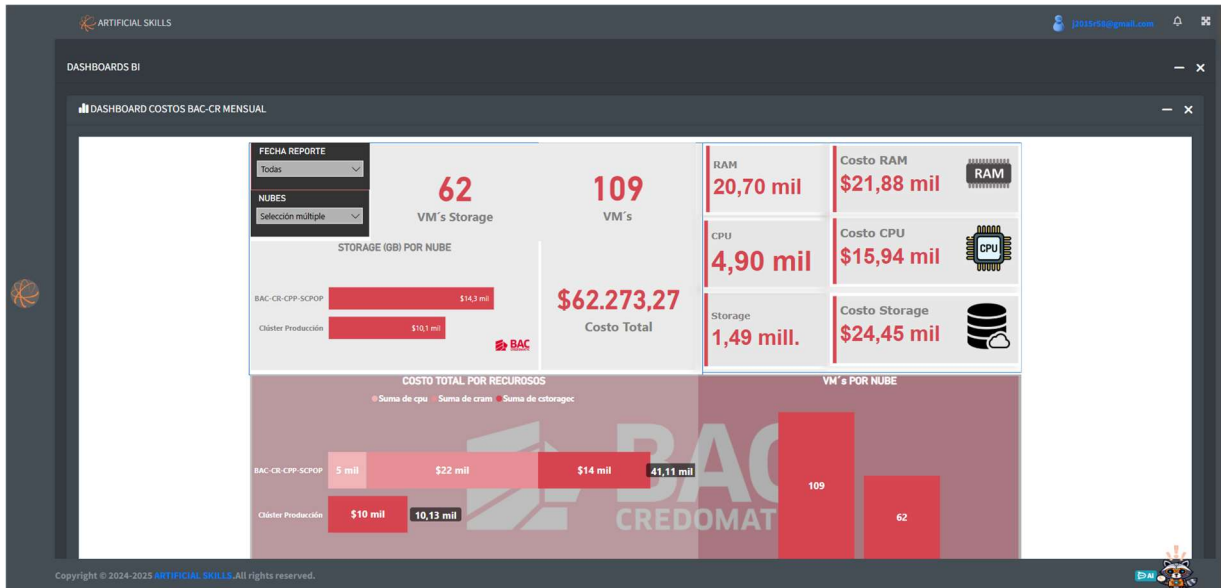
Página generación de reportes



Fuente: Elaboración propia, 2024.

Figura 36

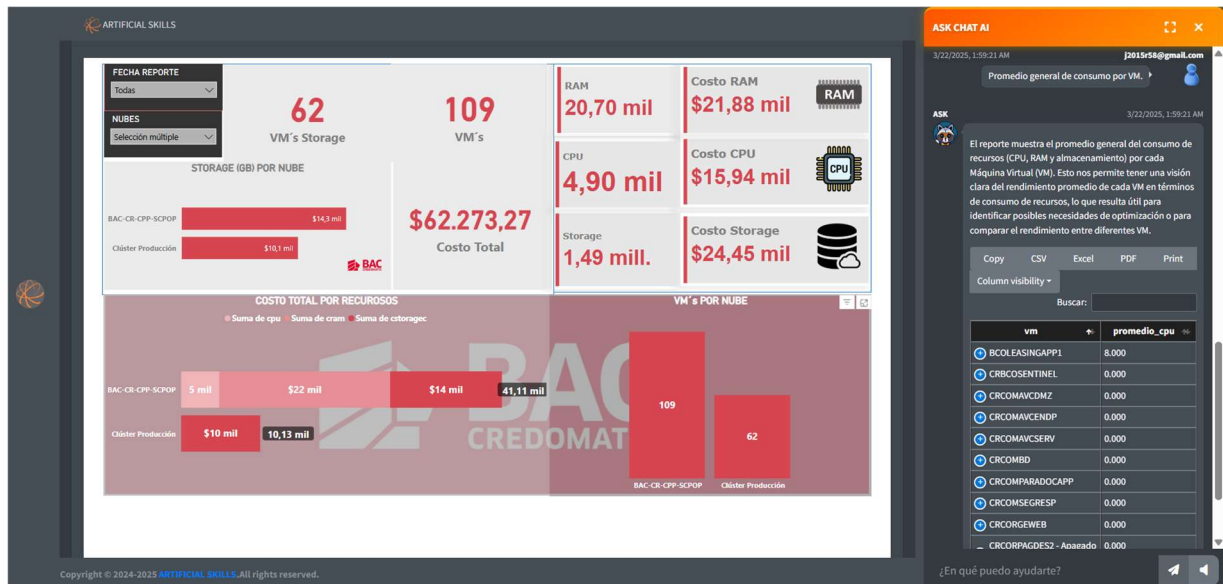
Página visualización de reportes



Fuente: Elaboración propia, 2024.

Figura 37

Página asistente virtual



Fuente: Elaboración propia, 2024.

5.1.11 Herramientas de inteligencia artificial y analítica de datos.

Tras analizar en detalle el funcionamiento operativo de Artificial Skills y definidos los casos de uso y requerimientos funcionales necesarios para el desarrollo de este proyecto, se ha determinado un conjunto de herramientas esenciales para automatizar de manera eficiente los procesos de gestión de datos. El objetivo es minimizar las tareas manuales que actualmente realizan los colaboradores, optimizando el flujo de trabajo y aumentando tanto la precisión como la rapidez en la ejecución de actividades clave. Este enfoque no solo busca reducir la intervención manual, sino también asegurar que los procesos estén alineados con los objetivos y esquemas operativos de la empresa, fomentando así una mejora continua. A continuación, se expondrán las herramientas necesarias para llevar a cabo la propuesta de automatización; la identificación de estas se basa en el análisis realizado a lo largo del proyecto, destacando

funcionalidades clave y los beneficios esperados de su aplicación, asegurando nuevamente que las mismas cumplan con las expectativas y necesidades expuestas anteriormente.

API PyVmomi: Automatización recolección de datos VMware vCenter

Inicialmente para el desarrollo de este proyecto, se propone la implementación de PyVmomi, la cual como se menciona en el marco teórico es una API de código abierto basada en el lenguaje de programación Python diseñada específicamente para que actúe como un puente entre la plataforma VMware vSphere y los desarrolladores. Esta herramienta, que opera sobre el protocolo SOAP (Simple Object Access Protocol), está diseñada específicamente para optimizar la conexión, autenticación y gestión de infraestructuras virtualizadas. Como destaca Broadcom Inc. (2023), PyVmomi proporciona una capa de conveniencia sobre las APIs fundamentales de VMware, simplificando tareas complejas y mejorando significativamente la eficiencia operativa (Broadcom Inc, 2023).

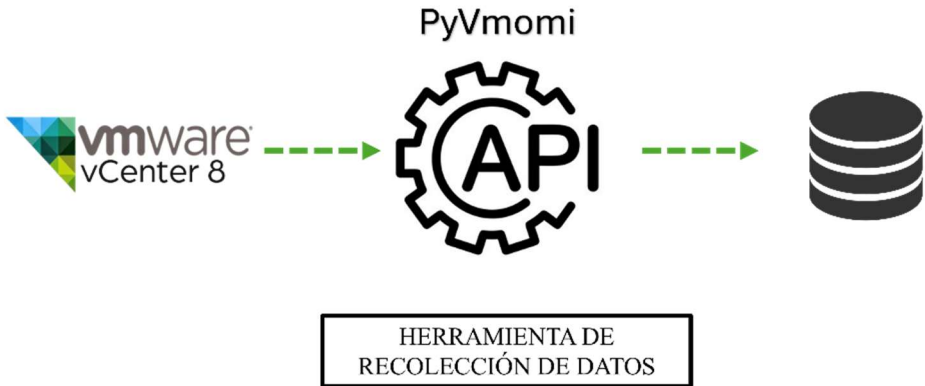
La relevancia del uso de las APIs en el desarrollo tecnológico, como señala Red Hat (2023), radica en su capacidad para estandarizar la comunicación entre sistemas (Red Hat, 2023). PyVmomi puede aprovechar esta característica permitiendo una recolección automatizada de los datos clave de la infraestructura virtual, eliminando la necesidad de incurrir en procesos manuales que, según el análisis realizado en el capítulo IV, generan retrasos y errores humanos. Además, esta API facilita la interacción con servicios básicos de VMware tal como vCenter Single Sign-On, ESX Agent Manager y Storage Policy API, integrando sistemas clave para el manejo eficiente de la infraestructura virtual (Broadcom Inc, 2023).

El propósito principal de integrar PyVmomi en esta propuesta es reemplazar los procesos manuales realizados por los operadores de infraestructura en la recolección de datos, sustituyéndolos por tareas automatizadas programadas. Este cambio permitirá una gestión

integral de la automatización; donde el flujo propuesto comienza con la configuración de una fuente de datos capaz de interactuar a través de una API. PyVmomi actuará como un puente que facilitará la comunicación entre la base de conocimiento y la fuente de datos externa (VMware vCenter), logrando la extracción automatizada de información.

Figura 38

Integración API PyVmomi

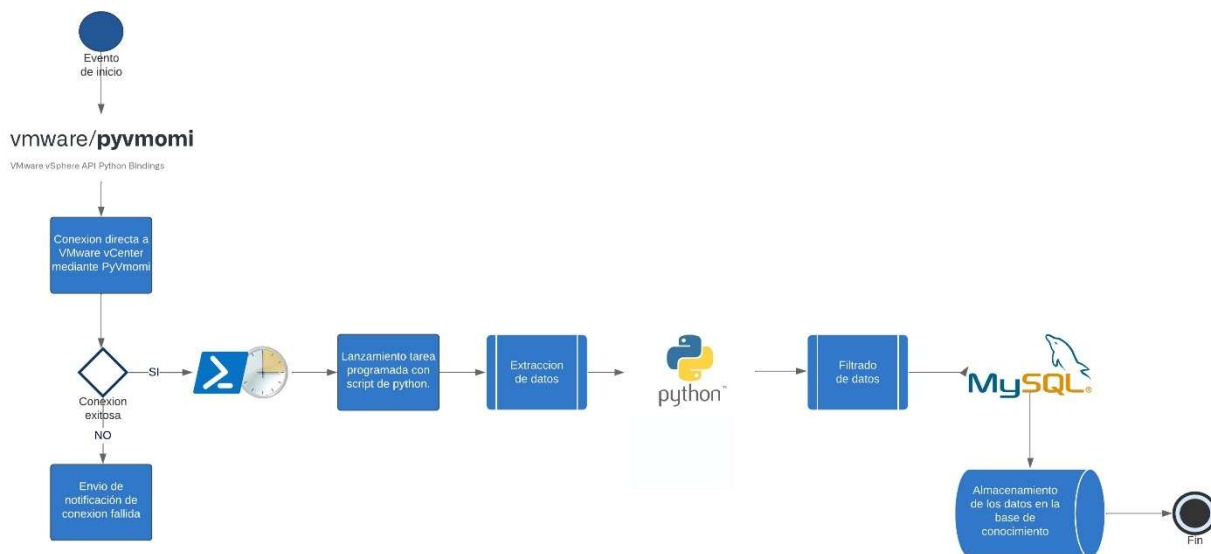


Fuente: Elaboración propia, 2024.

Como se muestra en la figura 23 – Integración API PyVmomi, la intención mencionada anteriormente se basa en que la API de PyVmomi se encuentre en el medio entre la infraestructura virtual en VMware vCenter y la base de conocimiento y que mediante la ejecución de tareas programadas esta pueda extraer los datos directamente desde VMware y almacenarlos en la base de conocimiento que será expuesta más adelante, sin la necesidad de una intervención humana, adicionalmente a continuación se muestra el diagrama de flujo del funcionamiento en el proceso de extracción de datos mediante la API de PyVmomi.

Figura 39

Diagrama de flujo extracción de datos



Fuente: Elaboración propia, 2024.

Entre los beneficios específicos del uso de la API de PyVmomi que se esperan obtener con el desarrollo de esta propuesta se encuentran la siguiente lista que podemos segregar en 3 puntos significativos:

1. Automatización en la extracción de datos.

La utilización de PyVmomi permitirá recolectar todos aquellos datos críticos mencionados en el capítulo IV sobre consumo de recursos tales como CPU, memoria y almacenamiento (ver tabla 5- Lista de datos incluidos en el informe); directamente desde VMware vCenter, reduciendo la interacción manual por parte de los colaboradores y por ende los tiempos operativos; asegurando datos precisos y confiables en tiempo real.

2. Optimización de tareas complejas.

Adicionalmente con el uso de PyVmomi se simplificará los procesos de administración, como la sincronización de tareas cliente-servidor y el seguimiento de procesos asíncronos, evitando cuellos de botella.

3. Reducción de errores humanos.

Y por último como se mencionó anteriormente, al automatizar el proceso actual de recolección de datos con PyVmomi se espera eliminar las inconsistencias derivadas de la intervención manual de los operadores y analistas; mejorando la calidad de los informes generados.

Cabe destacar que la implementación de PyVmomi, además de que se alinea con los principios fundamentales de las APIs al fomentar la interoperabilidad, escalabilidad y modularidad (Goodwin, 2024), proveen un diseño flexible, un costo de uso nulo y una amplia documentación sobre su uso proporcionada por el mismo fabricante, convirtiéndola en una herramienta clave para el desarrollo de este proyecto.

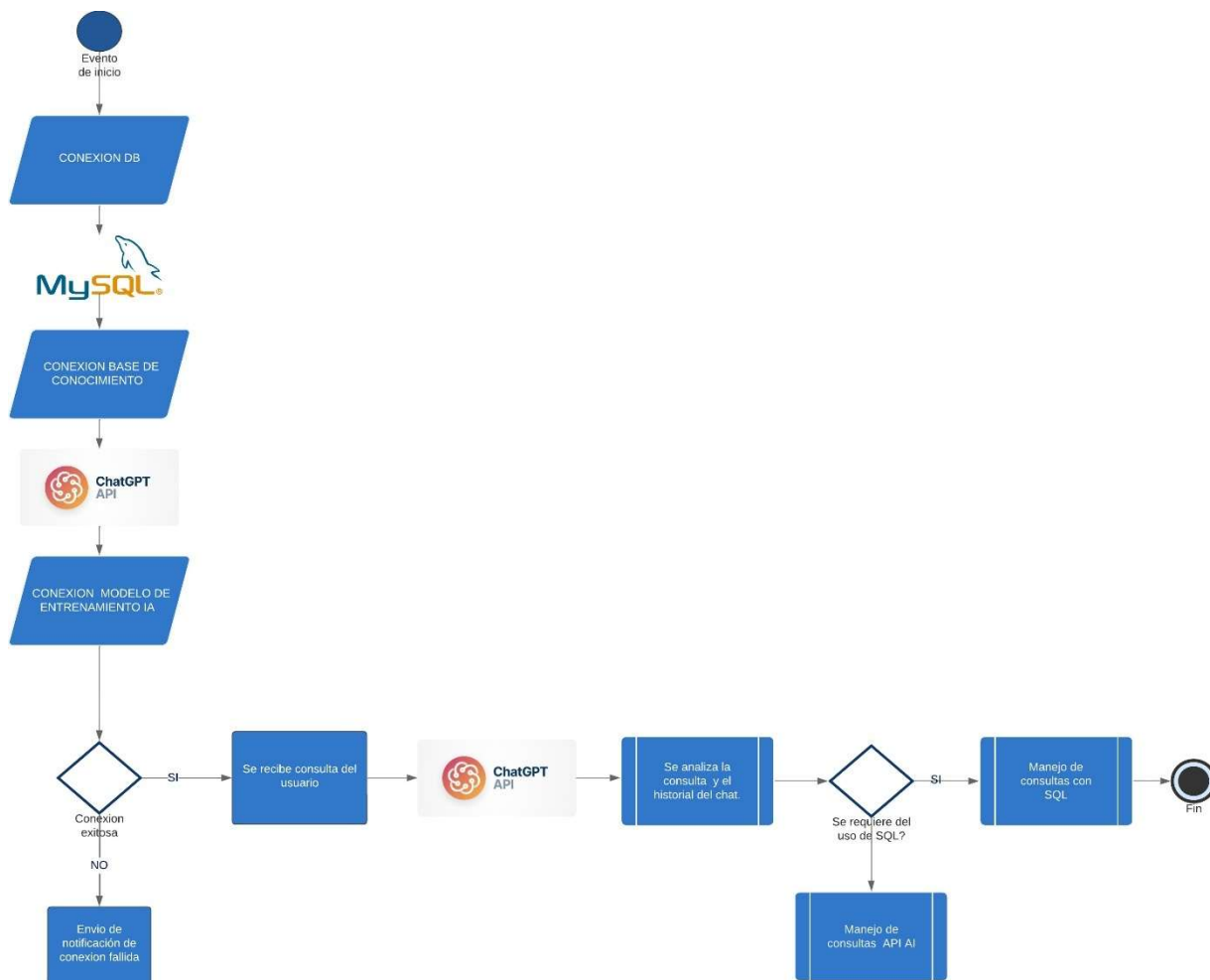
ChatGPT: Integración de la base de conocimiento con inteligencia artificial.

Otra de las herramientas identificadas clave para el desarrollo de esta propuesta es la integración de ChatGPT como herramienta de inteligencia artificial contra la base de conocimiento que se propondrá más adelante en este capítulo. La integración de inteligencia artificial a la base de conocimiento representa una solución innovadora para mejorar la eficiencia y accesibilidad de los datos; como se menciona en el marco teórico, ChatGPT es un modelo de lenguaje avanzado que utiliza inteligencia artificial para interactuar de manera conversacional, permitiendo la comprensión de preguntas, la corrección de errores y la mejora continua a través del aprendizaje interactivo (OpenAI, 2022).

Integrar ChatGPT a la base de conocimiento propuesta busca permitirle a los usuarios interactuar con los datos de una manera más intuitiva, formulando preguntas y obteniendo respuestas detalladas y precisas sobre el estado de las máquinas virtuales y los recursos consumidos de la infraestructura; siendo la intención de facilitar un acceso más eficiente a los colaboradores a dicha información, ya que ChatGPT puede realizar tareas como responder consultas complicadas, proporcionar resúmenes de grandes volúmenes de datos o incluso identificar patrones dentro de estos. Esto es sumamente valioso para los analistas y operadores, quienes podrán obtener respuestas instantáneas sin necesidad de navegar por sistemas complejos o la realización de consultas manuales.

Figura 40

Integración ChatGPT gestión de consultas



Fuente: Elaboración propia, 2024.

Basados en el contexto de la gestión de datos de la infraestructura virtual, esta funcionalidad será esencial para identificar tendencias de consumo de los recursos, detectar posibles problemas de rendimiento o incluso prever necesidades de expansión con respecto a las máquinas virtuales residentes en la plataforma. Además, la API de ChatGPT, como señala ChatGPT.es (2023), también tiene la capacidad de personalizar los modelos de lenguaje para adaptarlos a necesidades específicas, lo que podemos traducir que, mediante la retroalimentación continua,

se pueden entrenar modelos para generar respuestas y análisis más detallados y alineados con las operaciones de Artificial Skills (ChatGPT.es, 2023).

Figura 41

Integración API ChatGPT



Fuente: Elaboración propia, 2024.

El beneficio principal esperado de la integración de la API de ChatGPT en este proyecto se base en no solo optimizar la interacción con la base de conocimiento, sino también busca mejorar el proceso de análisis de los datos; ofreciendo resúmenes automatizados, recomendaciones personalizadas y un análisis constante de las interacciones de los usuarios con la infraestructura; proporcionándole a ASK una herramienta que mejore significativamente la toma de decisiones con datos accesibles en tiempo real, generando decisiones basadas en datos, incrementando la productividad, la confiabilidad y la eficiencia operativa de la organización obteniendo una ventaja competitiva al ofrecérsela a sus diferentes clientes.

Power BI: Analítica y presentación de los datos.

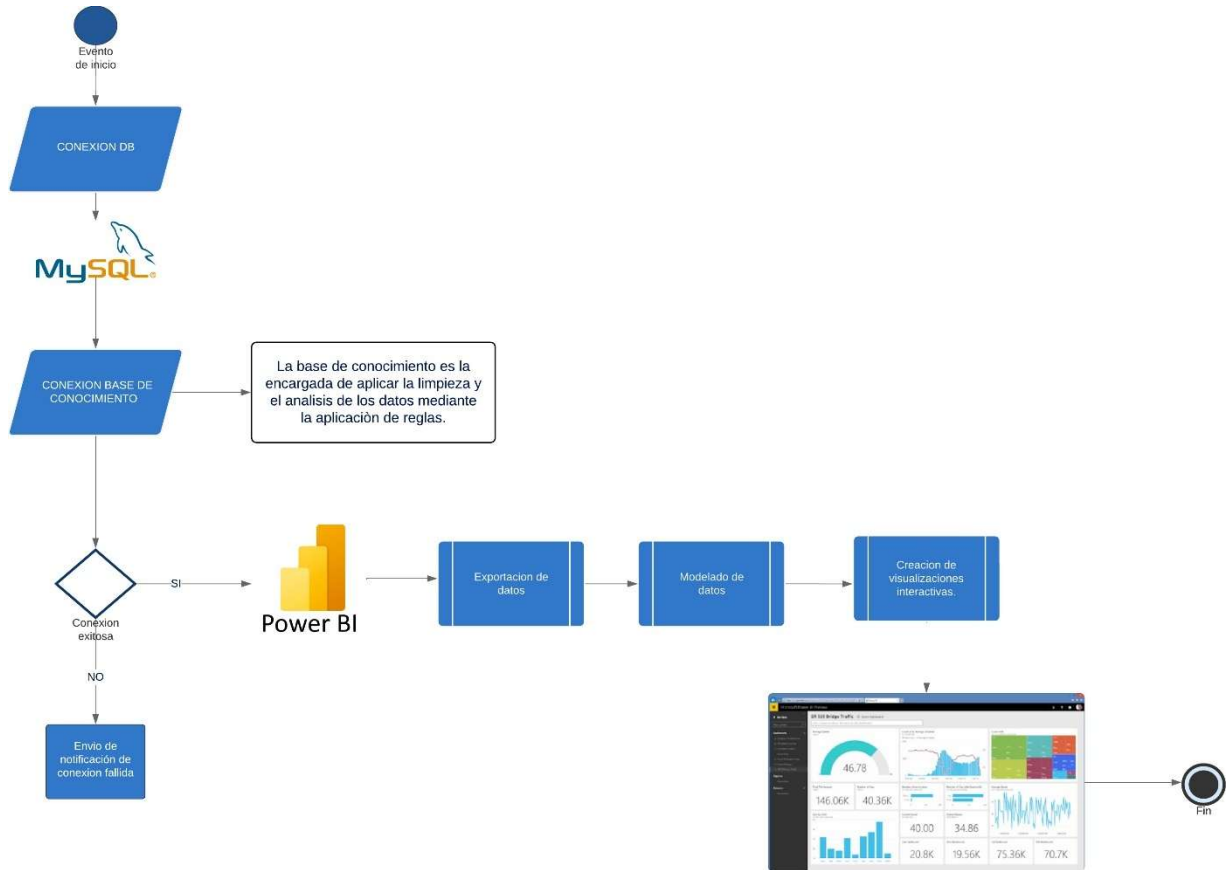
Por último, para el desarrollo de este proyecto, se ha seleccionado la herramienta de Power BI como la herramienta clave para la analítica y visualización de los datos de la infraestructura

virtual de Artificial Skills. Haciendo referencia al marco teórico del presente proyecto, Power BI herramienta desarrollada por Microsoft, es una plataforma que permite transformar los datos complejos en información relevante, visualmente atractiva y fácil de interpretar para los usuarios que la consuman. La misma es de gran apoyo para el desarrollo de este proyecto debido a su capacidad para conectar diversas fuentes de datos, procesar grandes volúmenes de información y generar informes interactivos; convirtiéndola en una solución indispensable para apoyar la toma de decisiones informadas (Microsoft, 2024).

La intención de utilizar Power BI en nuestro proyecto se basa en integrarla como la herramienta central para la creación de paneles de control dinámicos que muestren, de una forma más atractiva visualmente y en tiempo real, los datos recolectados y procesados a través de otras herramientas, como PyVmomi y almacenados en la base de conocimiento. Buscando permitirle a la gerencia tecnológica y a los diferentes colaboradores de Artificial Skills envueltos en el proceso, acceder a los KPIs o indicadores claves de rendimiento (Key performance indicators, por su definición en inglés); identificar patrones y tendencias claras de los datos, realizando un seguimiento continuo de los recursos de infraestructura virtual en tiempo real

Figura 42

Integración PowerBi, visualización de datos



Fuente: Elaboración propia, 2024.

Otra de las ventajas significativas de la integración de Power BI que fue tomada en cuenta para su elección, se fundamenta en su capacidad de combinar datos estructurados y no estructurados, como señala Schlegel y otros (2024); esto resulta fundamental para detectar anomalías, comprender el comportamiento de los diferentes recursos y anticiparse a posibles problemas o incidentes derivados a una mala interpretación de los datos (Schlegel , y otros, 2024). Además, los paneles interactivos que se desarrollen serán diseñados para facilitar la navegación intuitiva,

ofreciendo a los colaboradores con menos conocimientos y a las gerencias la posibilidad de explorar los datos y obtener respuestas clave sobre el estado y desempeño de la infraestructura.

Entre los beneficios esperados de la integración de Power Bi en nuestro proyecto los podemos dividir en 3 pilares fundamentales que se encuentran muy alineados a lo esperado con el desarrollo de esta propuesta:

1. Visualización de los datos en tiempo real.

Todos los datos procesados serán presentados mediante paneles dinámicos que se actualizan automáticamente, lo que permite a los equipos y gerentes tomar decisiones basadas en información actualizada y precisa; a diferencia del proceso actual en la cual deben de generar los gráficos y demás herramientas de visualización de manera manual y bajo demanda; generando retrasos operativos altos.

2. Facilidad de uso.

Power BI es una herramienta que ofrece una interfaz sumamente intuitiva, lo que les permite a los usuarios que la utilicen explorar los datos, realizar análisis personalizados en bases a sus necesidades; lo que en este proyecto es de suma importancia ya que al ASK trabajar con diferentes clientes les permitirá crear paneles en base a los requerimientos de cada cliente. Esto va a permitirle a ASK compartir insights con otros equipos o con los mismos cliente sin necesidad de conocimientos avanzados en analítica de datos, reduciendo significativamente la brecha de conocimiento.

3. Identificación de tendencias en los datos.

La identificación de tendencias en los datos es un pilar altamente significativo para ASK, Power BI permite identificar patrones clave, realizar proyecciones y detectar anomalías en los recursos

de la infraestructura virtual, apoyando no únicamente la toma de decisiones y la automatización de procesos en la realización de informes; sino que a la vez le permite a ASK contar con una ventaja competitiva al poder visualizar el estado de su infraestructura virtual o bien de las máquinas virtuales en tiempo real asegurando una estado de salud optimo de las mismas en cuanto a la utilización de los recursos asignados.

Figura 43

Integración Power BI



Fuente: Elaboración propia, 2024.

Todos estos beneficios esperados anteriormente se pueden respaldar gracias al reconocimiento de Power BI como líder en el Cuadrante Mágico de Gartner en la categoría de plataformas de análisis e inteligencia empresarial (ver figura 10 - Cuadrante mágico de Gartner para plataformas de análisis e inteligencia empresarial) reforzando su relevancia y eficacia como una herramienta estratégica en la analítica de datos. Este prestigioso posicionamiento refleja su capacidad para satisfacer las necesidades de las organizaciones modernas en la interpretación y visualización de datos complejos, destacando la innovación continua de Power BI al proporcionar soluciones avanzadas que habilitan a las empresas para tomar decisiones fundamentadas en insights claros y accesibles (Schlegel , y otros, 2024).

5.1.12 Diseño de la base de conocimiento propuesta.

Seguidamente en este apartado se presenta el diseño conceptual de la base de conocimiento estructurada, siendo esta el principal pilar para el desarrollo de esta propuesta de automatización en la gestión de los datos de la infraestructura virtual de Artificial Skills. Lo que se busca lograr con el desarrollo de esta base de conocimiento es automatizar el proceso de análisis y recolección de los datos extraídos de la plataforma de VMware vCenter y adicional a esto centralizar la información recolectada, garantizando su almacenamiento seguro accesible a lo largo del tiempo y que permita su integración con herramientas de analítica e inteligencia artificial. Esto responde a la necesidad de transformar los datos recolectados en un recurso estratégico para la toma de decisiones, reduciendo al máximo problemas como la duplicidad de esfuerzos y la pérdida de información.

Se propone la utilización de la herramienta de PostgreSQL como sistema gestor de bases de datos, la cual según la página oficial de The PostgreSQL Global Development Group (2024) se define como: “un potente sistema de base de datos relacional de objetos de código abierto con más de 35 años de desarrollo activo que le ha valido una sólida reputación por su confiabilidad, solidez de funciones y rendimiento” (The PostgreSQL Global Development Group, 2024). Entre sus principales beneficios se destaca su compatibilidad con el modelo relacional y el soporte para datos no estructurados, lo que permite almacenar información compleja y diversa.

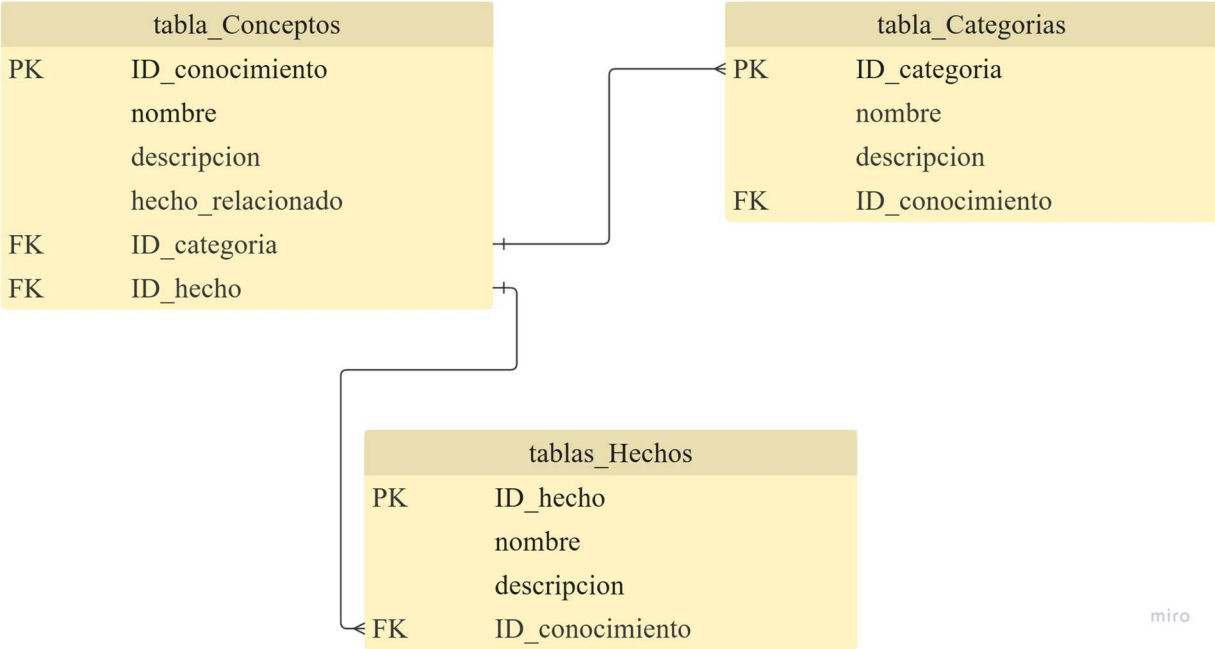
PostgreSQL ofrece una funcionalidad muy útil para el desarrollo de este proyecto como lo es el soporte para transacciones ACID (Atomicidad, Consistencia, Aislamiento y Durabilidad), facilitando la integridad de los datos incluso en escenarios de alta demanda como el que se presenta en Artificial Skills; otra ventaja significativa de PostgreSQL se centra en su flexibilidad, permitiendo la integración con herramientas de analítica avanzada como Power BI,

lo que facilita la generación de informes interactivos basados en datos fiables y actualizados en tiempo real. Por último, es una plataforma escalable de código abierto como se mencionó anteriormente, capaz administrar el gran volumen de datos y cantidad de usuarios, lo que asegura su viabilidad a largo plazo para los requerimientos de Artificial Skills (The PostgreSQL Global Development Group, 2024).

La base de conocimiento que se propone está diseñada para estructurar y organizar la información clave relacionada con la administración y operación de una infraestructura virtual basada en VMware vCenter. Este diseño busca automatizar la gestión del conocimiento al clasificar los diferentes términos necesarios para proporcionar una referencia clara y organizada sobre los conceptos relevantes, primer punto vital para facilitar la comprensión y el uso del conocimiento dentro del sistema.

Figura 44

Diagrama Entidad-Relación Conceptos



Fuente: Elaboración propia, 2024.

Para estructurar y organizar el conocimiento dentro de la base de conocimiento se definen tres categorías principales, la cual cada una de ellas se encuentra enfocada en un aspecto clave en relación con la infraestructura de VMware y su respectiva administración. Entre las categorías seleccionadas podemos encontrar primeramente la categoría de infraestructura (KC_infra), la cual busca incluir los componentes físicos y virtuales dentro de VMware vCenter, como hosts, clústeres, máquinas virtuales etc. Seguidamente la categoría de gestión (KC_gestion) centrándose en la administración de clientes, permisos, configuraciones y asignaciones dentro de vCenter; finalmente la categoría de administración (KC_admin) agrupando procesos relacionados con costos, clientes y departamentos, asegurando una correcta gestión financiera y organizativa.

Tabla 34 *Tabla de categorías.*

ID categoría	Nombre de la categoría	Descripción
KC_infra	Infraestructura	Componentes físicos y virtuales de VMware vCenter.
KC_gestion	Gestión	Administración de clientes, permisos y configuraciones.
KC_admin	Administración	Procesos relacionados con costos, clientes y departamentos.

Fuente: Elaboración propia, 2024.

Posteriormente se propone la tabla de conceptos la cual busca organizar y almacenar información clave sobre los principales conocimientos relacionados con la gestión de infraestructura virtual facilitando el entendimiento de la base de conocimiento. Cada concepto está categorizado según su función dentro del sistema y se asocia con hechos específicos para

facilitar la consulta y el análisis. Entre los conocimientos registrados, se encuentra vCenter (KN_vCenter), que representa la plataforma de gestión de infraestructura virtual y está categorizada dentro de infraestructura, con hechos relacionados en hechos_vCenter y hechosTemporal_Datos. Asimismo, el cliente corporativo (KN_cliente) se refiere a las organizaciones que tienen asignado un vCenter específico y pertenece a la categoría de gestión, con hechos documentados en hechos_Clientes.

Otro concepto clave de nuestro sistema es el clúster (KN_cluster), definido como un grupo de hosts dentro de un mismo data center, clasificado en la categoría de infraestructura y vinculado a hechos_Cluster. Por otro lado, en el ámbito administrativo, la asignación de costos (KN_costos) registra los costos asociados a clientes y servicios, perteneciendo a la categoría de administración y relacionándose con hechos_CentralCosto. Y finalmente definimos las máquinas virtuales (KN_vms) las cuales representan entornos virtualizados que emulan sistemas operativos y hardware sobre hosts ESXi. Estas forman parte de la categoría infraestructura y tienen hechos asociados en hechosTemporal_Datos y hechos_VMs.

Tabla 35 *Tabla de conceptos*

ID Conocimiento	Nombre	Descripción	Categoría	Hecho Relacionado
KN_vCenter	vCenter	Plataforma de gestión de infraestructura virtual.	Infraestructura	hechos_vCenter hechosTemporal_Datos
KN_cliente	Cliente Corporativo	Cliente que tiene asignado un vCenter específico.	Gestión	hechos_Clientes
KN_cluster	Clúster	Grupo de hosts en un mismo data center.	Infraestructura	hechos_Cluster

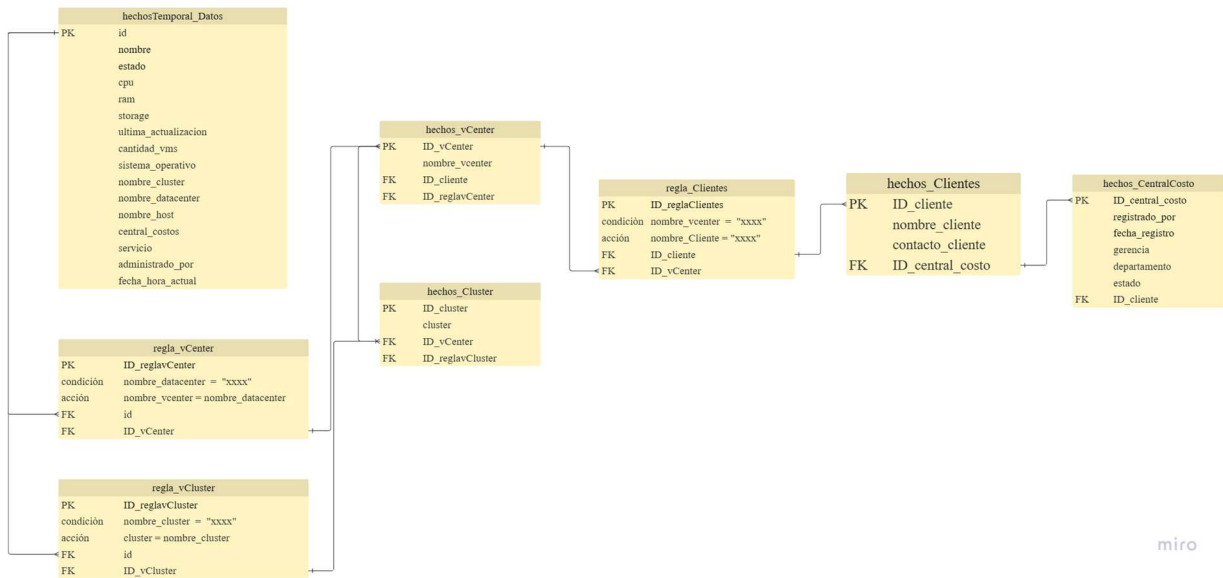
KN_costos	Asignación de Costos	Registro de costos asociados a clientes y servicios.	Administración	hechos_CentralCosto
KN_vms	VM (Virtual Machine)	Entorno virtualizado que emula un sistema operativo y hardware sobre un host ESXi.	Infraestructura	hechosTemporal_Datos hechos_VMs

Fuente: Elaboración propia, 2024.

Una vez el sistema comprenda los conceptos se presentan las tablas de hechos y reglas utilizadas para la gestión y análisis de cada uno de los datos importantes en un entorno de virtualización de VMware vCenter. Estas tablas están diseñadas para estructurar la información obtenida a través de la API mediante la aplicación de diferentes reglas que facilitan el filtrado, análisis y cálculo de costos en la infraestructura virtual. El modelado de la base de conocimiento se organiza en tres etapas clave que representan los principales procesos de gestión de datos que se pretende lograr con la base de conocimiento, a continuación, los mismos:

La primera de las etapas que se plasman en la propuesta es la obtención de datos claves como lo son el vCenter y clúster donde se alojan las diferentes máquinas virtuales de la infraestructura virtualizadas, donde se registran todos los hechos relacionados con la infraestructura, lo que permitirá posteriormente ligar cada una de estas infraestructuras, con un cliente específico.

Tablas obtención de datos vCenter

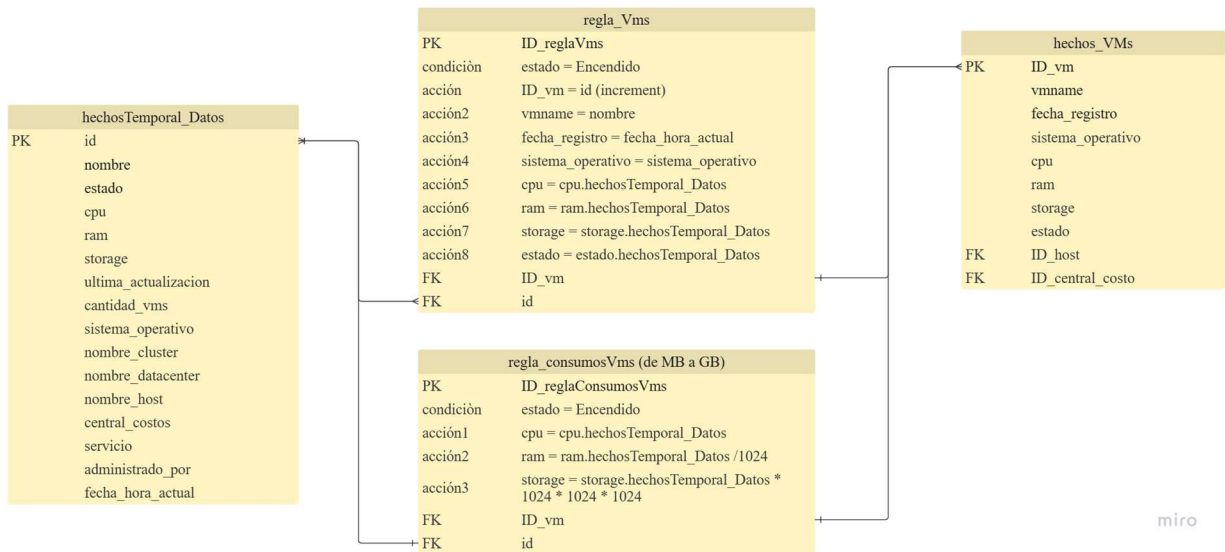


Fuente: Elaboración propia, 2024.

Una vez la base de conocimiento identifique el vCenter sobre el cual se está ejecutando el análisis y lo relacione con su determinado cliente, el sistema debe continuar con el filtrado y análisis de datos de las máquinas virtuales (VMs), lo cual se busca lograr mediante la aplicación de diferentes reglas a nuestra tabla de datos temporal permitiendo extraer únicamente la información relevante para la toma de decisiones. La aplicación de estas permite obtener los datos filtrados para el cálculo de costos de las máquinas virtuales, tomando en cuenta únicamente las que se encuentren en estado de energía “encendido” ya que aquellas que se encuentren apagadas no generan costo alguno que se deba incluir en los reportes; adicionalmente también permite convertir los datos de consumos tanto de RAM, CPU y memoria de MB a GB para facilitar posteriormente el cálculo de los costos.

Figura 46

Tablas filtrado de datos VMs

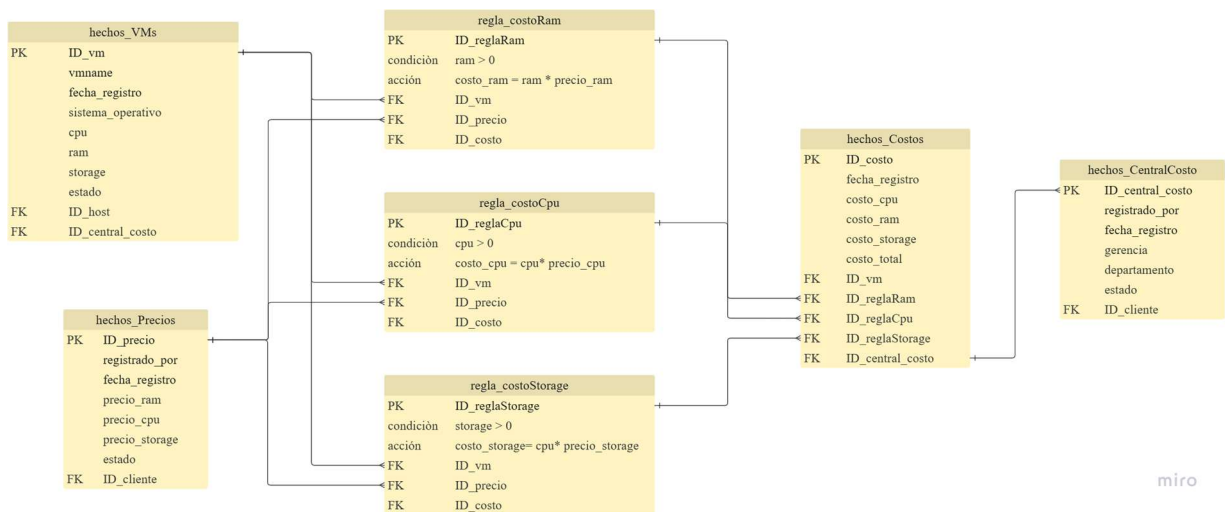


Fuente: Elaboración propia, 2024.

Seguidamente una vez el sistema obtenga los datos de las diferentes máquinas virtuales filtrados, la última de las etapas es el cálculo de costos, donde se propone la aplicación de reglas para determinar el impacto financiero de los recursos utilizados automáticamente mediante nuestra base de conocimiento.

Figura 47

Tablas cálculos costos

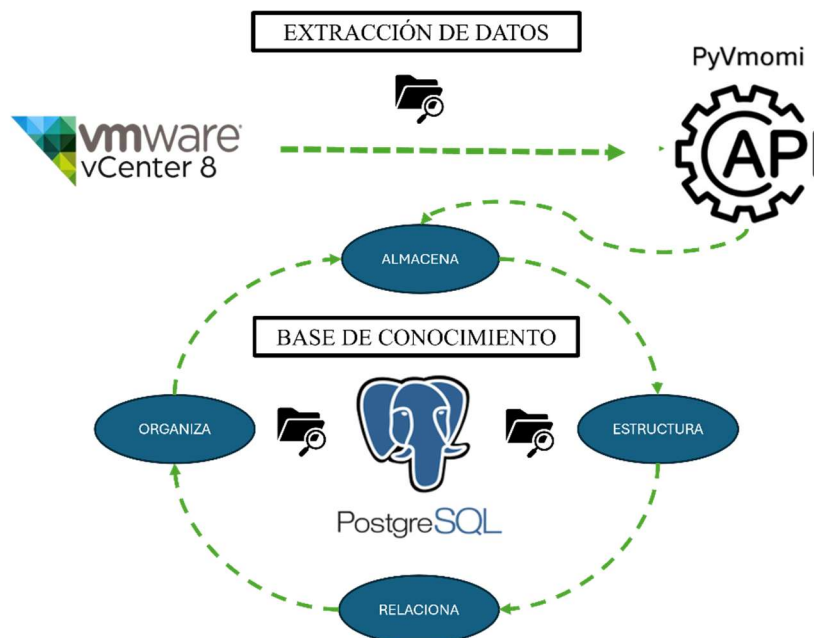


Fuente: Elaboración propia, 2024.

Con la utilización de esta base de conocimiento no únicamente se solventa la problemática actual de Artificial Skills con respecto al almacenamiento seguro y perdurable en el tiempo de los datos de las máquinas virtuales de su infraestructura virtual, lo que actualmente genera pérdida de información y por ende esfuerzos operativos dobles o repetitivos, sino que a la vez permite agilizar en gran manera la gestión completa de los datos, que sumado a otras herramientas como el uso de la API PyVmomi permite un flujo de trabajo mucho más optimizado, reduciendo la gran cantidad de tareas manuales que se ejecutan actualmente que generan riesgos operativos significativos.

Figura 48

Flujo base de conocimiento



Fuente: Elaboración propia, 2024.

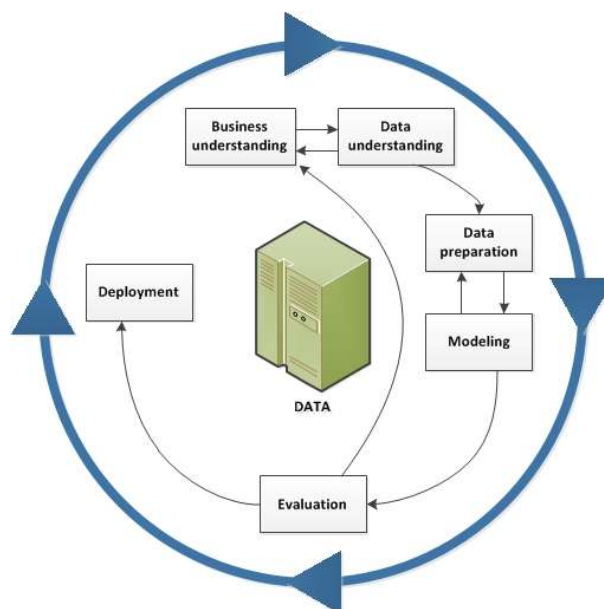
5.1.13 Propuesta de automatización.

A continuación, se presentara el flujo de trabajo optimizado, donde se integran todos los componentes expuestos a lo largo de este proyecto; formando la propuesta de automatización con una base de conocimiento estructurada tal como se menciona en el apartado 5.1.3 e integrando la misma con las herramientas de inteligencia artificial y analítica de datos identificadas en el punto 5.1.2; proponiendo una base integrada para una adecuada gestión de los datos de las máquinas virtuales de la plataforma VMware vCenter en Artificial Skills, que ayude a reducir los tiempos operativos, solventar los riesgos de pérdida información y por ultimo que le ayude a ASK a contar con una base de conocimiento que permita mejorar la toma de decisiones informadas.

Para el desarrollo de esta propuesta de automatización, se propone el empleo del modelo de proceso CRISP-DM (Cross-Industry Standard Process for Data Mining), un modelo ampliamente reconocido en el ámbito de la minería de datos por ofrecer un marco estructurado y efectivo para la gestión de proyectos relacionados con datos, que según IBM (2021) al usarse como un modelo de procesos, nos permite identificar el ciclo vital de la minería de datos, para obtener un máximo provecho de los datos (IBM Corporacion, 2021). Durante las diferentes fases de este modelo de proceso, se utilizarán las diversas herramientas identificadas anteriormente que facilitarán la automatización total o parcial de los procesos.

Figura 49

Ciclo de vida de minería de datos



Fuente: Adaptado de IBM Corporación, 2021, IBM Corporación (<https://www.ibm.com/docs/es/spss-modeler/saas?topic=dm-crisp-help-overview>)

En la primera fase de comprensión del negocio, gracias al diagnóstico de la situación actual realizado en el capítulo IV del presente proyecto se han identificado las áreas críticas donde la automatización generará mayor impacto debido a las necesidades identificadas como la forma actual de recolección, limpieza y almacenamiento de los datos de las máquinas virtuales provenientes de VMware vCenter. Durante la fase de preparación de los datos, se emplearán herramientas como la API de PyVmom, para automatizar la limpieza y estructuración de los grandes volúmenes de información extraídos, asegurando que solo los datos más relevantes sean procesados en las etapas posteriores; buscando optimizar el tiempo invertido por los colaboradores en esta actividad y a la vez mejorar la calidad de los datos, haciendo que los modelos sean más precisos y efectivos.

En las fases de modelado y evaluación, la API seleccionada de ChatGPT, desempeñarán un rol fundamental al automatizar procesos de análisis y cálculos con inteligencia artificial mediante un chatbot, permitiendo una ejecución más rápida y precisa y nuevamente minimizando la intervención humana en el proceso. Finalmente, en la fase de despliegue, los resultados y reportes generados se integrarán directamente en los sistemas internos de Artificial Skills mediante herramientas como Power BI, garantizando una toma de decisiones ágil y basada en datos confiables, eliminando la necesidad de procesos manuales adicionales.

Figura 50

Propuesta base de conocimiento integrada.



Fuente: Elaboración propia, 2024.

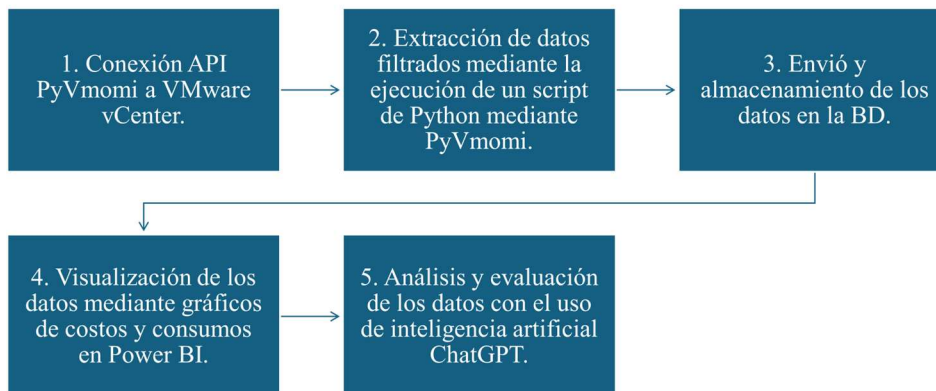
Como se puede observar en la propuesta de integración (ver figura 35 – Propuesta de base de conocimiento integrada) la misma está compuesta por un flujo de datos que involucra varias etapas y herramientas conectadas entre sí para la recolección, almacenamiento, visualización y análisis de la información.

vCenter: Representa la plataforma base de esta propuesta, siendo un servidor o sistema centralizado, que podría estar relacionado con la virtualización o la gestión de entornos virtuales. Siendo una de las herramientas más utilizada para la gestión de infraestructuras virtuales en VMware.

1. API PyVmomi (Preparación de datos): Posteriormente en medio de la plataforma de vCenter y la base de datos se encuentra la API (Interfaz de Programación de Aplicaciones) de PyVmomi, la cual sirve como el puente para comunicar diferentes sistemas, permitiendo el intercambio de datos entre el vCenter y la base de datos. La API se asegura de que los datos puedan ser solicitados o enviados a la base de datos desde el vCenter, que estos sean filtrados automáticamente, asegurándose de la extracción de los datos realmente relevantes únicamente.
2. Base de datos (BD): Aquí es donde se almacenan los datos obtenidos a través de la API. La base de datos guarda la información para que pueda ser posteriormente utilizada o consultada para su posterior análisis a lo largo del tiempo.
3. Visualización de datos (Modelado de los datos): Este componente muestra que la información almacenada en la base de datos puede ser transformada en gráficos o reportes visuales, utilizando herramientas de visualización como Power BI.
4. Inteligencia artificial API ChatGPT (Evaluación): Como ultimo componente se propone el uso de la API de ChatGPT la cual servirá como una herramienta de inteligencia artificial esencial para un análisis más profundo de los datos, fungiendo como un chatbot integrado a la plataforma, permitiéndole a los usuarios realizar consultas sobre las tendencias de los datos de las máquinas virtuales y obtener insights valiosos para la toma de decisiones.

Figura 51

Diagrama de flujo aplicando propuesta de automatización.



Fuente: Elaboración propia, 2024.

En el flujo expuesto podemos ver que la API de PyVmomi es la encargada de realizar una petición API solicitando los datos desde VMware vCenter directamente, por lo general estos datos contienen una gran cantidad de información que en muchos casos excede lo que realmente se necesita; tal como se evidencio en el diagnóstico de la situación actual. Por ello, es importante realizar un filtrado de datos para extraer únicamente los campos esenciales; este proceso de limpieza se puede llevar a cabo mediante un script de Python lenguaje de programación en la que está basado PyVmomi; esto no solo ayuda a reducir el tamaño de los datos almacenados, sino que también contribuye a acelerar las operaciones posteriores, ya que trabajar con un conjunto de datos más pequeño permite consultas más rápidas y menor carga en el almacenamiento.

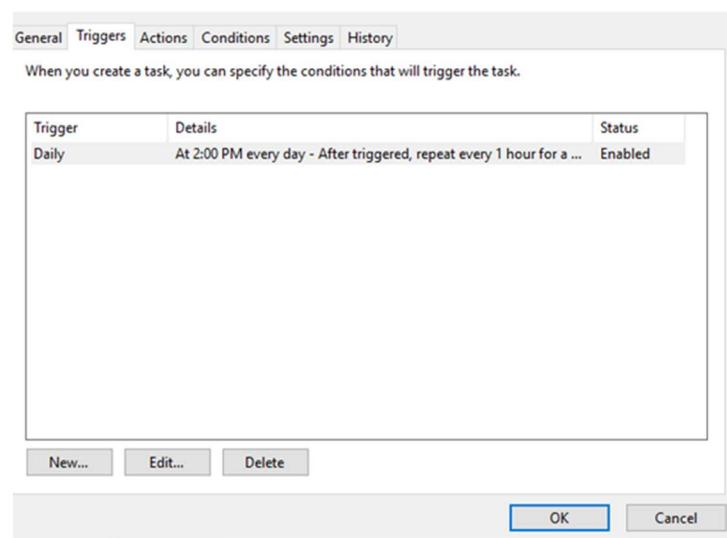
Para evitar que los colaboradores de Artificial Skills (ASK) tengan que ejecutar manualmente tareas repetitivas de carácter rutinario, se propone el uso de la herramienta Task Scheduler de

Windows. Esta solución permite configurar la ejecución automática de procesos, eliminando la necesidad de intervención humana de los colaboradores con respecto al proceso actual. Con Task Scheduler, es posible programar actividades para que se realicen en momentos específicos, adaptándose a las necesidades operativas. Por ejemplo y en base a las necesidades de Artificial Skills, el código puede ejecutarse automáticamente en horarios definidos, ya sea de forma diaria, mensual o bien bajo demanda según lo requerido la gerencia de TI en determinados momentos.

Otra de las principales ventajas de esta herramienta es su flexibilidad, que permite ajustar la programación según los requisitos del sistema sin depender de la disponibilidad de un usuario. Además, es posible añadir condiciones avanzadas, como programar la ejecución en periodos de baja actividad del sistema o en respuesta a eventos específicos, optimizando así el uso de los recursos disponibles; esto permite asegurarse que las tareas se lleven a cabo de forma puntual y confiable, eliminando la intervención manual y reduciendo al mínimo el riesgo de errores humanos.

Figura 52

Ejemplos tareas programadas Windows.

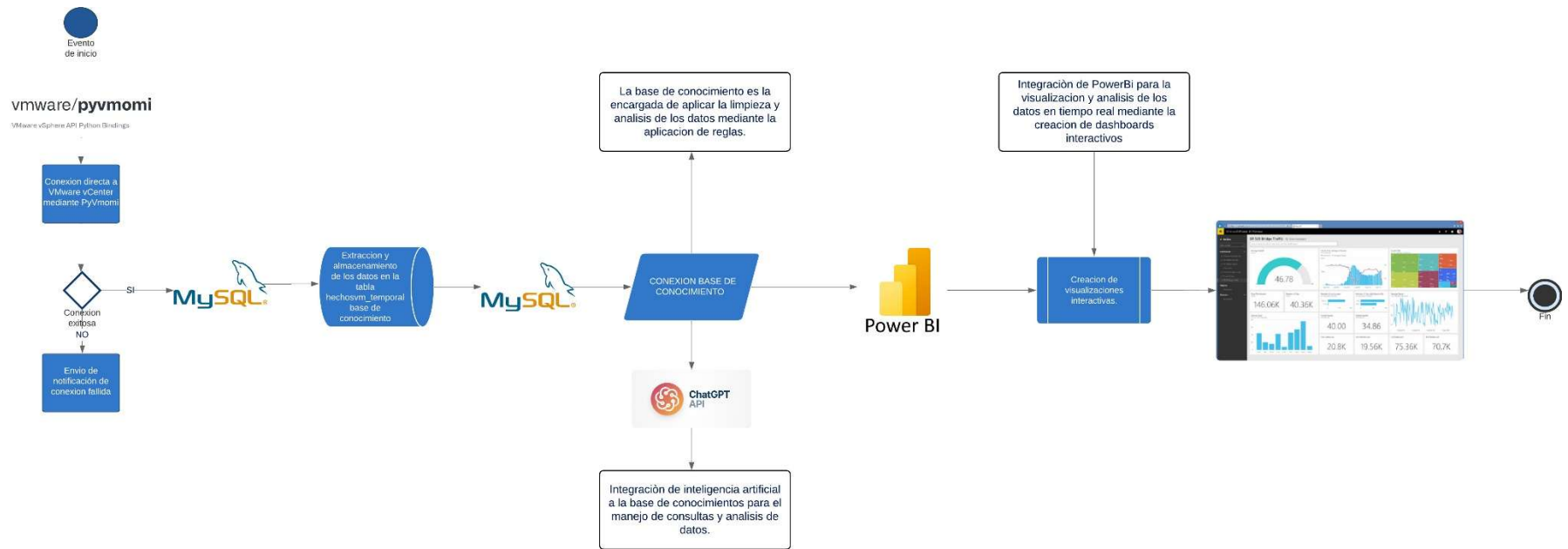


Fuente: Elaboración propia, 2024.

La propuesta de automatización integrada como se mencionó anteriormente busca mejorar la eficiencia operativa, lo que le permite a la vez al personal de Artificial Skills enfocarse en actividades más estratégicas y de mayor valor, en lugar de tareas repetitivas que les consumen un gran porcentaje de su tiempo operativo. Asimismo, la programación constante y regular de esta propuesta garantiza la continuidad de las operaciones de TI, evitando interrupciones en el servicio y asegurando el cumplimiento de los plazos de entrega de informes ya sea con la gerencia de TI o bien con sus clientes, contribuyendo a la mejora continua del flujo de trabajo.

Figura 53

Diagrama de flujo automatización mediante base de conocimiento



Fuente: Elaboración propia, 2024.

CAPÍTULO VI: CONCLUSIONES Y RECOMENDACIONES

En este último capítulo final se presentan las conclusiones y recomendaciones derivadas del desarrollo de este proyecto, con el objetivo de brindar una visión clara y fundamentada de los resultados obtenidos y las oportunidades de mejora identificadas. Las conclusiones representan ideas precisas y contundentes que permiten definir el grado de cumplimiento de los objetivos planteados al inicio del proyecto, destacando los logros alcanzados, las decisiones clave tomadas y los hallazgos más relevantes que aportan valor a Artificial Skills (ASK). Por otro lado, las recomendaciones proponen acciones adicionales necesarias para fortalecer y complementar la implementación de esta propuesta.

6.1 CONCLUSIONES

6.1.1 Conclusión 1: Análisis del proceso actual.

Se determina que en el proceso actual de gestión de los datos de las máquinas virtuales de la plataforma de VMware vCenter en ASK presenta deficiencias significativas. El análisis del proceso actual de gestión de los datos reveló deficiencias significativas, principalmente debido al desconocimiento del concepto y aplicación de la automatización, sumado a la ausencia de herramientas tecnológicas adecuadas que faciliten la ejecución del proceso. Adicionalmente se identificó una alta intervención manual en la recolección y análisis de datos, lo que genera retrasos en su disponibilidad y aumenta el riesgo de errores en la toma de decisiones, afectando negativamente la eficiencia operativa de ASK. Gracias a este análisis se percibe la necesidad de optimizar estos procesos para mejorar la gestión de información estratégica.

6.1.2 Identificación herramientas de inteligencia artificial y analítica de datos.

Se concluye que en ASK una de las principales problemáticas con respecto al proceso actual es la utilización de herramientas básicas que si bien es cierto cumplen con el objetivo no son las

óptimas para una eficiencia operativa real, es por lo que gracias a la presente investigación se encuentra y determina un conjunto de herramientas que pueden ejercer un papel clave para la automatización de la gestión de datos en VMware vCenter, específicamente las herramientas de PyVmomi, ChatGPT y Power BI.

Estas herramientas según el análisis demuestran teóricamente que cumplen con los requerimientos técnicos del proyecto y ofrecen funcionalidades esenciales para la automatización de tareas repetitivas, la generación de reportes dinámicos y el análisis predictivo, que le permitan a ASK mejorar no solo su eficiencia operativa sino a la vez la toma de decisiones gerenciales. Otro hallazgo importante con respecto a estas herramientas es su capacidad de integración con la infraestructura tecnológica de ASK las cuales cuentan con versiones gratuitas que permiten su implementación sin una inversión inicial, con la opción de versiones pagas que ofrecen funcionalidades avanzadas.

6.1.3 Base de conocimiento estructurada.

Se concluye que la implementación de una base de conocimiento centralizada permitirá gestionar eficientemente los datos de las máquinas virtuales en VMware vCenter, asegurando su disponibilidad, trazabilidad y organización. Actualmente, la ausencia de un sistema estructurado dificulta el acceso a información clave y genera problemas como la duplicidad de datos y la pérdida de conocimiento a lo largo del tiempo lo que nos permitió ejecutar un estudio el cual demuestra que la aplicación de un modelo basado en el modelo CRISP-DM proporcionaría una metodología estructurada para la recopilación, almacenamiento y análisis de los datos, facilitando su procesamiento y optimización continua.

6.1.4 Propuesta de integración de inteligencia artificial y analítica de datos.

El estudio determinó que la integración de inteligencia artificial y analítica de datos en la gestión de VMware vCenter representa una oportunidad clave para optimizar la eficiencia operativa y mejorar la toma de decisiones en ASK. Se identificó que una solución basada en inteligencia artificial permitiría procesar y analizar grandes volúmenes de datos de manera automatizada, facilitando la identificación de patrones y tendencias en el uso de recursos. Asimismo, la aplicación de una base de conocimiento con un enfoque estructurado, como el modelo CRISP-DM, permitiría gestionar la implementación de estas tecnologías de manera ordenada y eficiente, asegurando que los procesos de extracción, análisis y visualización de datos estén alineados con las necesidades estratégicas de la empresa.

6.2 RECOMENDACIONES

Las recomendaciones presentadas en este capítulo están orientadas a garantizar el éxito de la implementación de la propuesta de automatización y a maximizar los beneficios obtenidos por Artificial Skills (ASK). Estas sugerencias surgen a partir de los hallazgos y limitaciones identificados durante el desarrollo del proyecto, así como de elementos que no se abordaron dentro del alcance inicial pero que resultan relevantes para el futuro de la organización.

El objetivo de estas recomendaciones es proporcionar a ASK una guía práctica que permita optimizar la ejecución, mejorar la adopción de las herramientas tecnológicas propuestas y asegurar la sostenibilidad de los resultados a largo plazo, fomentando así una mejora continua en la gestión de datos y procesos operativos.

6.2.1 Implementación de la propuesta

Se recomienda a Artificial Skills implementar progresivamente la automatización en los procesos de gestión de datos en VMware vCenter, priorizando aquellos procesos en los que existe una mayor intervención manual. Como primer paso, se sugiere la creación de una base de datos centralizada en un motor de base de datos como PostgreSQL para garantizar un almacenamiento estructurado y seguro de la información crítica. Posteriormente, se recomienda integrar herramientas como PyVmmomi para la extracción automatizada de datos, reduciendo la dependencia de procesos manuales; el uso de inteligencia artificial, como ChatGPT que les permita analizar ampliamente la información generada y finalmente sumado a herramientas de análisis de datos como Power BI la cual facilitará la visualización y generación de reportes estratégicos, mejorando la toma de decisiones y la eficiencia operativa en general.

6.2.2 Implementación y capacitación en herramientas de inteligencia artificial y analítica de datos

Como segunda recomendación a ASK la implementación progresiva de las herramientas identificadas de inteligencia artificial y analítica de datos como PyVmmomi, ChatGPT y Power BI; con el objetivo de optimizar mayormente la gestión de datos en VMware vCenter y mejorar la eficiencia operativa. Se sugiere integrar PyVmmomi para la automatización de la extracción de datos, el punto de dolencia más importante del proceso actual, reduciendo la intervención manual y garantizando una mayor precisión y disponibilidad en tiempo real de los datos; seguidamente, se recomienda implementar ChatGPT como apoyo en el análisis e interpretación de la información, facilitando la identificación de patrones y la generación de respuestas automatizadas a las necesidades de las diferentes áreas. Finalmente, se propone incorporar

Power BI para la visualización de datos estratégicos, permitiendo la creación de reportes dinámicos y análisis predictivo que respalden la toma de decisiones gerenciales.

Dado que estas herramientas cuentan con versiones gratuitas, ASK puede iniciar su implementación sin una inversión inicial significativa. Para maximizar su aprovechamiento, se recomienda capacitar al equipo técnico en la configuración y uso de cada herramienta, asegurando su correcta integración con la infraestructura tecnológica actual, permitiendo que ASK no solo automatice procesos clave, sino que también fortalezca su capacidad de análisis y toma de decisiones estratégicas.

6.2.3 Implementación de una base de conocimiento centralizada

Se recomienda a Artificial Skills desarrollar e implementar una base de conocimiento centralizada para optimizar la gestión de los datos en VMware vCenter, en la cual se sugiere adoptar un modelo basado en la metodología CRISP-DM, permitiendo un proceso centralizado completo desde la recopilación, procesamiento y análisis estructurado de los datos de la plataforma. Además, se recomienda establecer controles de calidad de la información y procesos de actualización continua para evitar una duplicidad y pérdida de datos críticos que ayuden a brindar una mayor confiabilidad al proceso.

6.2.4 Propuesta de integración

Se sugiere a ASK desarrollar un sistema automatizado para la recopilación y procesamiento de datos, basados en una base de conocimiento centralizada que asegure que la información sea precisa y accesible en tiempo real que sumado a un mecanismo de análisis predictivo basado en inteligencia artificial permita identificar patrones y optimizar la asignación de recursos tecnológicos obteniendo un máximo provecho de los datos para la toma de decisiones

estratégicas. Finalmente, la visualización de datos debe ser mejorada mediante la aplicación e integración de herramientas de analítica avanzada como PowerBi, que le permitan a los gerentes o encargados de la toma de decisiones estratégicas dentro de la empresa tener información valiosa de una forma accesible y entendible.

BIBLIOGRAFÍA

(s.f.). Obtenido de ARTIFICIAL SKILLS: <https://artificial-skills.com/nosotros>

Arboleda Mazo, W. H., & Orozco Carvaja, L. J. (2018). Big Data, herramienta para el desarrollo empresarial. *UNAC*, 85-93.

Bonilla Botia, L., & Briceño Díaz, F. A. (junio de 2006). Sistemas de Información como apoyo a la toma de decisiones. *Sistema de Información Científica Redalyc*, 53-57. Recuperado el 30 de octubre de 2024

Breton, L. (2024). *Virtualización con VMware vSphere 8 Fundamentos*. Madrid: Ediciones ENI.

Broadcom Inc. (10 de octubre de 2005-2024). *VMWARE by Broadcom*. Obtenido de VMWARE by Broadcom: <https://www.vmware.com/products/cloud-infrastructure/vcenter>

Broadcom Inc. (04 de abril de 2023). *Broadcom Inc*. Obtenido de Broadcom Inc: <https://docs.vmware.com/en/VMware-vSphere/8.0/vsphere-apis-sdks-getting-started/GUID-729A9C47-60EA-4571-B7B0-7C98B7BA40DF.html>

ChatGPT.es. (8 de mayo de 2023). *ChatGPT.es*. Obtenido de ChatGPT.es: <https://chatgpt.es/api-de-chatgpt-y-sus-posibles-usos/>

Garrell, A., & Guilera, L. (2019). *La Industria 4.0 en la sociedad digital*. Valencia: Marge Books.

- Gómez López, J., Villar Fernández, E., & Alcayde García, A. (2011). *Seguridad en sistemas operativos Windows y GNU/Linux*. Paracuellos del Jarama, Madrid: Grupo Editorial RA-MA.
- Goodwin, M. (9 de abril de 2024). *IBM*. Obtenido de IBM: <https://www.ibm.com/mx-es/topics/api>
- Groover, M. (2014). *Automation, Production Systems, and Computer-Integrated Manufacturing*. Pensilvania: Pearson College Div.
- IBM. (18 de Julio de 2023). *International Business Machines Corporation*. Obtenido de IBM: <https://www.ibm.com/mx-es/topics/virtualization>
- IBM Corporacion. (08 de 17 de 2021). *IBM Corporacion*. Obtenido de IBM Corporacion: <https://www.ibm.com/docs/es/spss-modeler/saas?topic=dm-crisp-help-overview>
- Mamani, Y. (12 de marzo de 2018). *Researchgate*. Obtenido de Researchgate: https://www.researchgate.net/publication/323993348_Business_Intelligenceherramientas_para_la_toma_de_decisiones_en_procesos_de_negocio
- Marchionni, E. A., & Formoso, O. M. (2013). *Virtualización con VMware: Manuales Users*. Buenos Aires, Argentina: Creative Andina Corp.
- MEDINA ROMERO, M., ROJAS LEÓN, R., BUSTAMANTE HOCES, W., LOAIZA CARRASCO, R., MARTEL CARRANZA, C., & CASTILLO ACOBO, R. (2023). *Metodología de la investigación: Técnicas e instrumentos de investigación*. Urb. Ciudad Jardín Mz. B3 Lt. 2, Puno – Perú: Instituto Universitario de Innovación Ciencia y Tecnología Inudi Perú S.A.C.

Microsoft. (22 de 03 de 2024). *Microsoft Learn*. Obtenido de Microsoft Learn:
<https://learn.microsoft.com/es-es/power-bi/fundamentals/power-bi-overview>

Olavsrud, T. (30 de 06 de 2024). *IDG Communications Inc*. Obtenido de CIO:
<https://www.cio.com/article/1314134/sistemas-de-apoyo-a-la-toma-de-decisiones-la-herramienta-clave-para-el-analisis-de-datos-empresariales.html>

OpenAI. (30 de noviembre de 2022). *OpenAI*. Obtenido de OpenAI:
<https://openai.com/index/chatgpt/>

Oracle Corporation. (20 de 05 de 2024). *Oracle Corporation*. Obtenido de Oracle Corporation:
<https://www.oracle.com/mx/business-analytics/what-is-analytics/>

Ponce Gallegos, J. C., Torres Soto, A., Quezada Aguilera, F. S., Silva Sprock, A., Martínez Flor, E. U., Casali, A., . . . Pedreño, O. (2014). *Inteligencia Artificial*.
Morales: Iniciativa Latinoamericana de Libros de Texto Abiertos (LATIn).

Red Hat. (01 de 20 de 2023). *Red Hat Inc*. Obtenido de Red Hat Inc:
<https://www.redhat.com/es/topics/api/what-are-application-programming-interfaces>

Richardson, D. (23 de febrero de 2022). *Red Hat Inc*. Obtenido de Red Hat:
https://www.redhat.com/es/blog/industry-4-and-ot-transformation-technologies-benefits-and-challenges?sc_cid=7015Y0000048RsyQAE&gad_source=1&gclid=Cj0KCCQjw4Oe4BhCcARIsADQ0cskO5rg-l8rpozOHKr1Q-CG9B7FI_zRuP99_eO0CpYDYAOb1D6HyAT0aAmBIEALw_wcB

Rouhiainen, L. (2018). *Inteligencia artificial: 101 cosas que debes saber hoy sobre nuestro futuro*. Madrid: Planeta, S.A.

Sampieri, R. H. (2014). *Metodología de la investigación* (Sexta ed.). Iztapalapa. Mexico D. F.: McGraw-Hill Interamericana.

Schlegel, K., Ganeshan, A., Pidsley, D., Sun, J., O'Callaghan, G., Long, C., . . . O'Brien, J. (2024 de junio de 2024). *Gartner Inc*. Obtenido de Gartner: <https://www.gartner.com/doc/reprints?id=1-2HVUGEM6&ct=240620&st=sb>

SHARDA, R., DELEN, D., & TURBAN, E. (2013). *Business Intelligence and Analytics: Systems for Decision Support*. New Jersey: Pearson.

The PostgreSQL Global Development Group. (20 de 10 de 2024). *Postgresql Org*. Obtenido de Postgresql Org: <https://www.postgresql.org/>