



**UNIVERSIDAD AUTÓNOMA
DE AGUASCALIENTES**

**Centro de Ciencias Básicas
Departamento de Sistemas de Información**

**Diseño, Implementación y Evaluación Ágil de un Servicio de
TI -con Código Abierto-para Ciencia de Datos:**

Caso Prototipo para PYMES.

Trabajo Práctico que Presenta:

Juan Daniel Gallegos Rodríguez

Para Optar por el Grado de:

Maestría en Informática y Tecnologías Computacionales

Tutores

Dr. José Manuel Mora Tavarez

Dra. Paola Yuritzy Reyes Delgado

Integrantes del Comité Tutorial

MITC. Jorge Eduardo Macias Luevano

Aguascalientes, 26 de noviembre de 2024

CARTA DE VOTO APROBATORIO
INDIVIDUAL

M. EN C. JORGE MARTÍN ALFÉREZ CHÁVEZ
DECANO DEL CENTRO DE CIENCIAS BÁSICAS

PRESENTE

Por medio del presente como TUTOR designado del estudiante **JUAN DANIEL GALLEGOS RODRÍGUEZ** con ID **161692** quien realizó el trabajo práctico titulado: **Diseño, Implementación y Evaluación Ágil de un Servicio de TI -con Código Abierto-para Ciencia de Datos: Caso Prototipo para PYMES**, un trabajo propio, innovador, relevante e inédito y con fundamento en el Artículo 175, Apartado II del Reglamento General de Docencia doy mi consentimiento de que la versión final del documento ha sido revisada y las correcciones se han incorporado apropiadamente, por lo que me permito emitir el **VOTO APROBATORIO**, para que él pueda proceder a imprimirlo así como continuar con el procedimiento administrativo para la obtención del grado.

Pongo lo anterior a su digna consideración y sin otro particular por el momento, me permito enviarle un cordial saludo.

ATENTAMENTE
"Se Lumen Proferre"

Aguascalientes, Ags., a 26 día de noviembre de 2024.

DR. JOSÉ MANUEL MORA TAVAREZ
Tutor de trabajo práctico

El nombre completo que aparece en el Voto Aprobatorio debe coincidir con el que aparece en el documento pdf. No se puede abreviar, ni omitir nombres

c.c.p.- Interesado
c.c.p.- Secretaría Técnica del Programa de Posgrado

Elaborado por: Depto. Apoyo al Posgrado.
Revisado por: Depto. Control Escolar/Depto. Gestión de Calidad.
Aprobado por: Depto. Control Escolar/ Depto. Apoyo al Posgrado.

Código: DO-SEE-FO-07
Actualización: 01
Emisión: 17/05/19

M. EN C. JORGE MARTÍN ALFÉREZ CHÁVEZ
DECANO DEL CENTRO DE CIENCIAS BÁSICAS

PRESENTE

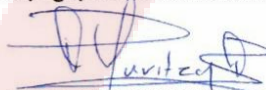
Por medio del presente como TUTORA designada del estudiante **JUAN DANIEL GALLEGOS RODRÍGUEZ** con ID **161692** quien realizó el trabajo práctico titulado: **Diseño, Implementación y Evaluación Ágil de un Servicio de TI -con Código Abierto-para Ciencia de Datos: Caso Prototipo para PYMES**, un trabajo propio, innovador, relevante e inédito y con fundamento en el Artículo 175, Apartado II del Reglamento General de Docencia doy mi consentimiento de que la versión final del documento ha sido revisada y las correcciones se han incorporado apropiadamente, por lo que me permito emitir el **VOTO APROBATORIO**, para que él pueda proceder a imprimirlo así como continuar con el procedimiento administrativo para la obtención del grado.

Pongo lo anterior a su digna consideración y sin otro particular por el momento, me permito enviarle un cordial saludo.

ATENTAMENTE

"Se Lumen Proferre"

Aguascalientes, Ags., a 26 día de noviembre de 2024.



DRA. PAOLA YURITZY REYES DELGADO
Tutora de trabajo práctico

El nombre completo que aparece en el Voto Aprobatorio debe coincidir con el que aparece en el documento pdf. No se puede abreviar, ni omitir nombres

c.c.p.- Interesado

c.c.p.- Secretaría Técnica del Programa de Posgrado

Elaborado por: Depto. Apoyo al Posgrado.
Revisado por: Depto. Control Escolar/Depto. Gestión de Calidad.
Aprobado por: Depto. Control Escolar/ Depto. Apoyo al Posgrado.

Código: DO-SEE-FO-07
Actualización: 01
Emisión: 17/05/19

CARTA DE VOTO APROBATORIO
INDIVIDUAL

M. EN C. JORGE MARTÍN ALFÉREZ CHÁVEZ
DECANO DEL CENTRO DE CIENCIAS BÁSICAS

PRESENTE

Por medio del presente como MIEMBRO DEL COMITÉ TUTORIAL designado del estudiante **JUAN DANIEL GALLEGOS RODRÍGUEZ** con ID **161692** quien realizó el trabajo práctico titulado: **Diseño, Implementación y Evaluación Ágil de un Servicio de TI -con Código Abierto-para Ciencia de Datos: Caso Prototipo para PYMES**, un trabajo propio, innovador, relevante e inédito y con fundamento en el Artículo 175, Apartado II del Reglamento General de Docencia doy mi consentimiento de que la versión final del documento ha sido revisada y las correcciones se han incorporado apropiadamente, por lo que me permito emitir el **VOTO APROBATORIO**, para que él pueda proceder a imprimirlo así como continuar con el procedimiento administrativo para la obtención del grado.

Pongo lo anterior a su digna consideración y sin otro particular por el momento, me permito enviarle un cordial saludo.

ATENTAMENTE
"Se Lumen Proferre"

Aguascalientes, Ags., a 26 día de noviembre de 2024.

MITC. JORGE EDUARDO MACIAS LUEVANO
Asesor de trabajo práctico

El nombre completo que aparece en el Voto Aprobatorio debe coincidir con el que aparece en el documento pdf. No se puede abreviar, ni omitir nombres

c.c.p.- Interesado
c.c.p.- Secretaría Técnica del Programa de Posgrado

Elaborado por: Depto. Apoyo al Posgrado.
Revisado por: Depto. Control Escolar/Depto. Gestión de Calidad.
Aprobado por: Depto. Control Escolar/ Depto. Apoyo al Posgrado.

Código: DO-SEE-FO-07
Actualización: 01
Emisión: 17/05/19



DICTAMEN DE LIBERACIÓN ACADÉMICA PARA INICIAR LOS TRÁMITES DEL EXAMEN DE GRADO



Fecha de dictaminación dd/mm/aaaa: 25/11/2024

NOMBRE: Juan Daniel Gallegos Rodríguez ID: 161692

PROGRAMA: Maestría en Informática y Tecnologías Computacionales LGAC (del posgrado): Gestión de Sistemas y Tecnologías de Información para mejorar competitividad, innovación y cambio organizacional

TIPO DE TRABAJO: () Tesis (X) Trabajo Práctico

TITULO: Diseño, Implementación y Evaluación Ágil de un Servicio de TI -con Código Abierto para Ciencia de Datos: Caso Prototipo para PYMES.

IMPACTO SOCIAL (señalar el impacto logrado): a los altos costos implementarios (recursos de hardware, software, y personal experto). Afortunadamente, existen plataformas de Ciencia de Datos de tipo Código Abierto (no costos de licencias para instalarlas y usarlas), pero su selección correcta y su proceso de instalación y uso demandan conocimiento experto.

INDICAR	SI	NO	N.A.	(NO APLICA)	SEGÚN	CORRESPONDA:
Elementos para la revisión académica del trabajo de tesis o trabajo práctico:						
SI						El trabajo es congruente con las LGAC del programa de posgrado
SI						La problemática fue abordada desde un enfoque multidisciplinario
SI						Existe coherencia, continuidad y orden lógico del tema central con cada apartado
SI						Los resultados del trabajo dan respuesta a las preguntas de investigación o a la problemática que aborda
SI						Los resultados presentados en el trabajo son de gran relevancia científica, tecnológica o profesional según el área
SI						El trabajo demuestra más de una aportación original al conocimiento de su área
SI						Las aportaciones responden a los problemas prioritarios del país
SI						Generó transferencia del conocimiento o tecnológica
SI						Cumple con la ética para la investigación (reporte de la herramienta antiplagio)
El egresado cumple con lo siguiente:						
SI						Cumple con lo señalado por el Reglamento General de Docencia
SI						Cumple con los requisitos señalados en el plan de estudios (créditos curriculares, optativos, actividades complementarias, estancia, predoctoral, etc)
SI						Cuenta con los votos aprobatorios del comité tutorial, en caso de los posgrados profesionales si tiene solo tutor podrá liberar solo el tutor
SI						Cuenta con la carta de satisfacción del Usuario
SI						Coincide con el título y objetivo registrado
SI						Tiene congruencia con cuerpos académicos
SI						Tiene el CVU del Conacyt actualizado
N.A.						Tiene el artículo aceptado o publicado y cumple con los requisitos institucionales (en caso que proceda)
En caso de Tesis por artículos científicos publicados						
N.A.						Aceptación o Publicación de los artículos según el nivel del programa
N.A.						El estudiante es el primer autor
N.A.						El autor de correspondencia es el Tutor del Núcleo Académico Básico
N.A.						En los artículos se ven reflejados los objetivos de la tesis, ya que son producto de este trabajo de investigación.
N.A.						Los artículos integran los capítulos de la tesis y se presentan en el idioma en que fueron publicados
N.A.						La aceptación o publicación de los artículos en revistas indexadas de alto impacto

Con base a estos criterios, se autoriza se continúen con los trámites de titulación y programación del examen de grado: Sí X No

Elaboró: FIRMAS
* NOMBRE Y FIRMA DEL CONSEJERO SEGÚN LA LGAC DE ADSCRIPCION: DR. CARLOS ARGELIO ARÉVALO MERCADO
NOMBRE Y FIRMA DEL SECRETARIO TÉCNICO: DRA. LIZETH ITZIGUERY SOLANO ROMO
* En caso de conflicto de intereses, firmará un revisor miembro del NAB de la LGAC correspondiente distinto al tutor o miembro del comité tutorial, designado por el Decano
Revisó: DR. ALEJANDRO PADILLA DÍAZ
NOMBRE Y FIRMA DEL SECRETARIO DE INVESTIGACIÓN Y POSGRADO:
Autorizó: M. EN C. JORGE MARTÍN ALFÉREZ CHÁVEZ
NOMBRE Y FIRMA DEL DECANO:

Nota: procede el trámite para el Depto. de Apoyo al Posgrado
En cumplimiento con el Art. 105C del Reglamento General de Docencia que a la letra señala entre las funciones del Consejo Académico: ... Cuidar la eficiencia terminal del programa de posgrado y el Art. 105F las funciones del Secretario Técnico, llevar el seguimiento de los alumnos.
Elaborado por: D. Apoyo al Posg.
Revisado por: D. Control Escolar/D. Gestión de Calidad.
Aprobado por: D. Control Escolar/ D. Apoyo al Posg.
Código: DO-SEE-FO-15
Actualización: 01
Emisión: 28/04/20



SITIS DE AGUASCALIENTES, S.C.
Mar Caribe 203-B Int. 5, Fracc. Las Brisas
Aguascalientes, Ags.
Tels. 449-912-1085 y 449-162-5786
e-mail: jemlsitis@gmail.com
sitisags@gmail.com

Aguascalientes, Ags. a 21 de Junio de 2024.

Dr. José Manuel Mora Tavaréz
Director de Tesis
Presente:

Estimado Dr. Mora: Reciba un cordial saludo de parte nuestra. Por medio de la presente me permito informarle que aceptamos el resultado del proyecto práctico "Diseño, Implementación y Evaluación Ágil de un Servicio de TI -con Código Abierto- para Ciencia de Datos: Caso Prototipo para PYMES" que realizó el alumno de la Maestría en Informática y Tecnologías Computacionales Juan Daniel Gallegos Rodríguez, en nuestra organización; el cual fue de mucha ayuda para apoyarnos en el análisis de datos históricos de nuestros sistemas.

Quedo a sus órdenes para cualquier duda respecto a la presente, en los teléfonos y correo arriba mencionado. Sin más por el momento queda de usted su servidor:

Atentamente

Ing. Jorge Eduardo Macías Luévano
Director General
SITIS de Aguascalientes, S.C.

Agradecimientos

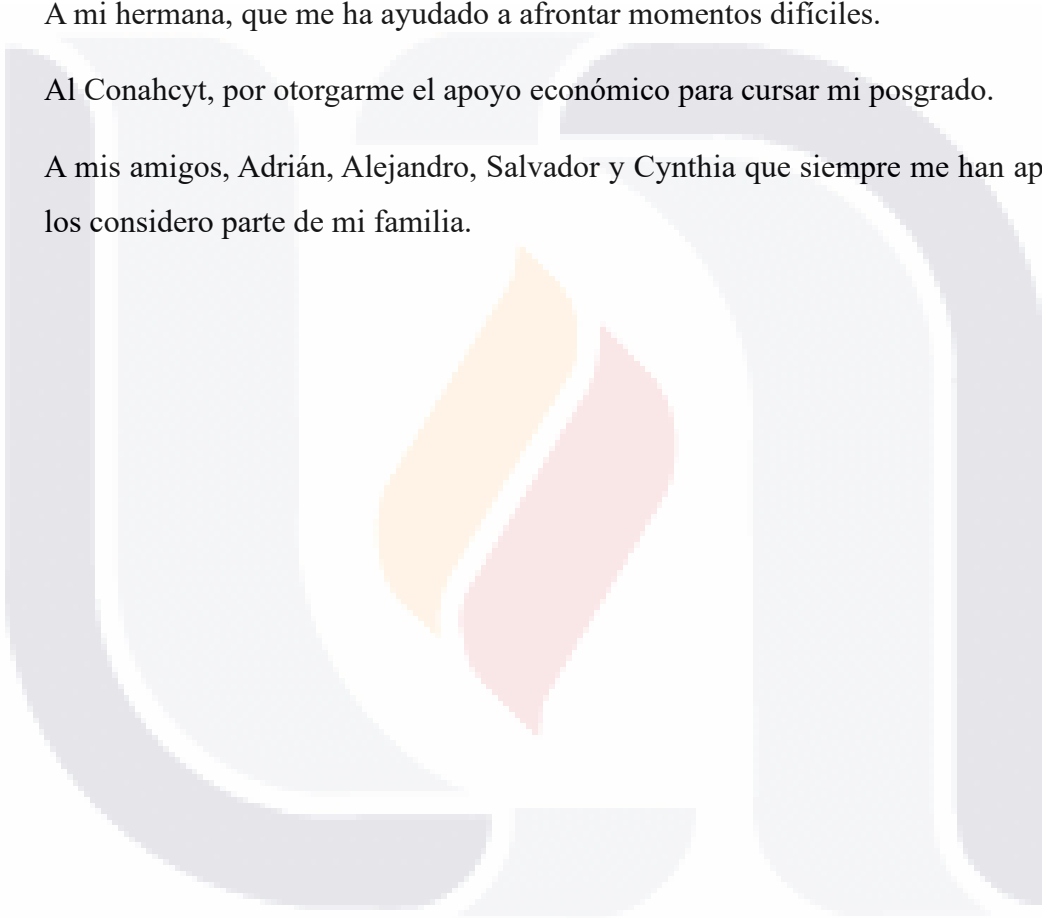
A mi Comité Tutorial por el asesoramiento, apoyo, paciencia y guía que me proporcionaron a lo largo del desarrollo del trabajo práctico.

A mi madre, que siempre ha estado a mi lado alentándome para lograr todo lo que me proponga.

A mi hermana, que me ha ayudado a afrontar momentos difíciles.

Al Conahcyt, por otorgarme el apoyo económico para cursar mi posgrado.

A mis amigos, Adrián, Alejandro, Salvador y Cynthia que siempre me han apoyado y los considero parte de mi familia.



Dedicatorias

Este trabajo está dedicado a mi madre, Alma Verónica Rodríguez Valdez que siempre ha sido mi apoyo incondicional y mi motor para seguir adelante.

A mi hermana, que juntos hemos pasado por situaciones muy difíciles pero siempre hemos estado el uno para el otro.



Índice General

1. Introducción del Caso Problema del Trabajo Práctico	9
1.1 Presentación del Trabajo Práctico	9
1.2 Objetivo General del Trabajo Práctico	11
1.3 Metodología Específica para Desarrollar Solución al Caso Problema.	12
1.4 Organización del Documento	14
2. Planteamiento del Caso Problema del Trabajo Práctico.	15
2.1 Descripción del Problema Específico del Caso.	15
2.2 Justificación de Intervención.	15
2.3 Objetivos Específicos de la Intervención en el Caso Problema.	16
2.4 Organización, Cliente, Usuarios Principales del Trabajo Práctico.	16
3. Fundamentación Teórica.	17
3.1 Revisión de Teorías Bases	17
3.2 Revisión de Intervenciones Similares.	37
3.3 Contribuciones y Limitaciones.....	43
4. Diseño de Intervención al Caso Problema.	47
4.1 Etapa 1. Elaboración del Backlog de MVITS: involucramiento y planificación	48
4.2 Etapa 2. Ciclos MVITS. Diseño, transición y construcción	61
4.3 Etapa 3. Versión final. Entrega y soporte.	74
5. Resultados de la Intervención.	78
6. Evaluación de la Intervención.	79
6.1 Datos Demográficos.....	79
6.2 Estadística Descriptiva	81
7. Conclusiones.....	85
7.1 Conclusiones del objetivo específico 1	85
7.2 Conclusiones del objetivo específico 2	85
7.3 Conclusiones del objetivo específico 3	86
7.4 Conclusiones del objetivo específico 4	86
7.5 Beneficios obtenidos	87
7.6 Problemas adicionales encontrados.....	87
7.7 Recomendaciones finales para el cliente y usuarios.....	88
7.8 Conclusiones finales.....	89
8. Glosario	90
9. Bibliografía.....	92
10. Anexos	96
Anexo 1	

Encuesta demográfica	97
Anexo 2	
Instrumento conceptual de métricas de usabilidad de tecnologías de información.	101
11. Apéndices.....	102
Apéndice 1	
Evidencia de funcionamiento del Rstudio Server instalado en el MVITS.	102
Apéndice 2	
Evidencia de funcionamiento del paquete R para CentOS 7.....	103
Apéndice 3	
Evidencia de instalación del sistema operativo CentOS 7 en la máquina virtual.....	104
Apéndice 4	
Evidencia del funcionamiento del workbench de MySQL instalado en el MVITS.....	105
Apéndice 5	
Evidencia del curso de capacitación destinado para usuarios finales.	106

Índice de Tablas

Tabla 1: Se muestra la clasificación de una empresa por número de empleados acorde al sector al que pertenecen.	10
Tabla 2: Clasificación de MPYMES acorde a sus ingresos anuales.	10
Tabla 3: Fases, Actividades y Calendarización.	12
Tabla 4: Tabla de contribuciones y limitaciones de teoría base.	44
Tabla 5: Tabla comparativa de trabajos similares.	46
Tabla 6: Descripción de las actividades involucradas en la tarea T.1.1 respecto a la metodología AVS_FD_MVITS.v1-2022.	49
Tabla 7: Hoja de backlog (inicial), tomada de la metodología AVS_FD_MVITS.v1-2022.	52
Tabla 8: Descripción de prácticas, tiempos de entrega y actividades a realizar en la tarea T.1.2 acorde a la metodología AVS_FD_MVITS.v1-2022.	53
Tabla 9: Matriz de asignación funcional, se toma como referencia las historias de usuario y los componentes TI de cada capa del diagrama de bloques (tomada de la metodología AVS_FD_MVITS.v1-2022).	55
Tabla 10: Matriz de asignación no-funcional acorde a las métricas de evaluación de servicio y componentes TI (tomada de la metodología AVS_FD_MVITS.v1-2022).	56
Tabla 11: Evidencias sobre la viabilidad del MVITS, es posible el desarrollo del proyecto al confirmar la historia de usuario 10 (tomada de la metodología AVS_FD_MVITS.v1-2022).	57
Tabla 12: Descripción de prácticas, tiempo de entrega y actividades involucradas en la tarea T.1.3 acorde a la metodología AVS_FD_MVITS.v1-2022.	58
Tabla 13: Proceso del cálculo de la velocidad de trabajo, los datos en color verde representan los datos acordados por el cliente y el equipo de desarrollo, los datos en azul se van calculando por medio de las fórmulas descritas en la metodología AVS_FD_MVITS.v1-2022.	59
Tabla 14: Hoja de backlog (reordenada) añadiendo los puntos de historia respectivos a cada historia de usuario, además de señalar las historias de usuario a realizar en cada ciclo.	60
Tabla 15: Descripción de prácticas ágiles, tiempos de entrega y actividades involucradas en la tarea T.2.1 acorde a la metodología AVS_FD_MVITS.v1-2022.	63
Tabla 16: Hoja de trabajo del ciclo #1.	64
Tabla 17: Hoja de trabajo del ciclo #2.	65
Tabla 18: Descripción de prácticas ágiles, tiempos de entrega y actividades involucradas en la tarea T.2.2 acorde a la metodología AVS_FD_MVITS.v1-2022.	66
Tabla 19: Ejemplo de una hoja de reuniones rápidas acorde a la metodología AVS_FD_MVITS.v1-2022.	66

Tabla 20: Descripción de prácticas ágiles, tiempos de entrega y actividades involucradas en la tarea T.2.3 acorde a la metodología AVS_FD_MVITS.v1-2022.	67
Tabla 21: Hoja de pruebas de aceptación y pruebas técnicas de cada historia de usuario que conforma el ciclo #1.....	68
Tabla 22: Hoja de pruebas de aceptación y pruebas técnicas de cada historia de usuario que conforma el ciclo #1.....	69
Tabla 23: Descripción de prácticas ágiles, tiempos de entrega y actividades involucradas en la tarea T.2.4 acorde a la metodología AVS_FD_MVITS.v1-2022.	70
Tabla 24: Descripción de prácticas ágiles, tiempos de entrega y actividades involucradas en la tarea T.2.5 acorde a la metodología AVS_FD_MVITS.v1-2022.	71
Tabla 25: Hoja de registro sobre la revisión del ciclo #1 acorde a la metodología AVS_FD_MVITS.v1-2022.	72
Tabla 26: Hoja de registro sobre la revisión del ciclo #1 acorde a la metodología AVS_FD_MVITS.v1-2022.	73
Tabla 27: Descripción de prácticas ágiles, tiempos de entrega y actividades involucradas en la tarea T.3.1 acorde a la metodología AVS_FD_MVITS.v1-2022.	75
Tabla 28: Documento de acuerdo de soporte para la versión final acorde a la metodología AVS_FD_MVITS.v1-2022.	76
Tabla 29: Descripción de prácticas ágiles, tiempos de entrega y actividades involucradas en la tarea T.3.2 acorde a la metodología AVS_FD_MVITS.v1-2022.	77
Tabla 30: Documento de aceptación de la versión final del MVITS acorde a la metodología AVS_FD_MVITS.v1-2022.	77
Tabla 31: Distribución de resultados de la encuesta demográfica.	79
Tabla 32: Resultados del análisis descriptivo del constructo utilidad.	82
Tabla 33: Resultados del análisis descriptivo del constructo facilidad de uso.	83
Tabla 34: Resultados del análisis descriptivo del constructo compatibilidad.	83
Tabla 35: Resultados del análisis descriptivo del constructo valor.	83

Índice de Figuras

Figura 1: Ciclo de vida del servicio ITILv3.	17
Figura 2: Actividades importantes a desempeñar dentro de la etapa de diseño de un servicio de TI acorde a ITILv3.	19
Figura 3: Mapa de las actividades que conforman la gestión de la cadena de valor.	20
Figura 4: Incrementos de conceptos populares tales como: ciencia de datos, analítica de datos, minería de datos, entre otros. Se muestra su comportamiento al pasar los meses.	26
Figura 5: Los tres principales tipos de analítica y sus puntos principales de cada una de estas.	29
Figura 6: Las cinco etapas genéricas dentro del proceso de Ciencia de Datos.	31
Figura 7: Distribución de aplicación de las herramientas Open-Source seleccionadas acorde a las fases genéricas de Ciencia de Datos.	36
Figura 8: Arquitectura de la plataforma Checkmetrix (tomada de “Data Analytics Platforms: Value Propositions and Adoption Challenges for Small Hospitality Businesses”).	39
Figura 9: Arquitectura de SciServer Compute y su interacción con diversos componentes (tomada de “SciServer Compute: Bringing Analysis Close to the Data”).	41
Figura 10: Componentes esenciales de la plataforma RADStack (tomada de “The RADStack: Open-Source Lambda Architecture for Interactive Analytics”).	43
Figura 11: Diagrama de la metodología AVS_FD_MVITS desarrollada por los investigadores --- Manual Mora, Olayele Adelakun, Yuritzzy Reyes y Oswaldo Díaz.	47
Figura 12: Etapa inicial de Involucramiento y Planificación del proyecto (tomada de “Metodología Ágil para el Desarrollo de un Nuevo Servicio de TI Mínimo Viable --- AVS_FD_MVITS.v1-2022”).	48
Figura 13: Diagrama de contexto de las demandas del MVITS acorde a la metodología AVS_FD_MVITS.v1-2022.	50
Figura 14: Diagrama de casos de uso, tomada de la metodología AVS_FD_MVITS.v1-2022.	51
Figura 15: Diagrama de bloques de los componentes TI implementados en el MVITS (entorno estudiantil) acorde a la metodología AVS_FD_MVITS.v1-2022.	54
Figura 16: Etapa de Diseño, Transición y Construcción (tomada de “Metodología Ágil para el Desarrollo de un Nuevo Servicio de TI Mínimo Viable (AVS_FD_MVITS.v1-2022)”).	62
Figura 17: Etapa Entrega y Soporte (tomada de “Metodología Ágil para el Desarrollo de un Nuevo Servicio de TI Mínimo Viable (AVS_FD_MVITS.v1-2022)”).	74
Figura 18: Base de datos acorde a la información recopilada con ayuda artefacto descrito en el anexo 2.	81
Figura 19: Características definidas para cada variable de la base de datos.	82

Resumen en español

En México, en los últimos años, hemos podido observar y presenciar el desarrollo y crecimiento de este tipo de empresas PYMES. Este crecimiento es mérito principalmente a los apoyos creados y otorgados por el Gobierno, además de la creación e implementación de nuevos reglamentos para dichas empresas, todo esto con el fin de impulsar y acelerar el crecimiento económico y por consecuencia reducir los índices de pobreza del país. Actualmente en México, la cantidad de empresas clasificadas como PYMES oscilan entre 4.9 millones, que en otras palabras, representan al 99.8% de establecimientos del sector privado y paraestatal (conforme con los datos recopilados por el INEGI). Además en los últimos años, el desarrollo y creación de servicios de Tecnologías de Información (TI) se ha vuelto de suma importancia en cualquier sector empresarial y en casi todas las empresas (sin importar tamaño o giro empresarial), ya que éstas, pueden llegar a ser imprescindibles en la toma de decisiones.

El presente trabajo práctico muestra el desarrollo y evaluación de una plataforma de código abierto para desempeñar tareas propias de la ciencia de datos para una empresa local.

Se implementó una metodología ágil para la construcción del nuevo servicio de TI y para poder evaluarlo se tomaron artefactos demográficos y métricas de evaluación los cuales determinaran el grado de adopción.

Resumen en inglés

Mexico, in recent years, we have been able to observe and witness the development and growth of this type of SME companies. This growth is mainly due to the support created and granted by the Government, in addition to the creation and implementation of new regulations for such companies, all of this with the purpose of boosting and accelerating economic growth and consequently reducing the country's poverty indexes. Currently in Mexico, the number of companies classified as SMEs ranges between 4.9 million, which in other words, represent 99.8% of private and parastatal sector establishments (according to data compiled by INEGI). In addition, in recent years, the development and creation of Information Technology (IT) services has become extremely important in any business sector and in all companies (regardless of size or business line), since they can become essential in decision making.

This practical work shows the development and evaluation of an open-source platform to perform data science tasks for a local company.

An agile methodology was implemented for the construction of the new IT service and to evaluate it, demographic artifacts and evaluation metrics were used to determine the degree of adoption.

1. Introducción del Caso Problema del Trabajo Práctico.

1.1 Presentación del Trabajo Práctico.

Es evidente que en los últimos 40 años el uso de las Tecnologías de Información (TI) se han vuelto de suma importancia para cualquier sector empresarial sin importar el tamaño de la empresa, ya que es imprescindible la toma de decisiones y es aquí donde entran en juego los servicios y/o sistemas que ofrecen las TI (Meléndez et al., 2016).

Gracias a la TI podemos comunicar grandes cantidades de información a personas y/u organizaciones de todo el mundo (en un abrir y cerrar de ojos). Desde las computadoras, calculadoras y equipos médicos, hasta satélites o Wall Street, son herramientas poderosas e invaluable que nos ayudan a hacer nuestro trabajo y permiten que estemos conectados entre todos. En pocas palabras, la TI se puede definir como la tecnología digital, como computadoras y el Internet, para almacenar y procesar datos en información útil. La TI no solo tiene que ver el lado técnico, sino que también involucra la interacción entre las personas y tales tecnologías computacionales. Como indican (Olmera & Rosas, 2011): ¿de qué sirve la Tecnología de Información, si las personas no pueden usarla o darle sentido? Para tal fin, las TI en las organizaciones deben estar disponibles en los horarios esperados, y tener la capacidad, continuidad, seguridad y rendimiento esperados por los acuerdos de Servicios de TI.

A nivel mundial, no solo las grandes y medianas organizaciones son usuarias de las TI. También las pequeñas y micro organizaciones (PYMES) en las últimas décadas se han vuelto usuarios activos. PYMES son definidas por la Secretaría de Economía de México como empresas que cuentan con cierta cantidad de empleados por sector al que pertenecen (industria, comercio o servicios), en la siguiente tabla se muestra la cantidad de empleados específica por sector para que una empresa se considere micro, pequeña, mediana o grande (Alcaraz Rodríguez, 2011).

Sector	Tamaño de la empresa	Número de empleados
Industria	Micro	1-30
	Pequeña	31-100
	Mediana	101-500
	Grande	Más de 500
Comercio	Micro	1-5
	Pequeña	6-20
	Mediana	21-100
	Grande	Más de 100
Servicios	Micro	1-20
	Pequeña	21-50
	Mediana	51-100
	Grande	Más de 100

Tabla 1: Se muestra la clasificación de una empresa por número de empleados acorde al sector al que pertenecen.

Estratificación		
Tamaño	Sector	Rango de ventas (ingresos) anuales (mdp)
Micro	Comercio	Hasta \$4
	Servicio	
	Industria	
Pequeña	Comercio	Desde \$4.01 hasta \$100
	Servicios	
	Industria	
Mediana	Comercio	Desde \$100.01 hasta \$250
	Servicios	
	Industria	

Tabla 2: Clasificación de MPYMES acorde a sus ingresos anuales.

En México, en los últimos años hemos podido observar y presenciar el desarrollo y crecimiento de este tipo de empresas PYMES. Este crecimiento es mérito principalmente a los apoyos creados y otorgados por el Gobierno, además de la creación e implementación de nuevos reglamentos para dichas empresas, todo esto con

el fin de impulsar y acelerar el crecimiento económico y por consecuencia reducir los índices de pobreza del país (Heredia, 2014).

Desafortunadamente, a pesar de que las PYMES necesitan las TI y pueden obtener múltiples beneficios organizacionales, hemos detectado que los estudios y/o investigaciones acerca de la integración de los Servicios de TI (Soporte Técnico, Redes Computacionales, Sistemas Operativos, Sistemas Administrativos y de Infraestructura y Seguridad) en PYMES son escasos. En contraste, podemos darnos cuenta de que tales estudios abundan en cuanto a grandes y medianas empresas se refiere, pero si se trata de las pequeñas y micros empresas es todo lo contrario y esto genera desaliento en utilizarlas y varias desventajas tales como: tener recursos limitados, capacidades tecnológicas restringidas y con planes de poca adaptabilidad. Es por lo anterior que las PYMES aún tienen problemas para implementar TI, ya que el proceso deber ser monitoreado para ver que vaya cumpliendo con los objetivos planteados, además de que obviamente se debe invertir en dicho proceso (Aguirre, 2015).

Adicionalmente, a nivel mundial emergen relevantes innovaciones en las TI, tales como el Internet de las Cosas (Rose et al., 2015), la Industria 4.0 (Boone, 2022), la Digitalización de los Negocios (Bouwman et al., 2018), y la Ciencia de Datos, la cual incluye la Analítica de Negocios. Sin embargo, las PYMES, por las razones anteriormente expuestas, usualmente quedan fuera de poder utilizarlas y por lo tanto obtener sus beneficios potenciales. En particular, la Ciencia de Datos es una de las TI más avanzadas actualmente y esta diciplina surge gracias a la necesidad de procesar, analizar y visualizar los datos internamente generados por la organización tanto como los datos externos disponibles, para transformarlos en información útil de fácil acceso y entendimiento para la mejora en la toma de decisiones organizacionales.

1.2 Objetivo General del Trabajo Práctico.

Actualmente, la utilización de plataformas de Ciencia de Datos es ampliamente usadas por grandes y medianas empresas, ya sea instaladas o en nubes de proveedores externas. Sin embargo, las PYMES usualmente no tienen acceso a TI avanzadas como plataformas de Ciencias de Datos debido a los altos costos de implementarlas (recursos

de hardware, software, y personal experto). Afortunadamente, existen plataformas de Ciencia de Datos de tipo Open-Source (no costos de licencias para instalarlas y usarlas), pero su selección correcta y su proceso de instalación y uso demandan conocimiento experto. Este Trabajo Práctico busca atender la problemática de **"no utilización de Plataformas de Ciencia de Datos en PYMES por altos costos de implantación"** mediante el objetivo general de:

- **"Presentar un caso prueba de concepto prototipo del diseño ágil, implantación y evaluación de un servicio de TI web de Plataforma de Ciencia de Datos (tipo Open-Source), que sea evaluado adecuadamente en las métricas de utilidad, facilidad de uso, valor e intención de uso".**

1.3 Metodología Específica para Desarrollar Solución al Caso Problema.

1.3.1 Actividades principales

Fase	Actividad	Calendarización			
		2022-2	2023-1	2023-2	2024-1
Planeación	Inducción	X			
	Diagnóstico		X		
	Diseño General del servicio solicitado		X		
Implementación	Diseño Detallado del servicio solicitado			X	
	Construcción del servicio solicitado			X	
	Instalación del servicio solicitado			X	
Evaluación	Beta				X
	Final				X
	Reporte Final				X

Tabla 3: Fases, Actividades y Calendarización.

1.3.2 Materiales y Equipo

1.3.2.1 Documentación (metodología a elegir)

- Metodología: AVS_FD_MVTIS.v1-2022 (open-source).

1.3.2.2 Equipos y Software

- Equipo de cómputo (descripción de mi computadora actual):
 - Procesador: Intel(R) Core(TM) i5-11400H.
 - RAM: 8.00GB.
 - Disco duro: 500 GB.
 - Sistema Operativo: Windows 11 64 bits.
- Posibles softwares por usar (AVS_FD_MVTIS.v1-2022):
 - 1 servidor virtual CENTOS 7.0 de Plataforma Ciencia de Datos (incluye Base de Datos MySQL, RStudio Server Community Edition, RShiny Server Community Edition, R packages, Jupyter Lab, y Apache SuperSet).

1.3.3 Métodos de Evaluación

La evaluación de la plataforma open-source desarrollada para la empresa “Nombre de la empresa” (la selección de la empresa aún está por ser definida o acordada con mi comité tutorial) será calificada por mi Comité Tutorial, además de la Junta Directiva de dicha empresa mediante un instrumento de medición de percepción usado comúnmente en la literatura científica de Sistemas de Información. Dicho instrumento evaluará las siguientes métricas:

- Utilidad.
- Facilidad de uso.
- Compatibilidad
- Valor

1.3.4 Limitaciones

Ya que cada empresa **tipo PYME** cuenta con su propia infraestructura y esto ocasiona que los requerimientos cambien respecto a cada una de ellas, se decidió enfocarse directamente sobre la empresa indicada dentro del estado de Aguascalientes, México.

1.4 Organización del Documento.

Este documento está organizado en siete capítulos:

1. Introducción del caso problema del trabajo practico: se describe la presentación del trabajo, así como el objetivo general y la metodología específica para el desarrollo.
2. Planteamiento del caso problema del trabajo práctico: aquí se describe el caso problema al que se busca dar solución, su justificación, los objetivos específicos al igual que la organización, clientes y usuarios a quienes está enfocado el trabajo práctico.
3. Fundamentación teórica.
4. Diseño de intervención al caso problema.
5. Resultados y validación de intervención: se describen y explican los resultados obtenidos de las evaluaciones por parte del comité tutorial, usuario principal y grupo piloto de usuarios al haber implementado la plataforma open-source.
6. Evaluación de la intervención: se describen las evaluaciones realizadas y la relación de los mismos respecto a los objetivos específicos.
7. Conclusiones: se argumentan las conclusiones sobre los resultados obtenidos sobre la investigación, el proceso realizado y las conclusiones. Adicional, se describen los beneficios obtenidos, así como los problemas que surgieron durante el proceso y las recomendaciones finales.

2. Planteamiento del Caso Problema del Trabajo Práctico.

2.1 Descripción del Problema Específico del Caso.

La empresa de computación “Nombre de la empresa” líder en su sector de producción entra dentro del espectro de microempresa ya que cuenta únicamente con 3-5 empleados, manejan datos de autoría privada por lo que se deben de manejar son cierto cuidado (aquí se hace énfasis en la posible capa de seguridad que debe requerir la plataforma de ciencia de datos terminada).

Esta Organización se ubica únicamente en el Estado de Aguascalientes de la República Mexicana. Cuentan con un departamento de informática el cual está al tanto de la importancia de una plataforma de Ciencia de Datos y su valor que pudiera traer directamente a su departamento de peticiones de servicios.

2.2 Justificación de Intervención.

El departamento de Tecnologías Computaciones de la empresa son los encargados de administrar toda la información pertinente a la petición de servicios y encuentran útil analizar estos datos por medio de herramientas de Ciencias de Datos para poder extraer características de interés para el negocio.

Al toparse con el impedimento de los recursos financieros optan por contratar algún servicio de código abierto que pueda satisfacer sus necesidades actuales, pero no encuentran alguno que se adapte a sus requerimientos o hardware actual.

Es aquí donde entra la importancia de crear una plataforma de código abierto capas de ser adaptada a las necesidades y equipo básico de las empresas micro o pequeñas, ya que según sea el caso se pueden cargar ciertos paquetes en un servidor virtual que sea operable por un usuario final y pueda obtener los resultados que espera con las herramientas proporcionadas.

2.3 Objetivos Específicos de la Intervención en el Caso Problema.

El desarrollo de este Trabajo Práctico tiene planteado los siguientes Objetivos Específicos a lograr:

1. Elaborar una revisión comparativa sumariada entre los procesos de diseño de Servicios de TI principales de Rigor (ITIL v3.v2011) y de Agilidad (ITIL v4, AVS_FD_MVITS).
2. Revisar las funcionalidades y requerimientos de utilización de las principales herramientas de tipo Open Source para las diversas fases de la Ciencia de Datos, y seleccionar las herramientas a utilizar en este Trabajo Práctico.
3. Implementar un Servicio de TI (prototipo) de Plataforma Open-Source Web de Ciencia de Datos para PYMES, usando la metodología ágil AVS_FD_MVITS.
4. Obtener una evaluación piloto del Servicio de TI (prototipo) implementado respecto a la utilidad, facilidad de uso, compatibilidad y valor, por parte de un grupo de usuarios de la Organización Cliente contactada.

2.4 Organización, Cliente, Usuarios Principales del Trabajo Práctico.

La organización principal para la cual se realiza este trabajo practico es “Nombre de la empresa” la cual entra dentro del espectro de microempresa y su principal modelo de negocios es brindar servicios de computación. Como cliente directo tenemos a su departamento de tecnologías computaciones los cuales dan soporte y serán los usuarios finales de la plataforma.

3. Fundamentación Teórica.

3.1 Revisión de Teorías Bases.

3.1.1 Fundamentos de Diseño de Servicios de TI basados en Marcos Internacionales de Gestión de Servicios de TI.

3.1.1.1 Estructura (Roles, Actividades, y Artefactos) de los Procesos de Diseño de Servicios de TI de Rigor (ITIL v3-2011).

ITILv3-2011

Debemos recordar que la Librería de Infraestructura de TI (“ITIL” por sus siglas en inglés) nos otorga una “guía” de las mejores prácticas destinadas para la Gestión de servicios de TI (“ITSM” por sus siglas en inglés) (Cartlidge et al., n.d.) En este apartado se analizarán los elementos que conforman el marco de rigor de la versión de ITILv3 (2011) en cuanto a roles, actividades y artefactos se refiere, obviamente enfocándonos en la etapa de Diseño del Servicio. Cabe recalcar que nos enfocamos en la sección del Diseño del Servicio del ciclo de vida del servicio (Figura 1) descrito en ITIL, ya que el objetivo final de este trabajo es lograr el Diseño de un entorno o servicio funcional con las mínimas especificaciones requeridas para la aplicación de Ciencia de Datos.

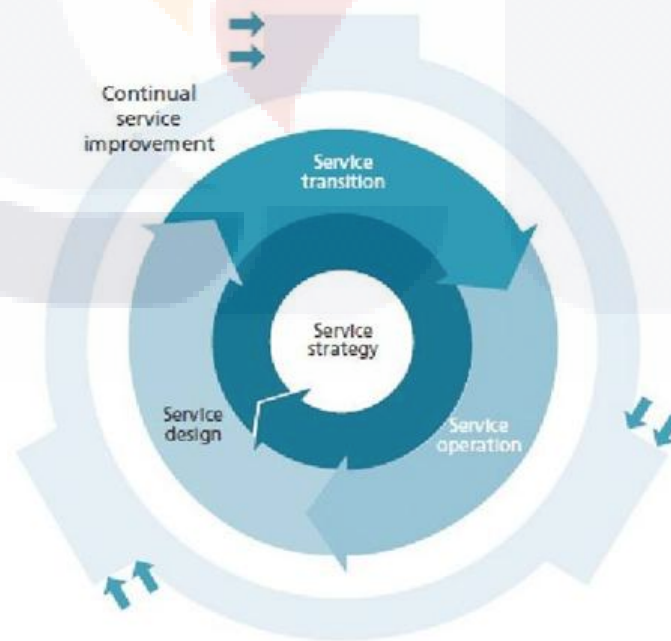


Figura 1: Ciclo de vida del servicio ITILv3.

Roles

ITILv3 (2011) nos indica y/o sugiere roles genéricos y específicos, además de proporcionarnos ejemplos según el tamaño de la organización, no obstante, podemos englobarlos en los siguientes roles:

- 1) Gestor de diseño del servicio.
- 2) Diseñador/Arquitecto de TI.
- 3) Gestores del proceso de diseño del servicio.
- 4) Cliente.
- 5) Usuario.

Actividades y/o tareas

ITILv3 (2011) no ofrece un proceso de diseño de servicio de TI específico o “base” para todas las organizaciones, sin embargo, nos sugiere ciertas actividades para lograr el éxito de esta etapa de diseño. Dichas actividades son las siguientes:

- 1) **Identificar los requerimientos del servicio:** al momento de diseñar un nuevo servicio se debe considerar cada uno de los elementos como un “todo” (el servicio, sus componentes y sus interrelaciones)(Reyes-Delgado et al., 2023).
- 2) **Identificar y documentar los requerimientos del negocio y sus “drivers”:** para ofrecer un catálogo de servicios con un nivel aceptable de calidad, es imprescindible la información sobre las necesidades empresariales (Reyes-Delgado et al., 2023).
- 3) **Diseñar soluciones de servicio:** las tareas y/o actividades destinadas para la creación o modificación de servicios deben desarrollarse pensando en los posibles costos, su utilidad, su garantía y los tiempos acordados (Reyes-Delgado et al., 2023).
- 4) **Evaluar soluciones alternativas:** en caso de contar con servicios y soluciones aportadas por proveedores externos puede ser necesaria esta actividad para su evaluación (Reyes-Delgado et al., 2023).
- 5) **Adquirir la mejor solución:** frecuentemente es inevitable usar elementos externos proporcionados por algún proveedor en las soluciones a implementar,

es por esto por lo que antes se debe determinar los términos y condiciones de cualquier nueva contratación y adquisición de la solución seleccionada (Reyes-Delgado et al., 2023).

- 6) **Desarrollar la solución de servicio:** por último, se debe construir un plan de desarrollo y/o reutilización de componentes necesarios para proveer el servicio y su implementación (Reyes-Delgado et al., 2023).

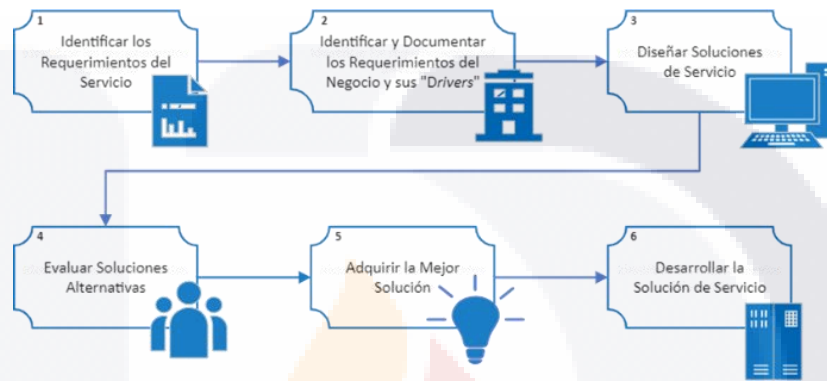


Figura 2: Actividades importantes a desempeñar dentro de la etapa de diseño de un servicio de TI acorde a ITILv3.

Artefactos

Denominado como “paquete de diseño del servicio” es el producto o resultado esperado una vez completada la etapa de Diseño. En este paquete se deben detallar todos los aspectos del servicio y sus requerimientos teniendo en cuenta cada etapa del ciclo de vida del servicio (Hunnebeck et al., 2011):

- **Requerimientos**
 - Requerimientos del negocio.
 - Aplicabilidad del servicio.
 - Contactos del servicio.
- **Diseño del servicio**
 - Requisitos funcionales del servicio.
 - Requisitos de nivel del servicio.
 - Requisitos de servicio y gestión operacional.
 - Topología y diseño del servicio.
- **Evaluación del grado de preparación de la organización**

- **Plan del ciclo de vida del servicio**

- Programa de servicio.
- Plan de transición del servicio.
- Continuidad del plan de transición del servicio.
- Plan de aceptación operativa del servicio.
- Criterios de aceptación del servicio.

3.1.1.2 Estructura (Roles, Actividades, y Artefactos) de los Prácticas de Diseño de Servicios de TI con Agilidad (ITIL v4, AVS_FD_MVITS v.2022).

ITILv4-2019

ITIL 4 proporciona la orientación que las organizaciones necesitan para afrontar los nuevos retos de la gestión de servicios y aprovechar el potencial de la tecnología moderna. Está diseñada para garantizar un sistema flexible, coordinado e integrado para el gobierno y la gestión eficaces de los servicios posibilitados por TI.

El agregado en esta versión de ITIL es el concepto de “cadena de valor” (Figura 3), sin embargo, nos seguimos enfocando en la etapa de diseño ya que es el objetivo principal de la creación de un servicio TI de ciencia de datos para PYMES. En este nuevo modelo se agrupa tanto el diseño como la transición.

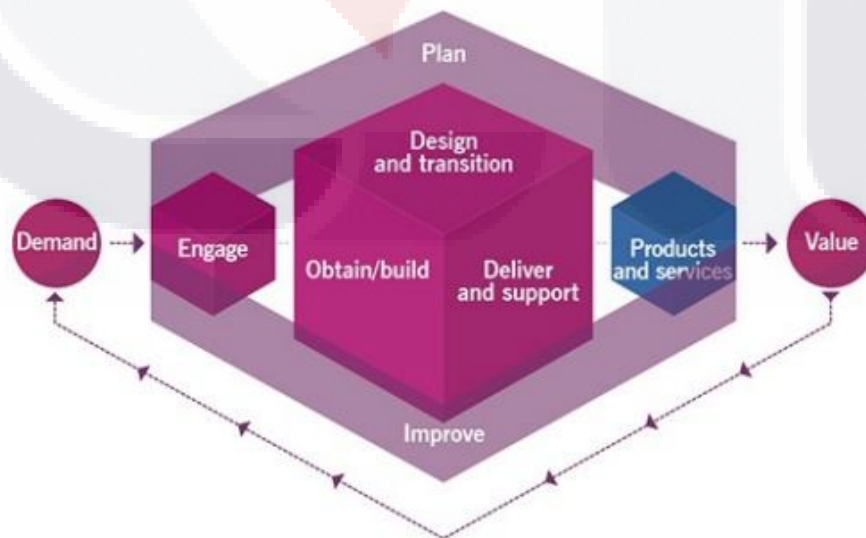


Figura 3: Mapa de las actividades que conforman la gestión de la cadena de valor.

Roles

Conjunto de responsabilidades, actividades y autoridades asignadas a una persona o equipo. Un rol se define en un proceso o función. Una persona o un equipo pueden tener varias funciones:

- 1) Coach de diseño del servicio ágil.
- 2) Equipo central de diseño del servicio de TI ágil.
- 3) Equipo de soporte de diseño del servicio de TI ágil.
- 4) Cliente
- 5) Usuario

Actividades y/o prácticas

1. **Involucrarse:** proporcionar una buena comprensión de las necesidades de las partes interesadas, transparencia, y un compromiso continuo y buenas relaciones con todas las partes interesadas (Stationery Office, 2012).
2. **Plan:** el propósito de la actividad de la cadena de valor del plan es garantizar una comprensión compartida de la visión, el estado actual y la dirección de mejora de las cuatro dimensiones y de todos los productos y servicios de la organización (Stationery Office, 2012).
3. **Diseño y transición:** diseñar productos y servicios que se ajusten a su finalidad, se adapten a su uso y puedan ser suministrados por la organización y su ecosistema. Esto incluye la planificación y organización de las personas, socios y proveedores, información, comunicación, tecnología y prácticas para productos y servicios nuevos o modificados, y la interacción entre la organización y sus clientes (Stationery Office, 2012).
4. **Obtener/Construir:** garantizar que los componentes del servicio estén disponibles cuando y donde se necesiten, y que cumplan las especificaciones acordadas (Stationery Office, 2012).
5. **Diseño y transición:** garantizar que los productos y servicios satisfagan continuamente las expectativas de las partes interesadas en cuanto a calidad costes y plazos de comercialización (Stationery Office, 2012).

6. **Entregar y dar soporte:** es garantizar que los servicios se prestan y apoyan de acuerdo con las especificaciones expectativas de las partes interesadas.

Artefactos

- 1) Modelo “MVS” (servicio mínimo viable) para servicio de TI ágil.
- 2) Decisión financiera del servicio de TI ágil.
- 3) Modelo de diseño de compromisos del servicio de TI ágil.
- 4) Modelo de diseño del servicio de TI ágil.
- 5) Realización del diseño del servicio de TI ágil.

AVS_FD_MVITS v.2022

Marco de metodología propuesto para casos en los que se requieran optar por una versión mínima de un servicio de TI, ya sea por falta de recursos o de infraestructura.

Roles

- 1) **Cliente:** este rol corresponde al propietario del producto, o el cliente propietario del servicio ITILv4. Además, este rol establece el alcance y el valor del “MVITS”. También es quien establece el nivel de aceptación de las pruebas. De igual forma es quien representa al usuario final (Mora, Adalakun, et al., n.d, 2019).
- 2) **Experto:** para poder desempeñar dicho rol es necesario dominar y dirigir el flujo de trabajo ágil para el desarrollo del MVITS. Gestiona el repositorio del MVITS. Por último, es el encargado de presentar y cerrar el proyecto AVS_FD_MVITS (Mora et al., 2019).
- 3) **Equipo de desarrollo:** son los encargados de establecer el flujo de trabajo para cada ciclo del proyecto (establecen el máximo de tareas a completar). También, se ocupan de elaborar un diseño simple de la arquitectura del MVITS. Por último, el equipo de desarrollo son los que diseñan pruebas técnicas y aplicarlas para probar el diseño del servicio (Mora et al., 2019).
- 4) **Consultor externo:** colabora en el diseño simple del servicio y del incremento en la ingeniería (si es necesario) (Mora et al., 2019).

Actividades y/o tareas

1. Fase 1: Comprometerse y planificar

- a. **Formulación del backlog:** elaborar el diagrama de contexto de demandas, además de describir de la problemática además de una lista de necesidades (Mora et al., 2019).
- b. **Elaboración de la arquitectura:** el equipo de desarrollo elabora iterativamente un diseño simple (un boceto de la arquitectura). Se debe incluir: un diagrama de componentes informáticos, software del sistema y base de datos, hardware computacional, red y la infraestructura del servicio de TI (Mora et al., 2019).
- c. **Negociación del backlog:** el equipo de desarrollo estima un puntaje para las “historias” (acciones que deben ser cubiertas por el servicio de TI) para ayudar a la planificación de estas, este puntaje va entre los valores de 1, 2, 3 o 5. Después se reordenarán las historias de ser necesario en base a los puntajes anteriormente otorgados (Mora et al., 2019).

2. Fase 2: Diseño, Transición y Construir-Obtener

- a. **Planificación del sprint backlog:** el cliente define los objetivos de cada “vuelta” o revisión del servicio de TI, indica las funcionalidades que deben implementarse en cada ciclo, además de establecer el nivel de las métricas del servicio que se deben lograr (Mora et al., 2019).
- b. **Reuniones rápidas de 20-30 minutos:** el encargado del proyecto (experto) organiza reuniones rápidas con el equipo de desarrollo para poder informar al cliente aspectos tales como: tareas que se llevaron a cabo un día anterior, tareas y/o actividades que se realizarán el día actual y sus posibles impedimentos. Se llevará un registro de estas reuniones para poder evaluar avances (Mora et al., 2019).
- c. **Aceptación del cliente y elaborar pruebas técnicas:** el cliente debe establecer las condiciones de aceptación para cada “historia” para cada ciclo de evaluación del sistema, el cliente usa una plantilla

proporcionada por el experto. El equipo de desarrollo elabora pruebas de aceptación proporcionadas por el cliente, también se usan platillas específicas para la tarea (Mora et al., 2019).

- d. **Incrementar la ingeniería:** el equipo de desarrollo aplica las pruebas técnicas y de aceptación (de acuerdo con lo propuesto por el cliente) para el nuevo incremento en el modelo MVITS. Esta tarea se centra en evitar cualquier posible fallo que puede surgir en la revisión de cada ciclo. En caso de surgir una falla en alguna de las “historias” propuestas (no puede ser implementada como se acordó), el equipo de desarrollo debe comunicarlo al cliente para poder llevar una negociación (Mora et al., 2019).
 - e. **Revisión del sprint:** el experto debe planificar reuniones para la revisión de cada ciclo del MVITS. El cliente usará el nuevo incremento al ciclo del MVITS para poder aceptar o rechazar cada historia propuesta en la fase 1. Además, el cliente y el equipo de desarrollo revisan en conjunto los objetivos de cada ciclo y lo que se haya logrado o hecho (Mora et al., 2019).
 - f. **Retrospectiva del sprint:** el equipo de desarrollo planifica reuniones internas para intercambiar ideas y/o emociones que se deban tratar para modificar o lograr el siguiente ciclo. Se debe elaborar un calendario “Niko-Niko” para poder informar y darse una idea grafica del estado emocional de los miembros del equipo de desarrollo (Mora et al., 2019).
3. **Fase 3: Entrega de versión – Entrega y soporte**
 - a. **Soporte a la ingeniería de la versión final:** el equipo de desarrollo elabora un acuerdo de soporte para la presentación de las versiones del MVITS (similar al acuerdo de nivel de servicio), en dicho acuerdo se deben tratar los siguientes puntos: descripción del servicio, calendario de actividades, métricas del servicio (disponibilidad, capacidad, seguridad), datos del cliente y recomendaciones al utilizar el servicio (Mora et al., 2019).

- b. **Reunión de entrega de la versión final:** el cliente presenta a los usuarios finales del servicio desarrollado. El experto presenta el paquete de la versión final del MVITS al cliente, dicho paquete debe abarcar los siguientes aspectos: acuerdo de soporte, documento de aceptación del lanzamiento de la versión, repositorio de la versión y la presentación de la versión. Por último, el cliente y el experto cierran el ciclo de liberación de la versión final con un acuerdo de aceptación (se debe firmar por ambas partes) (Mora et al., 2019).

Artefactos

- 1) Paquete backlog MVITS (PKG.01)
- 2) Paquete sprint MVITS (PKG.02)
- 3) Paquete de versión de entrega MVITS (PKG.03)

3.1.2 Fundamentos de Ciencia de Datos y Analítica de Negocios

3.1.2.1 Conceptos Básicos (Ciencia de Datos, Analítica de Negocios, Tipos Descriptivo, Predictivo y Prescriptivo de Analíticas, Fases de la Ciencia de Datos (Data Science Pipeline)).

Ciencia de Datos

Desde principios del siglo actual, se pronosticaba el gran valor que poseían los datos y su fuerte potencial que podrían tener en tomas de decisiones empresariales. Sin embargo, gran parte del sector no tomaba con seriedad la nueva ciencia ya que se enfocaban en los métodos que se conocían por su éxito y funcionalidad. Posteriormente se comprobaron las grandes conexiones que se podían aprovechar y convertirse en ventajas competitivas gracias a la exploración de la información. Fue entonces, que el sector empresarial reconoció su significancia lo que desencadenó un aumento drástico en la demanda de profesionales calificados en dicha área. Por lo tanto, hoy en día la escena empresarial está intentando destacarse y/o competir usando sus grandes volúmenes de datos (en estos días, la cantidad de información que se genera empresarial y personalmente es abrumadora, todo gracias a la basta cantidad de TI a nuestro alcance), encontrando en ellos una gran oportunidad dependiendo de la calidad

de esfuerzos analíticos que se empleen. También es importante mencionar las grandes aportaciones en el desarrollo de tecnología que otorgan herramientas necesarias para desempeñar el procesamiento, visualización u otras fases involucradas en esta ciencia (Thomas H. Davenport & DJ Patil, 2012).

Vivimos en la era de la "ciencia de datos y la analítica avanzada", en la que casi todo en nuestra vida cotidiana se registra digitalmente como datos. Los datos pueden ser estructurados, semiestructurados o no estructurados, lo que aumenta día a día. La ciencia de datos suele ser un "concepto para unificar la estadística, el análisis de datos y sus métodos relacionados" para comprender y analizar los fenómenos reales con datos. La ciencia de datos es la ciencia de los datos o el estudio de los datos, donde un producto de datos es un entregable de datos, habilitado o guiado por datos, que puede ser un descubrimiento, una predicción, un servicio, una sugerencia, una visión para la toma de decisiones, un pensamiento, un modelo, un paradigma, una herramienta o un sistema. La ciencia de datos aumenta día a día (Figura 4) según los datos con las tendencias de Google de los últimos 5 años. Por lo general, la ciencia de datos es el campo de la aplicación de métodos analíticos avanzados y conceptos científicos para obtener información empresarial útil a partir de los datos (Sarker, 2021).

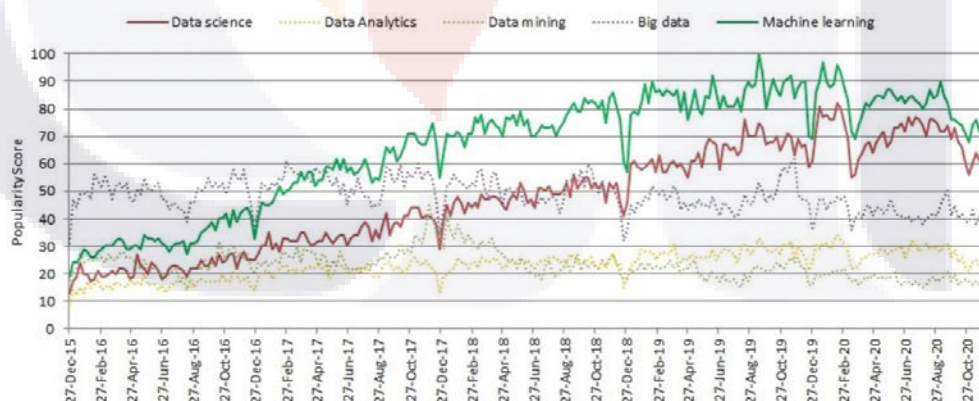


Figura 4: Incrementos de conceptos populares tales como: ciencia de datos, analítica de datos, minería de datos, entre otros. Se muestra su comportamiento al pasar los meses.

El área de la ciencia de datos es enorme, y nos centramos principalmente en obtener información útil mediante análisis avanzados, cuyos resultados se utilizan para tomar decisiones inteligentes en diversas áreas de aplicación del mundo real. En seguida se

muestran conceptos básicos definidos desde el punto de la ciencia de datos, estos términos son críticos para comprender el campo de estudio (Sarker, 2021):

- **Análisis de datos** se refiere al tratamiento de datos mediante teorías, tecnologías y herramientas convencionales para extraer información útil y con fines prácticos (Sarker, 2021).
- **Analítica de datos**, por su parte, se refiere a las teorías, tecnologías, instrumentos y procesos que permiten una comprensión y exploración en profundidad de la información procesable. El análisis estadístico y matemático de los datos es la principal preocupación en este proceso (Sarker, 2021).
- **Minería de datos** es otro término popular en la última década, que tiene un significado similar con varios otros términos como minería de conocimiento a partir de datos, extracción de conocimiento, descubrimiento de conocimiento de datos (KDD), análisis de datos/patrones, arqueología de datos y dragado de datos (Sarker, 2021).
- **Big data** es otro término popular hoy en día, que puede cambiar los enfoques estadísticos y de análisis de datos, ya que tiene las características únicas de masivo, de alta dimensionalidad, heterogéneo, complejo, no estructurado, incompleto, ruidoso y erróneo. Los big data pueden ser generados por dispositivos móviles, redes sociales, Internet de las Cosas, multimedia y muchas otras aplicaciones nuevas (Sarker, 2021).

Analítica avanzada

La analítica avanzada consiste más en anticiparse al uso de los datos para detectar patrones que permitan determinar qué es probable que ocurra en el futuro. La analítica avanzada es un paso adelante para ofrecer una comprensión más profunda de los datos y ayudar a analizar los datos granulares, que son los que nos interesan. En el campo de la ciencia de datos, son populares varios tipos de analítica, como la **analítica descriptiva**, que responde a la pregunta de qué ha ocurrido; la **analítica diagnóstica**, que responde a la pregunta de por qué ha ocurrido; la **analítica predictiva**, que predice lo que ocurrirá en el futuro; y la **analítica prescriptiva**, que prescribe qué medidas

deben tomarse. Diversos métodos analíticos avanzados, como el modelado de aprendizaje natural, el análisis de sentimientos, las redes neuronales o las redes neuronales o el análisis de aprendizaje profundo datos y, por lo tanto, pueden utilizarse para desarrollar aplicaciones inteligentes basadas en datos. La analítica avanzada da un paso adelante al ofrecer una comprensión más profunda de los datos y ayuda a analizar datos granulares. La analítica avanzada se caracteriza o define como el análisis autónomo o semiautónomo de datos o contenidos mediante técnicas y métodos avanzados para descubrir perspectivas más profundas, predecir o generar recomendaciones, normalmente más allá de la inteligencia o analítica empresarial tradicional (Sarker, 2021).

La analítica facilita lograr los objetivos empresariales mediante la recopilación de datos para analizar tendencias, crear modelos predictivos para estar preparados para futuros problemas y analizar y/u optimizar procesos empresariales dentro de una organización. Retomando las tres principales categorías (Figura 5) de la analítica (descriptiva, predictiva y prescriptiva) es importante analizarlas con mayor profundidad ya que pueden ser los pilares a la hora de desarrollar un nuevo servicio. **La analítica descriptiva** utiliza la información para responder preguntas tales como “¿Qué paso?” o “¿Qué está pasando?”. El principal objetivo de este tipo de analítica es identificar oportunidades o problemas respecto al negocio. Por otro lado, tenemos a **la analítica predictiva** la cual emplea técnicas matemáticas en los datos para descubrir patrones explicativos y predictivos los cuales representan las relaciones existentes entre los datos de entrada y los de salida. De igual forma este tipo de analítica busca responder ciertas preguntas las cuales se enfocan en “¿Qué va a pasar?” y “¿Por qué va a pasar?”, además, su principal resultado que se busca al usar algún modelo predictivo es tener una proyección clara de lo que puede pasar en el futuro y el porqué de ello. Por último, **la analítica prescriptiva** implementa algoritmos matemáticos para determinar acciones alternativas o decisiones dadas por complejos grupos de objetivos, requerimientos y limitaciones con el objetivo de mejorar el rendimiento empresarial (Delen & Demirkan, 2013).

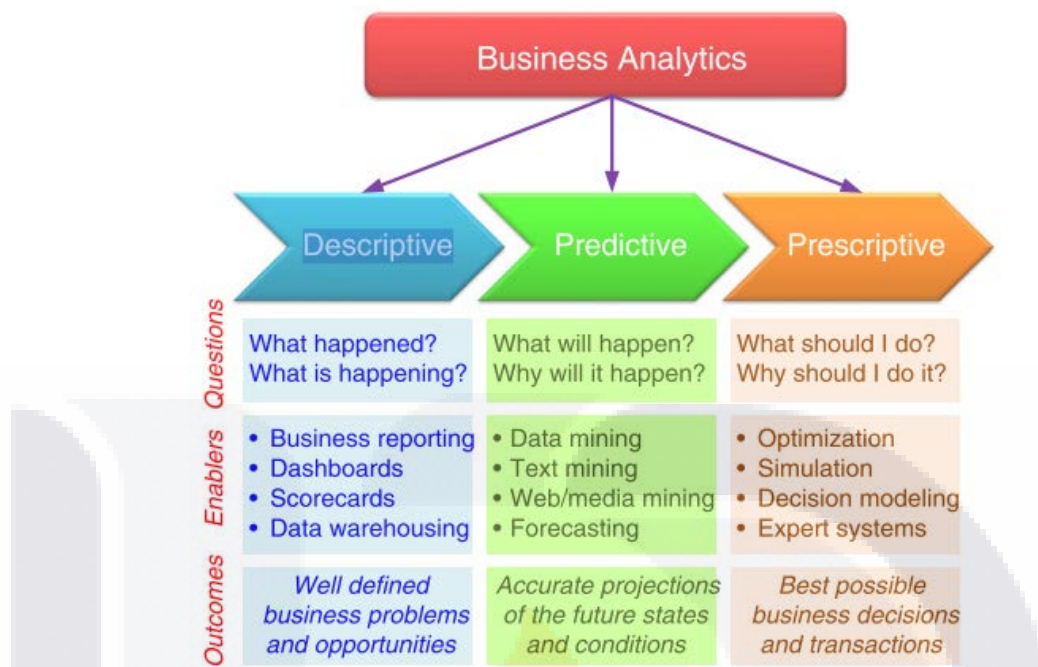


Figura 5: Los tres principales tipos de analítica y sus puntos principales de cada una de estas.

3.1.2.2 Fase, Propósitos y Ejemplos de Herramientas Open-Source de la Ciencia de Datos (Data Science Pipeline).

Un sistema basado en Ciencia de Datos cuenta con 5 fases habituales, estas fases son las siguientes:

- 1) **Identificar fuentes y adquirir datos brutos:** Identificar el conjunto de fuentes de datos brutos para la canalización de análisis de Big data, acordar legalmente su accesibilidad, recopilar los datos brutos acordados, transmitirlos y registrarlos. Principales tareas: Identificación de las fuentes de disponibles, Análisis de las fuentes de datos brutos disponibles, Seleccionar y acordar las fuentes de datos, Recoger y transmitir los datos brutos y, Registrar los datos. Principales problemas: Variedad de formatos (texto, imágenes, video, etc.), Velocidad (indicies de generación de datos), Volumen (tamaño de los registros de datos), Veracidad (nivel de confianza de los registros), Valor (necesidad empresarial del registro de datos) (Mora et al., 2023).

- 2) **Preprocesamiento de los registros de información:** Aplicar métodos de preprocesamiento a los registros de datos. Principales tareas: preprocesar los datos usando técnicas de normalización, comprimir, descomprimir, limpieza, eliminación de falsos positivos, transformación, entre otras según sea pertinente para el estudio. Problemáticas: riesgos de seguridad de la información y rendimiento de las plataformas de preprocesamiento (Mora et al., 2023).
- 3) **Almacenamiento y procesamiento de los datos:** Es fundamental quedarse solamente con los datos de interés para el caso de estudio, por lo que es necesario aplicar técnicas de procesamiento adecuadas para posteriormente almacenar esta información en plataformas diseñadas para tal tarea. Principales tareas: integrar, agregar y representar los datos, replicar los datos y procesar los datos. Problemáticas: se debe tener en cuenta el rendimiento con el que cuente la plataforma de almacenamiento, además de conocer si se cuenta con almacenamiento en la nube. Por otro lado, también se debe tener en cuenta el rendimiento que otorgue la plataforma seleccionada para desempeñar la tarea de procesar los datos de interés. Por último, se debe tener en cuenta el ámbito de la seguridad de la información y asegurar la privacidad de los datos de ser necesario (Mora et al., 2023).
- 4) **Modelado y análisis de los datos:** esta fase se caracteriza por enfocarse en elaborar modelos basados en los datos y aplicar estos procedimientos analíticos a las metas específicas del negocio. Principales tareas: hacer un análisis exploratorio y descriptivo (procesamiento analítico en línea (OLAP), estadística descriptiva), aplicar un análisis predictivo (clasificación, regresión, clusterización, asociación) y, por último, apoyarse en un análisis prescriptivo (optimización, simulación, métodos heurísticos, sistemas expertos). Problemáticas: rendimiento de procesamiento de los servidores de plataformas encargadas de desempeñar métodos de analítica. Clasificación y desempeño de los métodos de análisis exploratorio-descriptivo, predictivo y/o prescriptivo (Mora et al., 2023).
- 5) **Acceso y uso de los datos:** Utilizar modelos basados en datos en sistemas autónomos y/o integrados en sistemas de control automáticos o de usuario final

para objetivos empresariales específicos. Principales tareas: análisis visual interactivo, desarrollo de un sistema de ciencia de datos para el usuario final y desarrollo de un sistema de control automático de datos. Problemáticas: sistemas de transmisión de información, métricas de rendimiento, facilidad de uso y metas empresariales (Mora et al., 2023).

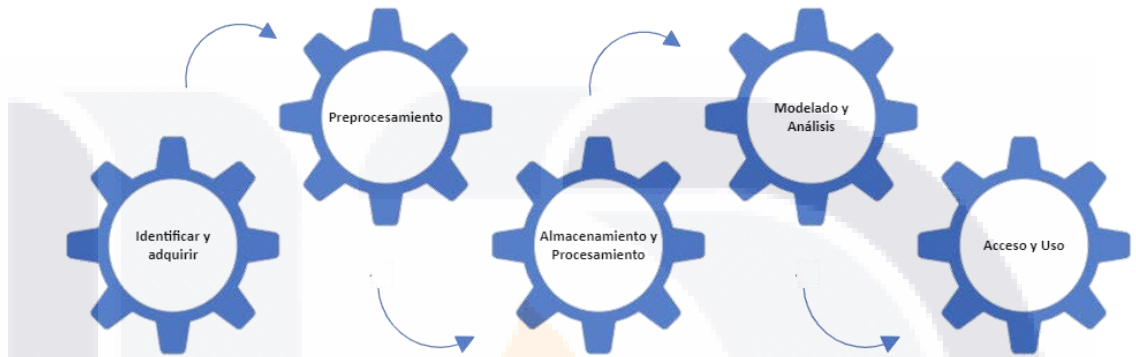


Figura 6: Las cinco etapas genéricas dentro del proceso de Ciencia de Datos.

3.1.2.3 Análisis y Selección de las Herramientas Open-Source de las Fases de la Ciencia de Datos (Data Science Pipeline).

Es ampliamente reconocido que los **recursos de código abierto** son de suma importancia en el avance de la ciencia (prácticamente cualquier ciencia) por lo que es crucial para la investigación. El modelo único de desarrollo del código abierto se caracteriza por la colaboración masiva distribuida, la revisión por pares, el desarrollo incremental y la evolución mediante la rápida publicación de versiones. La participación de los usuarios es la piedra angular del éxito del código abierto, y es también donde se difumina la distinción entre usuarios y desarrolladores (Zhao & Deek, 2005).

A continuación, se mencionan una herramienta de código abierto por cada categoría (plataforma en la nube, gestión de clústeres de servidores de procesamiento, motores transmisión/procesamiento de eventos complejos(CEP)), además de tener en cuenta cada fase del desarrollo de un sistema de ciencia de datos:

1. Identificación y adquisición

a. *Plataforma en la nube*

Apache CloudStack: Plataforma código abierto desarrollada para poder implementar, gestionar y escalar infraestructura computacional en la nube. Es una excelente opción para las empresas gestionar redes públicas, privadas e híbridas a través de una interfaz de usuario centralizada. El servicio es gratuito y adaptable a la infraestructura actual de cada empresa, logrando ser accesible para organizaciones con recursos limitados (The Apache Software Foundation, 2020).

b. *Clústeres de servidores de procesamiento*

Apache Mesos: Servicio TI para gestión/administración de clústeres de código abierto desarrollado para poder aglomerar recursos computacionales de diferentes servidores en uno solo para facilitar su monitoreo. Su trabajo es similar a un sistema operativo administrando los recursos informáticos destinados a las aplicaciones. Este servicio permite ejecutar diferentes aplicaciones en un mismo conjunto de servidores, además de ser idóneo para trabajar con grandes volúmenes de información (Hindman et al., n.d.).

c. *Máquinas de streaming*

Apache Storm: Plataforma de código abierto desarrollada para poder trabajar con información en tiempo real. Cuenta con las siguientes características: escalabilidad, tolerancia a fallas y ofrece compatibilidad con otros servicios aumentando su capacidad de análisis de grandes conjuntos de datos. Este servicio TI es ideal para casos en los que las aplicaciones web necesiten tiempos de respuesta estables, además de que garantiza continuidad de procesamiento en caso de posibles fallas (Apache Software Foundation, 2023a).

2. Preprocesamiento

a. *Gestión de clústeres de servidores de almacenamiento*

Apache Hadoop/HDFS: Software de código abierto desarrollado para el almacenamiento y análisis de grandes volúmenes de información en clústeres

de computadoras. Este software TI es idóneo en casos en los que se tenga que trabajar con datos estructurados, semiestructurados y no estructurados, además de ofrecer escalabilidad y cierta tolerancia a fallas. Por último, su implementación reduce costos ya que usa hardware estándar (International Business Machines (IBM), 2023).

b. Plataformas de gestión de lagos de datos

Delta Lakehouse: Es una plataforma de código abierto para el almacenamiento de datos que combina las mejores características entre los centros de datos tradicionales (data warehouses) y los lagos de datos modernos (data lakes). Ofrece un entorno de trabajo adecuado para almacenar, procesar, analizar datos estructurados, semiestructurados y no estructurados manteniendo la integridad de los datos (Databricks, 2023).

c. Almacenes de macrodatos:

Apache Hive: Servicio TI diseñado para el almacenamiento de información compatible con Apache Hadoop. Ofrece una interfaz de usuario similar a la de SQL para hacer consultas y analizar considerables volúmenes de datos, se recomienda el análisis por lotes, es muy amigable para usuarios no técnicos/inexpertos gracias al tipo de lenguaje similar a SQL. También ofrece ciertas características para la mejora de rendimiento al momento de realizar una consulta en conjuntos de datos enormes (Apache Software Foundation, 2023b).

3. Almacenamiento y procesamiento

a. Plataformas de reducción de mapeo SQL

Apache Impala: Servicio TI desarrollado para realizar consultas SQL de código abierto diseñado para ofrecer análisis de datos y presentarlos de forma interactiva, además de realizar procesos a alta velocidad sobre las bases de datos para contar con una baja latencia (The Apache Software Foundation, 2023b).

b. Plataformas no-SQL

Apache Hbase: Gestor de bases de datos de código abierto desarrollado para trabajar con volúmenes de datos estructurados considerablemente grandes en tiempo real. Cuenta con una gran capacidad de almacenamiento y otorga acceso

rápido al usuario en cuanto a consultas se refiere (The Apache Software Foundation, 2023a).

c. Plataformas de procesamiento analítico en línea (OLAP)

Apache Kylin: Plataforma de análisis OLAP (Procesamiento Analítico en Línea) de código abierto desarrollada con el objetivo de analizar grandes volúmenes de datos compatible con Hadoop. Principalmente esta pensada para la aplicación en Big Data en aplicaciones de Inteligencia de Negocios ya que permite búsquedas en bases de datos de gran volumen con tiempos de respuesta muy bajos (Apache Software Foundation, 2023c).

4. Modelado y análisis

a. Librerías y/o plataforma de analítica para Big Data

Apache Mahout: Plataforma de código abierto desarrollada para poder crear algoritmos de aprendizaje máquina (aprendizaje automático). Objetivamente es una librería que permite implementar algoritmos de filtrado y clasificación para la toma de decisiones y minería de datos, sin embargo, hoy en día se puede utilizar la plataforma para el diseño de algoritmos de aprendizaje supervisado y no supervisado (The Apache Software Foundation, 2023c).

Como se observa existen múltiples herramientas open source de distintas categorías para cada fase y/o etapa del proceso que implica realizar un sistema de Ciencia de Datos, teniendo en cuenta los objetivos de este trabajo práctico se optó por enfocarse únicamente en tres etapas: **Preprocesamiento, Modelado y Análisis, Acceso y uso**. Es importante mencionar que la primera etapa **Identificación y adquisición** no se toma en consideración ya que el proyecto no se enfoca en la creación de base de datos, por lo general las empresas (no importando su tamaño o clasificación) ya cuentan con datos relevantes a su giro empresarial y serán estos los que se deben considerar para la toma de decisiones empresariales.

Por lo anterior, se tomó la decisión de unificar la fase de Preprocesamiento y Almacenamiento en una misma (esto ayuda a no tener dos herramientas y ahorrar en recursos computacionales), para esta fase se seleccionó la **herramienta MySQL**: ofrece un servidor de base de datos SQL (lenguaje de consulta estructurado) muy

rápido, multiproceso, multiusuario y robusto. MySQL es un sistema de gestión de bases de datos, dado que las computadoras son muy buenas para manejar grandes cantidades de datos, los sistemas de gestión de bases de datos desempeñan un papel central en la informática, como utilidades independientes o como parte de otras aplicaciones. Las bases de datos MySQL son relacionales, Una base de datos relacional almacena datos en tablas separadas en lugar de colocar todos los datos en un gran almacén. Las estructuras de la base de datos están organizadas en archivos físicos optimizados para la velocidad (Oracle Corporation, 2023).

Para la etapa de Modelado y análisis se decidió implementar la **herramienta RStudio** (dicha herramienta interactúa con el lenguaje de programación R) ya que es perfecta para todo el apartado de estadística y es una excelente aproximación a la Ciencia de Datos para aquellas personas inexpertas del área. Es un software de código abierto de gran calidad para Ciencia de Datos para enriquecer el conocimiento en dicha área. Además, se destaca su versatilidad al ser compatible con varios sistemas operativos: Windows, MacOS, Ubuntu, entre otros. También cabe mencionar que esta opción se acopla excelente al **Jupyter Notebook** ya que se cuenta con el kernel de R, esta opción nos deja acceder a los comandos predefinidos en RStudio. Debemos recalcar en que se usará el lenguaje y entorno gráfico R el cual proporciona una amplia variedad de técnicas estadísticas: modelado lineal y no lineal, pruebas estadísticas clásicas, análisis de series temporales, clasificación, agrupamiento, etc. Uno de los puntos fuertes de R es la facilidad con la que se pueden producir gráficos bien diseñados con calidad de publicación, incluidos símbolos y fórmulas matemáticas cuando sea necesario.

Por último, para la etapa de Acceso y uso (etapa en la que el usuario final podrá observar con más detenimiento lo obtenido del procesamiento, modelado y análisis de los datos empresariales) se seleccionó la herramienta **Jupyter Notebook**, dicha herramienta puede ser bastante amigable y de fácil entendimiento para usuarios finales gracias a su capacidad de crear “notebooks” el cual es un documento que se puede compartir que combina código de computadora, descripciones en lenguaje sencillo, datos, visualizaciones enriquecidas como modelos 3D, cuadros, gráficos y figuras, y controles interactivos. Jupyter notebook proporciona un entorno interactivo rápido para

crear prototipos y explicar código, explorar y visualizar datos y compartir ideas con otros. Esta herramienta nos permite implementar un **kernel de Python**, el cual es un lenguaje de programación muy usado en el ámbito de Ciencia de datos ya que cuenta con bastantes librerías dedicadas a esta rama de la ciencia que se pueden aprovechar e implementar en proyectos de la índole. También cabe mencionar que esta herramienta se complementa o interactúa muy bien con la herramienta seleccionada para la etapa previa Modelado y análisis (Jupyter Team, 2023).

Retomando las fases generales del proceso de Ciencia de datos se muestra el siguiente diagrama de distribución (Figura 7) con el que se especifica en que proporción se aplicaran las herramientas de código abierto (open-source) dentro del diseño de la plataforma web. También cabe mencionar que tanto RStudio cliente como Jupyter Notebook Server tienen excelente mancuerna entre ellos ya que acorde con diferentes referencias es posible ejecutar Scripts tanto de Python como de R en ambos entornos de desarrollo, esto otorga versatilidad al momento de elaborar el curso de capacitación para el sector de Tecnologías de Información (cliente) para que se pueda entender comandos o instrucciones que pudieran ser difíciles para un personal que no esté acostumbrado a un entorno de programación.

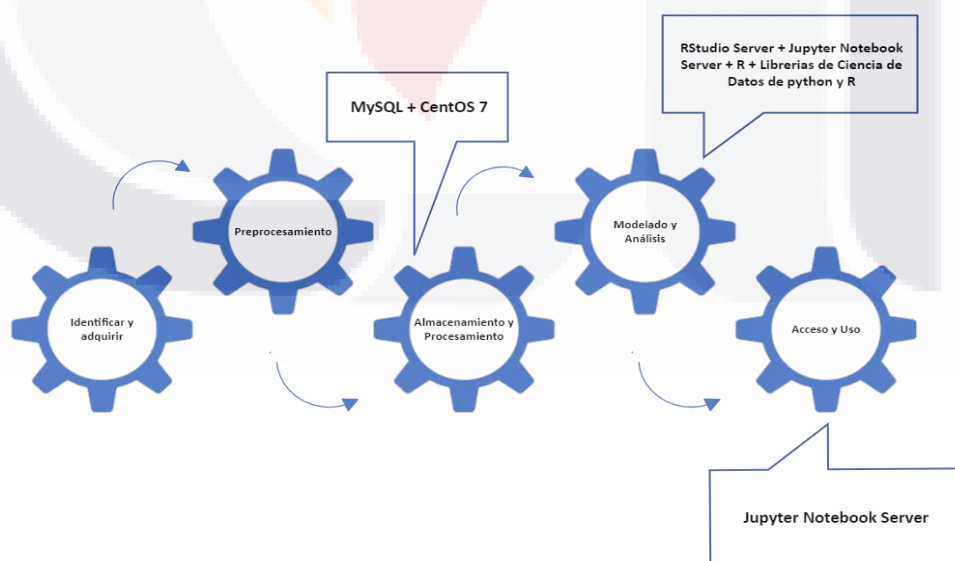


Figura 7: Distribución de aplicación de las herramientas Open-Source seleccionadas acorde a las fases genéricas de Ciencia de Datos.

3.2 Revisión de Intervenciones Similares.

3.2.1 Caso Similar #1 (Hosea A. Ofe, Mark de Reuver, Bouke Nederstigt y Marijin Janssen)

Descripción General

Nombre del artículo: “Plataformas de Análisis de Datos: Propuestas de Valor y Retos de Adopción para Negocios Pequeños de Hotelería”.

Los autores redactan de forma corta y concisa los términos pertinentes al momento de introducir al lector dentro del ámbito de la Analítica de Datos, conceptos tales como la definición, características generales, tipos, etc., sobre la ‘DA’ (Analítica de Datos, por sus siglas en inglés) son recursos fundamentales. Además, hacen la distinción entre los aportes destinados a empresas consolidadas que cuentan con grandes recursos para su inversión en DA, y las organizaciones que entran dentro de rubro de PYMES, las cuales son empresas recién fundadas que luchan por su supervivencia tratando de visualizar objetivos a largo plazo (Ofe et al., 2023).

El escrito se centra en dos objetivos:

- Comprender el valor de las propuestas que ofrece la Analítica de Datos para los negocios pequeños.
- Identificar los retos que los negocios pequeños encuentran al momento de implementar plataformas de ‘DA’.

Caso de estudio: plataforma de Analítica de Datos destinada para empresas pequeñas de hostelería lanzada por la organización *Checkmetrix*.

País: Holanda.

Año: 2023.

Consultores: Hosea A. Ofe, Mark de Reuver, Bouke Nederstigt y Marijin Janssen.

Problemática central

El sector hotelero holandés ha tardado en adaptar plataformas de Analítica de Datos debido a la diversidad de sistemas de puntos de ventas que ofrecen los proveedores, con una existencia de 130 proveedores registrados los cuales ofrecen más de 250 sistemas diferentes, sin mencionar que muchos de estos sistemas carecen de una interfaz abierta para acceder a los datos de venta (Ofe et al., 2023).

Solución

Los investigadores se centran en una plataforma de *DA* específica para poder entender de mejor manera el proceso de implantación y su posible adopción por empresas pequeñas. Dicha plataforma se caracteriza por su diseño destinado a la resolución del problema que atormenta a pequeños negocios como bares y restaurantes, el poder obtener información de sus ventas, también se hace la mención de poder conectar la plataforma a cualquier sistema de ventas. **Checkmetrix** comprende varios módulos para el tratamiento de la información obtenida por los sistemas de punto de venta (Ofe et al., 2023):

- Recopilar
- Limpiar
- Almacenar
- Categorizar
- Analizar

En la figura 8 se muestra el diagrama de la arquitectura que compone la plataforma.

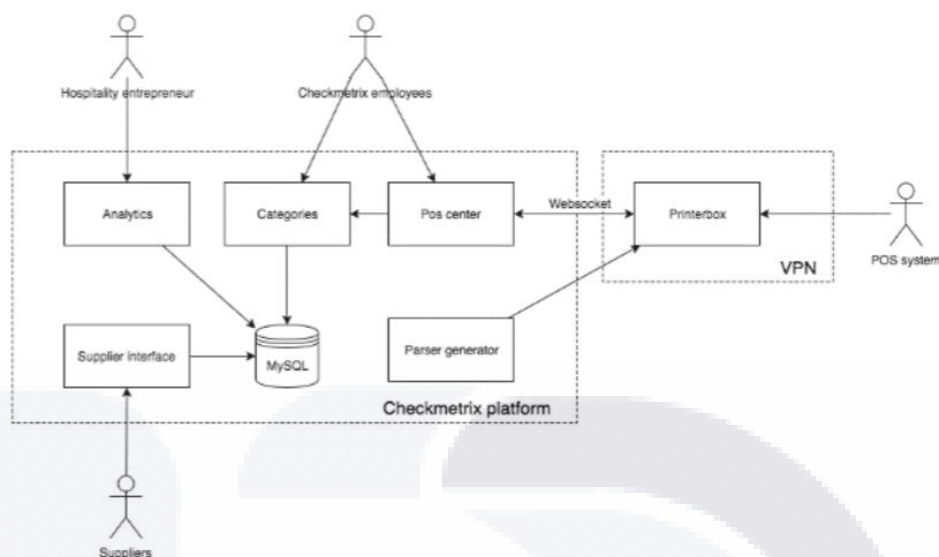


Figura 8: Arquitectura de la plataforma Checkmetrix (tomada de “Data Analytics Platforms: Value Propositions and Adoption Challenges for Small Hospitality Businesses”).

3.2.2 Caso Similar #2 (Lakshit Sharma, Dr. Vijay Gupta)

Descripción General

Nombre del artículo: “Un estudio comparativo de las herramientas de análisis de datos de código abierto”.

El objetivo de este artículo es poder realizar una comparativa fiable entre distintas plataformas de código abierto que se usan en la Analítica de Datos, la comparación abarca la interfaz, funcionalidad y los algoritmos que se pueden usar en cada una (Sharma & Gupta, 2019).

Caso de estudio: comparativa entre seis herramientas (Orange, Rapid Miner, Knime, Weka, Keel y R) de ‘DA’ donde se describe cada una de ellas, además de mostrar su interfaz, los recursos con los que cuenta y los algoritmos que se pueden implementar (Sharma & Gupta, 2019).

País: India.

Año: 2019.

Consultores: Lakshit Sharma y Dr. Vijay Gupta.

3.2.3 Caso Similar #3 (Dmitry Medvedev, Gerard Lemson y Mike Rippin)

Descripción General

Nombre del artículo: “SciServer Compute: Trayendo el Análisis Cerca de la Información”.

El proyecto SciServer Compute es desarrollado como una infraestructura modular y escalable para el almacenamiento, acceso, consulta y procesamiento de enormes conjuntos de datos científicos de nivel Petabyte (PB), que son equivalentes a 1000 Terabytes (TB) o 1000000 Gigabytes (GB). Cabe mencionar que este sistema fue construido en base a la infraestructura existente del proyecto astronómico Sloan Digital Sky Survey (SDSS). Metas principales del proyecto (Medvedev et al., 2016):

- Ser capaz de acceder y procesar grandes volúmenes de datos.
- Proveer un espacio de almacenamiento personalizado, donde el usuario pueda almacenar sus resultados de búsqueda en una base de datos y sus resultados de procesos en otra.
- Proveer un entorno colaborativo, donde los usuarios puedan compartir sus bases de datos con otros.
- Proveer capacidad de escalar los recursos computacionales para el procesamiento y análisis de estos volúmenes abrumadores de datos.

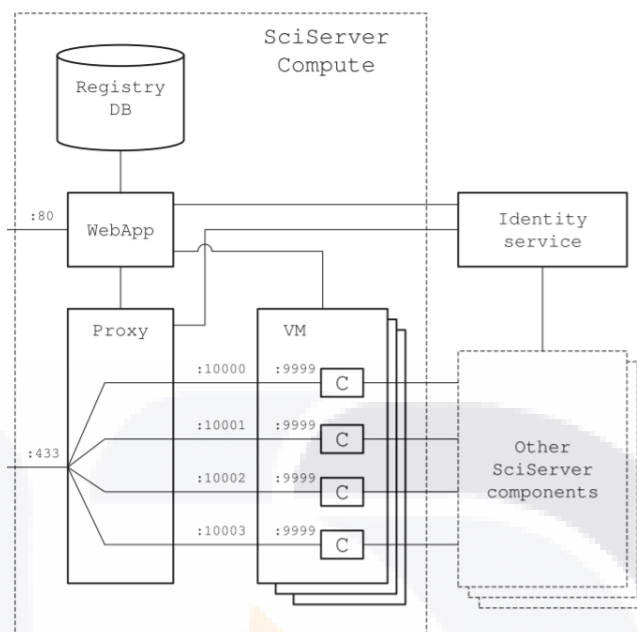


Figura 9: Arquitectura de SciServer Compute y su interacción con diversos componentes (tomada de “SciServer Compute: Bringing Analysis Close to the Data”).

Organización cliente: Universidad Johns Hopkins.

País: Estados Unidos de América.

Año: 2016.

Consultores: Dmitry Medvedev, Gerard Lemson y Mike Rippin.

3.2.4 Caso Similar #4 (Juan Cruz Pantano y María Romagnano)

Descripción General

Nombre del artículo: “Herramientas de software libre para el análisis y la inteligencia empresarial”.

Los autores dejan en claro la importancia y el potencial que las plataformas de BI (Inteligencia de Negocios, por sus siglas en inglés) y de BA (Analítica de Negocios, por sus siglas en inglés) pueden otorgar a las Pequeñas y Medianas Empresas (PYMES), este tipo de organizaciones dependen bastante de las decisiones empresariales que puedan tomar y al no aprovechar su información o contar con los

datos erróneos puede ocasionar pérdidas y vulnerabilidades dentro de su modelo de negocios. Por otra parte, el escrito proporciona un ejemplo de la implementación de un conjunto de herramientas (PowerBI, Orange y Knime) destinadas a la exploración y visualización de la información dentro de una empresa argentina ubicada en San Juan (empresa clasificada en el rubro de Mediana Empresa) con el objetivo de mostrar los beneficios que puede proporcionar el uso de BI y minería de datos aplicada a sus archivos de Excel existentes (Pantano & Romagnano, 2022).

Organización cliente: Empresa local de transportes ubicada en San Juan, Argentina.

País: Argentina.

Año: 2022.

Consultores: Juan Cruz Pantano y María Romagnano.

3.2.5 Caso Similar #5 (Fangjin Yang, Gian Merlino, Nelson Ray, Xavier Léauté, Himanshu Gupta, Eric Tschetter)

Descripción General

Nombre del artículo: “RADStack: Arquitectura Lambda de Código Abierto para Análisis Interactivos”.

Los investigadores proponen una arquitectura denominada lambda, la cual está constituida por un procesamiento por bloques/lotos, procesamiento por transmisión (streaming) y una capa de servicio. En este tipo de arquitectura los datos de entrada son transmitidos simultáneamente a la capa de procesamiento por bloques y a la capa de procesamiento por streaming. Se hace énfasis en la implementación del recurso de código abierto RADStack ya que ofrece flexibilidad, baja latencia en consultas analíticas de datos casi instantáneamente (Yang et al., 2017).

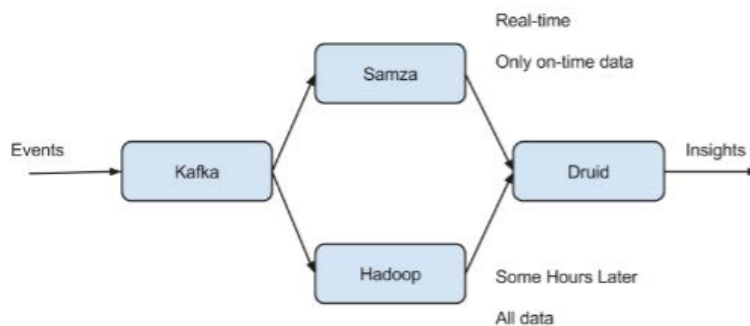


Figura 10: Componentes esenciales de la plataforma RADStack (tomada de “The RADStack: Open-Source Lambda Architecture for Interactive Analytics”).

3.3 Contribuciones y Limitaciones.

3.3.1 De Teoría Base

En el presente trabajo práctico se tomarán como base las guías de mejores prácticas para la Gestión de Servicios de TI: ITILv3 e ITILv4, ya que abarcan excelentes estrategias generales para cada fase del ciclo de vida del servicio de TI (se hace un énfasis en la etapa de Diseño que es el objetivo principal para desarrollar) que enriquecen los procesos del negocio y pueden ser fundamentales para la organización. Sin embargo, cabe mencionar que la metodología principal con la que se desarrollará el entorno virtual para desempeñar la Ciencia de Datos es la propuesta como AVS_FD_MVITS_v1-2022.

Estos marcos de procesos describen los roles, actividades y artefactos que influyen en la etapa de Diseño del servicio para su correcto desarrollo se debe tener en mente proporcionar utilidad y garantía ya que son dichos aspectos los que otorgan valor a un servicio, obvio teniendo en cuenta los criterios tanto del cliente como los de la empresa.

Metodología	Contribuciones	Limitaciones
ITIL v3 2011	Proporciona una guía de buenas prácticas que ayuda en el Diseño de Servicios teniendo en cuenta los procesos internos y políticas de la organización. Se realiza un diagnóstico de la	Practicas muy generales y falta de profundidad en la infraestructura de soporte de los servicios.

	infraestructura con la que se cuenta para que solo se requieran actualizaciones mínimas para garantizar el funcionamiento de los servicios.	
ITIL v4 2019	Nueva variante de un marco de Gestión de Servicios con una nueva perspectiva. Actualiza gran parte de las practicas ITSM en un contexto más amplio respecto a la perspectiva del cliente, los flujos de valor y la transformación digital. También se adoptan nuevas formas de trabajo.	Aplicar ITL a una empresa significa designar personal y recursos para su planificación sin garantizar resultados inmediatos, sin mencionar la necesidad de personal capacitado ya que estos marcos de trabajo requieren conocimiento teórico y terminología específica. Todos estos aspectos pueden estar fuera del alcance para las PYMES.
AVS_FD_MVITS_v1-2022	Metodología ágil que toma en cuenta los escasos recursos informáticos dentro de una organización con pequeños o medianos flujos de información, presenta un diseño de servicios con características mínimas viables en un periodo de máximo 12 semanas.	Se necesita contar con un experto que domine la metodología. La infraestructura y recursos empresariales disponibles limitan la variedad de herramientas para garantizar el valor del servicio.

Tabla 4: Tabla de contribuciones y limitaciones de teoría base.

3.3.2 De Intervenciones Similares

Los trabajos similares revisados se centran en dos grandes paradigmas: 1) Recursos existentes de código abierto (se destaca la importancia de conocer las herramientas existentes, ventajas y desventajas) y 2) Desarrollo de una plataforma (al carecer de una opción que satisfaga las necesidades empresariales se opta por el desarrollo), sea cual

sea el propósito es evidente que existe cierta predisposición al uso y aplicación de Ciencia de Datos, Aprendizaje máquina, Minería de Datos, etc.

También, se puede apreciar que existe la preocupación por parte de los investigadores de otorgar herramientas competentes a las organizaciones que les permitan el aprovechamiento óptimo del cúmulo de datos que poseen. Sin embargo, carecen los estudios e investigaciones sobre empresas que estén dentro del espectro de Micro, pequeñas y medianas empresas (μ PYMES).

Intervención	Contribuciones	Limitaciones
1) Plataformas de Análisis de Datos: Propuestas de Valor y Retos de Adopción para Negocios Pequeños de Hotelería.	Presenta un caso de estudio real sobre la adopción de una plataforma de Analítica de Datos para un negocio pequeño. Se documentan los retos que supone el adoptar este tipo de tecnología.	• No se aborda sobre el desarrollo de la plataforma. • No se mencionan otras alternativas.
2) Un estudio comparativo de las herramientas de análisis de datos de código abierto.	Comparación de diversas herramientas para desempeñar la Analítica de Datos, se enfoca en mostrar la construcción de la plataforma y hacer distinción en las funciones que se pueden usar en cada una de ellas.	• No existe algún desarrollo de alguna aplicación. • Falta de información sobre compatibilidades entre aplicaciones o sistemas operativos. • Se deja fuera de lado la discusión del comportamiento de estas herramientas dentro de un entorno PYMES.
3) SciServer Compute: Trayendo el Análisis Cerca de la Información.	Proyecto de infraestructura para el análisis de Big Data (grandes volúmenes de información) desarrollado por la universidad Johns Hopkins, la cual integra grandes bases de datos	• El demo de la aplicación usa la infraestructura de una plataforma anterior (toma de datos). • Conocimiento muy específico que se aleja completamente del

	sobre los campos de la astronomía, cosmología, oceanografía, etc.	entorno empresarial (investigación).
4) Herramientas de software libre para el análisis y la inteligencia empresarial.	Comparación de diferentes herramientas de código abierto para desempeñar Inteligencia de Negocios y Analítica de Negocios con la distinción de aplicarlas a un caso local de una empresa PYMES de transportes.	○ No se aborda sobre el proceso de desarrollo de ninguna de las aplicaciones que se describen. Se ignora los beneficios de implementar una herramienta u otra, además el tema de compatibilidad (datos, sistemas, aplicaciones) no se aborda.
5) RADStack: Arquitectura Lambda de Código Abierto para Análisis Interactivos.	Introducción a la arquitectura que conforma el recurso de código abierto RADStack, además de discutir los métodos encargados de la analítica interactiva y el procesamiento flexible de los datos.	○ Se desconoce si una empresa PYMES podrá adoptar este tipo de herramienta y cuáles serían los principales retos al intentar aplicarla dentro del sector.

Tabla 5: Tabla comparativa de trabajos similares.

4. Diseño de Intervención al Caso Problema.

El desarrollo de la intervención al caso problema consiste en la construcción de una plataforma web de Ciencia de Datos de código abierto que le permita al usuario desempeñar tareas que involucren todas las fases del análisis de información, para ello se tomará como apoyo la metodología “AVS_FD_MVITS.v1-2022” (figura 11), la cual fue diseñada para la creación de un nuevo servicio informático mínimo viable, dicha metodología cuenta con ventajas tales como: la plataforma puede adaptarse a la infraestructura con la que cuente la empresa, cuenta con un tiempo de entrega de 4-12 semanas, además de proporcionar plantillas para ayudar con el diseño, construcción y entrega del servicio.

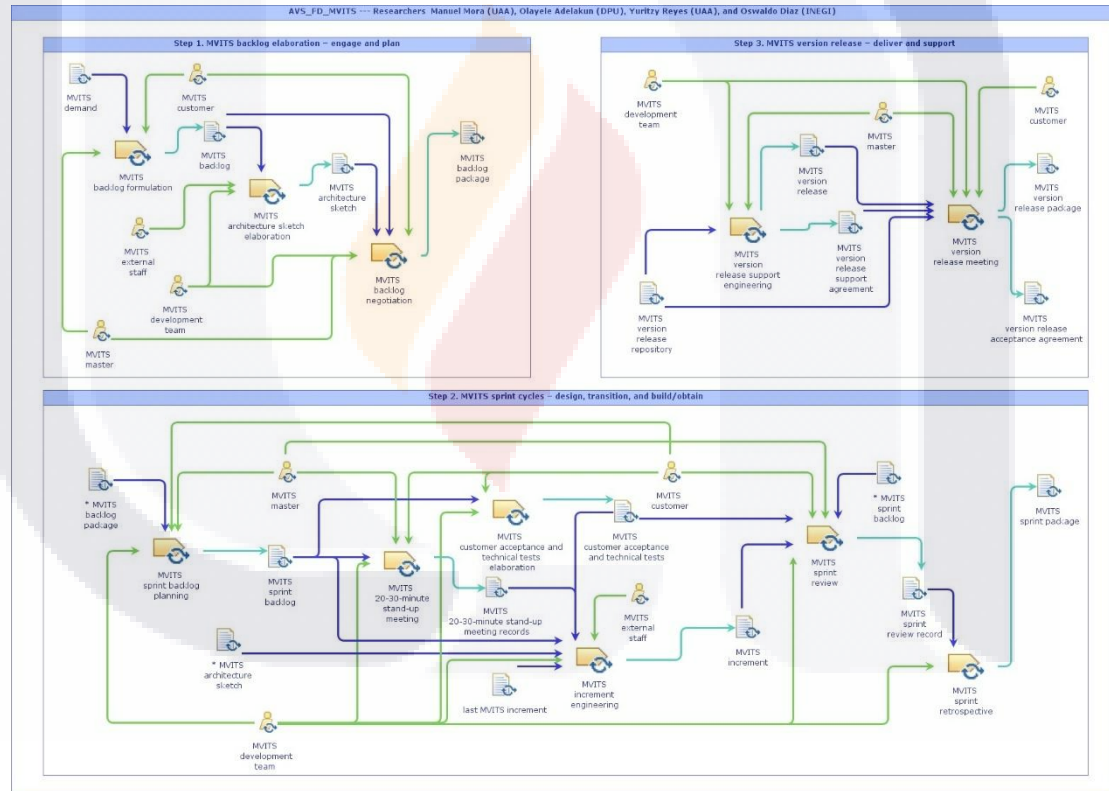


Figura 11: Diagrama de la metodología AVS_FD_MVITS desarrollada por los investigadores --- Manual Mora, Olayele Adelakun, Yuritzy Reyes y Oswaldo Díaz.

4.1 Etapa 1. Elaboración del Backlog de MVITS: involucramiento y planificación

Esta etapa inicial se compone de tres tareas: T.1.1 Formulación del backlog, T.1.2 Elaboración de un boceto de arquitectura y T.1.3 Negociación del backlog (figura 12), dicha etapa se inicia tomando en cuenta las demandas del cliente (problemática, necesidades y requerimientos generales) para posteriormente poder construir el paquete backlog del MVITS (artefacto: PKG.01), el cual se compone por los siguientes elementos: un diagrama de contexto, diagrama de casos de uso, hoja de backlog, diagrama de componentes TI, matriz de asignación funcional, matriz de asignación no funcional y pruebas de viabilidad. Esta etapa tiene un tiempo límite de entrega de una semana y los roles involucrados en ella son: Cliente (empresa Sitis De Aguascalientes S.C., Mtro. Jorge Eduardo Macias Luevano), Experto (Dr. José Manuel Mora Tavarez), Equipo de desarrollo (Ing. Juan Daniel Gallegos Rodríguez) y Personal externo (Dra. Yuritzy Reyes Delgado).

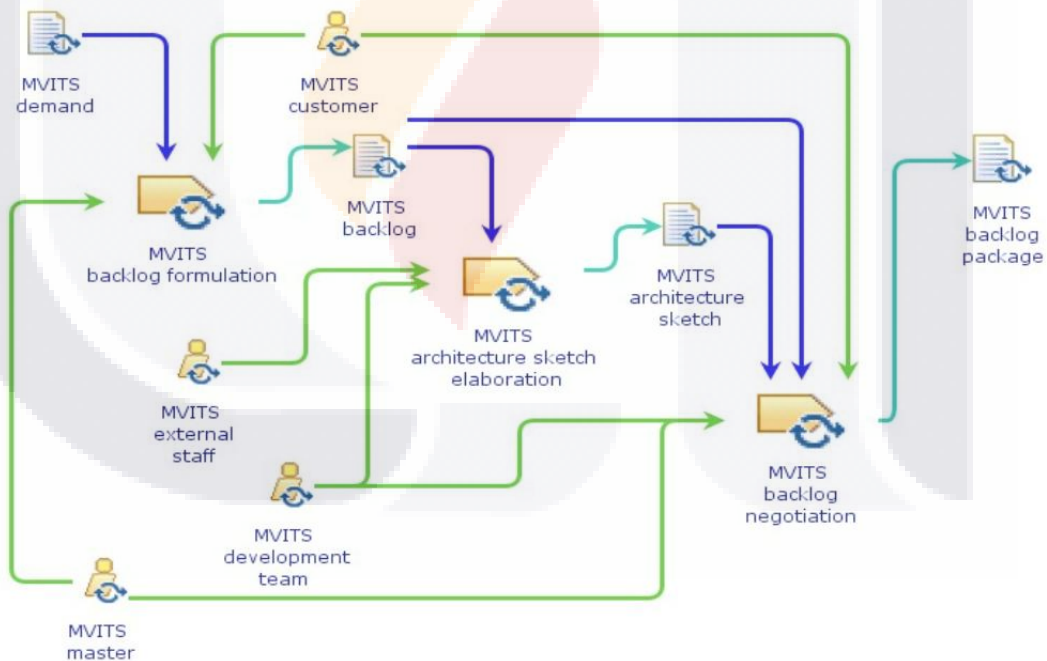


Figura 12: Etapa inicial de Involucramiento y Planificación del proyecto (tomada de “Metodología Ágil para el Desarrollo de un Nuevo Servicio de TI Mínimo Viable --- AVS_FD_MVITS.v1-2022”).

4.1.1 Tarea: T.1.1 Formulación del backlog

T.1.1 Formulación del Backlog	
Prácticas ágiles	Historias de usuario
Tiempo de entrega	2 días
Actividades: ○ El Experto y el Cliente elaboran el diagrama de contexto de las demandas (Figura 13) tomando en cuenta la problemática, necesidades y requerimientos iniciales. ○ El Cliente agrega una descripción de la visión del proyecto usando la siguiente plantilla (figura 13): “El/La [nombre del Servicio de TI mínimo viable (MVITS)] será usado por [tipos de usuarios] para [funciones principales] para tener [beneficios] con métricas de servicio previstas de (funcionalidad, nivel); (capacidad, nivel); (disponibilidad, nivel); (continuidad, nivel) y (seguridad, nivel)”. ○ El Experto y el Cliente elaboran en conjunto el diagrama de caso de uso (Figura 14), el cual ayuda a entender las demandas de las necesidades y así poder desarrollar historias de usuario específicas al proyecto. ○ El Cliente elabora un conjunto de historias de usuario usando la siguiente plantilla: “Como [tipo de usuario] puedo [acción] para [beneficios]”, además de asignar a cada historia un nivel de importancia de implementación: muy alto (VH), alto (H) o normal (N) (tabla 7). ○ El Experto activa un repositorio digital para el almacenamiento de todo lo relacionado al proyecto.	

Tabla 6: Descripción de las actividades involucradas en la tarea T.1.1 respecto a la metodología AVS_FD_MVITS.v1-2022.

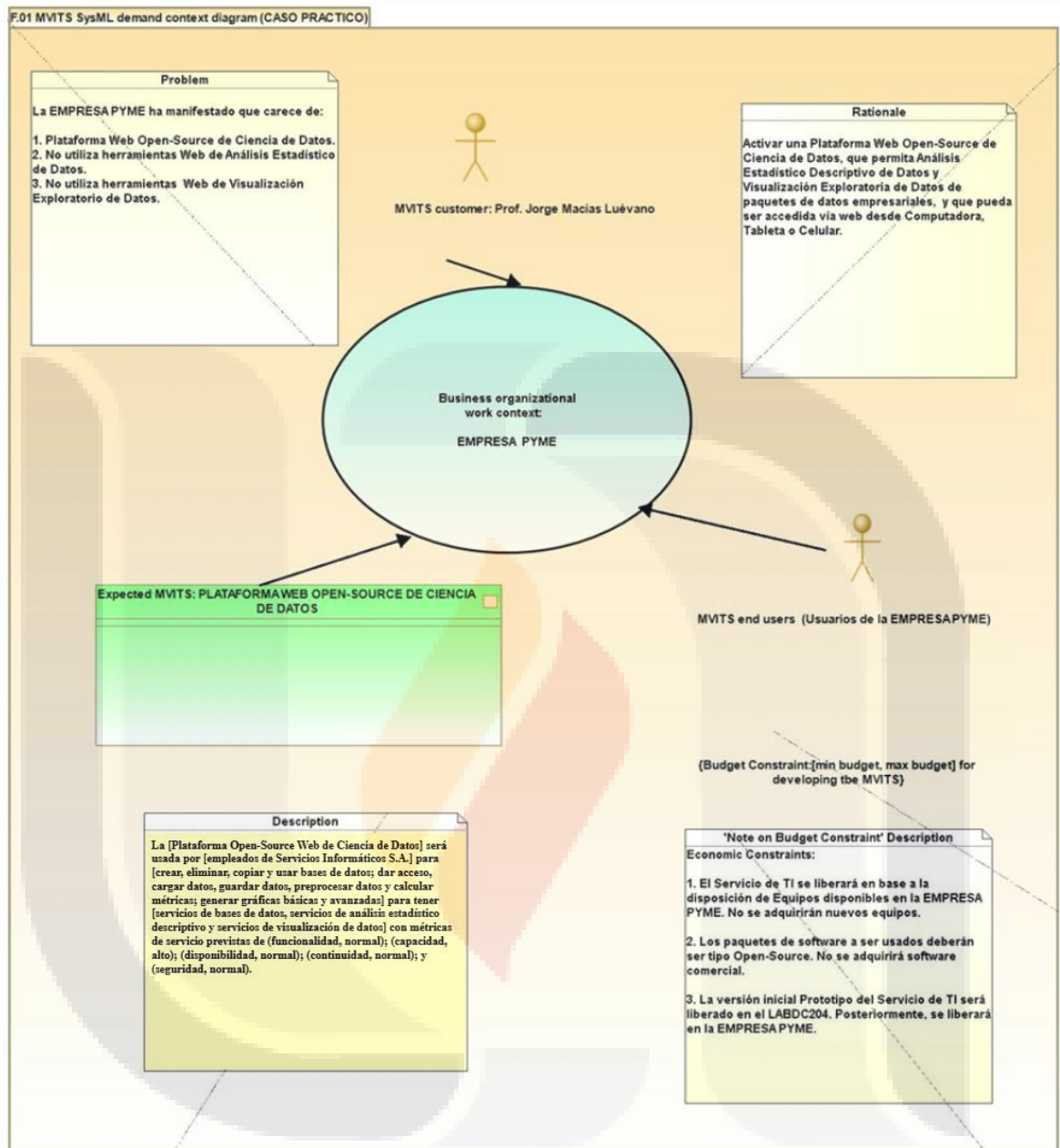


Figura 13: Diagrama de contexto de las demandas del MVITS acorde a la metodología AVS_FD_MVITS.v1-2022.

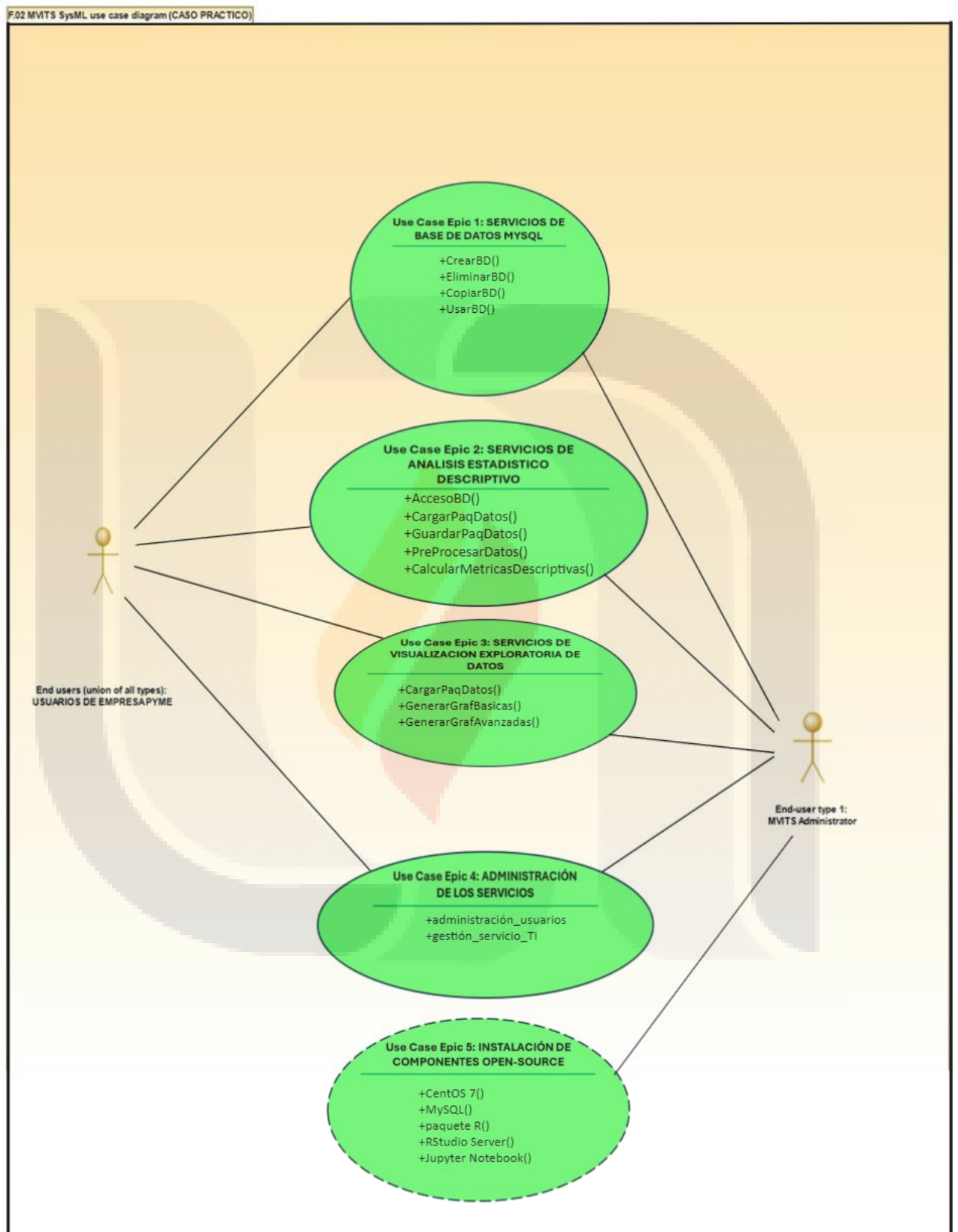


Figura 14: Diagrama de casos de uso, tomada de la metodología AVS_FD_MVITS.v1-2022.

Visión MVITS:	La [Plataforma Open-Source Web de Ciencia de Datos] será usada por [empleados de Sitis De Aguascalientes S.C.] para [crear, eliminar, copiar y usar bases de datos; dar acceso, cargar datos, guardar datos, preprocesar datos y calcular métricas; generar gráficas básicas y avanzadas] para tener [servicios de bases de datos, servicios de análisis estadístico descriptivo y servicios de visualización de datos] con métricas de servicio previstas de (funcionalidad, normal); (capacidad, alto); (disponibilidad, normal); (continuidad, normal); y (seguridad, normal).
---------------	---

# HISTORIA DEL USUARIO	ID EPIC (DIAGRAMA CASO DE USO)	HISTORIA DEL USUARIO	VALOR DE APLICACIÓN PARA EL CLIENTE	ID HISTORIA DEL USUARIO	PUNTOS DE HISTORIA	PLAN BACKLOG				
			1.VH, 2.H, OR 3.N		1, 2, 3 OR 5	CICLO 1	CICLO 2	CICLO 3	CICLO 4	CICLO 5
1	USO.CASO.1: Servicios de Base de Datos	Como [Administrador o Usuario] puedo [crear, eliminar y copiar bases de datos] para [mi análisis personal de datos].	1.VH	1.VH.STO.01.haseDatos						
2	USO.CASO.1: Servicios de Base de Datos	Como [Administrador o Usuario] puedo [usar (administrar tablas, consultar tablas, imprimir tablas, exportar, importar) bases de datos] para [mi análisis personal de datos].	2.H	2.H.STO.02.tablasDatos						
3	USO.CASO.2: Servicios de Análisis Estadístico Descriptivo	Como [Administrador o Usuario] puedo [acceder, cargar y guardar paquetes de datos] para [crear, administrar y monitorear variables de interés].	3.N	3.N.STO.03.paqueteDatos						
4	USO.CASO.2: Servicios de Análisis Estadístico Descriptivo	Como [Administrador o Usuario] puedo [preprocesar datos] para [garantizar que los datos estén listos para su análisis y facilite la extracción de información valiosa].	1.VH	1.VH.STO.04.preprocesar						
5	USO.CASO.2: Servicios de Análisis Estadístico Descriptivo	Como [Administrador o Usuario] puedo [calcular métricas descriptivas] para [proveer una interpretación personal de la información].	1.VH	1.VH.STO.05.metricasDes						
6	USO.CASO.3: Servicios de Visualización Exploratoria de Datos	Como [Administrador o Usuario] puedo [generar gráficas básicas] para [realizar una exploración inicial de manera visual].	2.H	2.H.STO.06.graficasBasicas						
7	USO.CASO.3: Servicios de Visualización Exploratorio de Datos	Como [Administrador o Usuario] puedo [generar gráficas avanzadas] para [la detección de patrones o diferencias, además de analizar su distribución].	2.H	2.H.STO.07.graficasAvanzadas						
8	USO.CASO.4: Administración de los Servicios	Como [Administrador o Usuario] puedo [gestionar la creación y administración de usuarios] para [llevar un control de usuarios que tienen acceso a los servicios y supervisar avances].	1.VH	1.VH.STO.08.usuarios						
9	USO.CASO.4: Administración de los Servicios	Como [Administrador o Usuario] puedo [gestionar los servicios de TI] para [controlar y mitigar fallas en los servicios].	1.VH	1.VH.STO.09.gestionServicio						
10	USO.CASO.5: Instalación de componentes Open-Source	Como [Administrador] debo [instalar las versiones adecuadas de los componentes Open-Source] para [garantizar el funcionamiento adecuado de los servicios].	1.VH	1.VH.STO.10.instalación						
TOTAL DE PUNTOS DE HISTORIA					0	0	0	0	0	0

Tabla 7: Hoja de backlog (inicial), tomada de la metodología AVS_FD_MVITS.v1-2022.

4.1.2 Tarea: T.1.2 Elaboración de boceto de arquitectura

T.1.2 Elaboración de un boceto de arquitectura	
Prácticas ágiles	Sesión rápida de diseño Diseño simple
Tiempo de entrega	2 días
Actividades: ○ El equipo de desarrollo elabora un diseño simple sobre la arquitectura que tendrá el servicio, se toma en cuenta la visión e historias de usuario proporcionadas por el Cliente. El esquema de la arquitectura debe incluir lo siguiente: ○ Diagrama de bloques sobre los componentes TI (figura 15): ▪ Software (aplicaciones) y Datos --- capa: APPSW-DATA. ▪ Software (sistema operativo) y Bases de datos --- capa: SYSSW-DBMS. ▪ Hardware computacional --- capa: HW. ▪ Red --- capa: NW. ▪ Infraestructura del entorno y Servicios informáticos externos --- capa: ENV-EXT. ○ Matriz de asignación de funcional (tabla 9) respecto a las historias de usuario (columnas) contra la lista de componentes TI (filas), marcando en ella con [✓] cada celda [componente TI – historia de usuario] donde el componente TI contribuya a lograr y/o desempeñar la historia de usuario. ○ Matriz de asignación no-funcional (tabla 10) respecto a los niveles métricos esperados (muy alto, alto y normal) de los parámetros a evaluar: funcionalidad, capacidad, disponibilidad, continuidad y seguridad (columnas), contra la lista de componentes TI (filas), evaluando en cada celda el nivel de contribución (contribución normal: [○] contribución crítica: [●]) de dicho elemento. ○ Pruebas de viabilidad (si las hay): en este documento se reportan las evidencias técnicas para constatar que es posible implementar las historias de usuario (tabla 11).	

Tabla 8: Descripción de prácticas, tiempos de entrega y actividades a realizar en la tarea T.1.2 acorde a la metodología AVS_FD_MVITS.v1-2022.

A partir de esta sección se usarán datos, componentes, infraestructura y evidencias dentro de un entorno estudiantil por temas de confidencialidad. Se toman estas medidas para poder garantizar la protección de los datos empresariales del Cliente del MVITS, no obstante, los procesos de la metodología son los mismos, simplemente cambian algunos componentes TI (versiones finales), un ejemplo de esto sería el método con el que se alberga la plataforma (SYSSW-DBMS): dentro de un servidor, aprovechar un

equipo de cómputo, o recurrir a la virtualización por medio de una máquina virtual (VM).

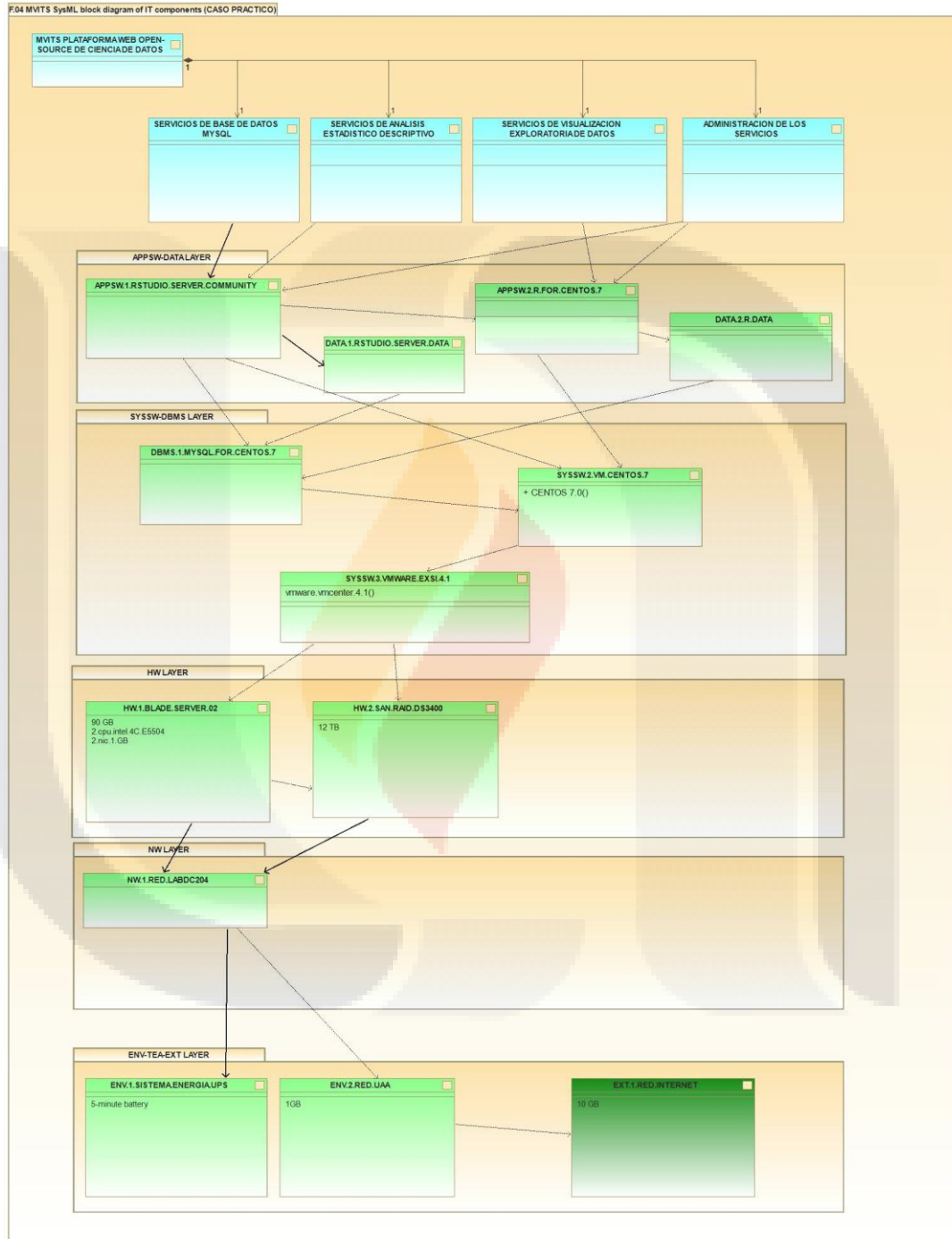


Figura 15: Diagrama de bloques de los componentes TI implementados en el MVITS (entorno estudiantil) acorde a la metodología AVS_FD_MVITS.v1-2022.

[illegible]

Tabla 9: Matriz de asignación funcional, se toma como referencia las historias de usuario y los componentes TI de cada capa del diagrama de bloques (tomada de la metodología AVS_FD_MVITS.v1-2022).

Tabla 10: Matriz de asignación no-funcional acorde a las métricas de evaluación de servicio y componentes TI (tomada de la metodología AVS_FD_MVITS.v1-2022).

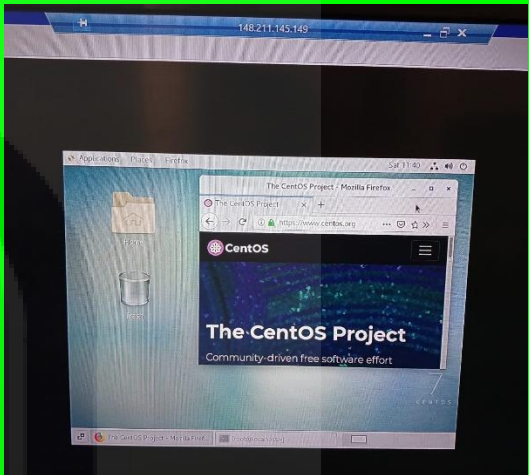
PROBLEMA DE VIABILIDAD TÉCNICA DESCONOCIDO	# HISTORIA DEL USUARIO	ID HISTORIA USUARIO	HISTORIA DE USUARIO RELACIONADA	EVIDENCIAS
ISSUE.1 ¿Se puede instalar y configurar el VM CENTOS 7 en el LABDC204?	10	1.VH.STO.10.instalación	Como [Administrador] debo [instalar las versiones adecuadas de los componentes Open-Source] para [garantizar el funcionamiento adecuado de los servicios].	Si, evidencia: 

Tabla 11: Evidencias sobre la viabilidad del MVITS, es posible el desarrollo del proyecto al confirmar la historia de usuario 10 (tomada de la metodología AVS_FD_MVITS.v1-2022).

4.1.3 Tarea: T.1.3 Negociación del backlog

T.1.3 Negociación del backlog	
Prácticas ágiles	Planificación Póker. Velocidad de trabajo del equipo de desarrollo.
Tiempo de entrega	1 día
Actividades: ○ El equipo de desarrollo estima los puntos de historia para cada historia de usuario usando la planificación póker (tabla 14). Las puntuaciones válidas solo pueden ser [1, 2, 3 o 5] puntos de historia . Un punto de historia corresponde a 8 horas de trabajo. ○ El equipo de desarrollo analiza y reordena (de ser necesario) las historias de usuario acorde a la dificultad de implementación y considerando el valor de aplicación para el cliente (tabla 14). ○ El equipo de desarrollo y el Cliente negocian el número de ciclos y la cantidad de historias de usuario por ciclo, basándose en la velocidad de trabajo calculada por ciclo (tabla 13).	

Tabla 12: Descripción de prácticas, tiempo de entrega y actividades involucradas en la tarea T.1.3 acorde a la metodología AVS_FD_MVITS.v1-2022.

		Datos de entrada	
		Datos calculados	
S-semanas/ciclo es el número de semanas acordadas por ciclo (2, 3 o 4).	S-semanas/ciclo	2	Semanas
N es el número de trabajadores por ciclo.	N	1	trabajadores
X-hr/día es el número de horas de trabajo reales asignadas al proyecto MVITS por trabajador.	X-hr/día	5	horas
RW es el total de horas de trabajo reservadas por ciclo	$RW = N * X\text{-hr/día} * 5\text{-días/semana} * S\text{-semanas/ciclo}$	50	horas
TW es el total de horas de trabajo por ciclo	$TW = N * 8\text{-hr/día} * 5\text{-días/semana} * S\text{-semanas/ciclo}$	80	horas
WF es el factor de carga de trabajo en porcentaje por ciclo	$WF = RW / TW$	62.5%	%
CA es la capacidad disponible en horas de trabajo por ciclo	$CA = WF * TW$ (hora-trabajada)	50	horas
CV es la velocidad calculada en puntos de historia por ciclo	$CV = CA / 8\text{-hrs-puntos de historia}$	6	puntos de historia
TSP es el total de puntos de historia estimados para el MVITS	TSP	14	puntos de historia
S es el número de ciclos necesarios	$S = TSP / CV$	2	Ciclos
DP es la duración del proyecto MVITS en semanas	$DP = S * S\text{-semanas/ciclo}$	4	semanas

Tabla 13: Proceso del cálculo de la velocidad de trabajo, los datos en color verde representan los datos acordados por el cliente y el equipo de desarrollo, los datos en azul se van calculando por medio de las fórmulas descritas en la metodología AVS_FD_MVITS.v1-2022.

El equipo de desarrollo asignó el valor de cada historia de usuario tomando en consideración su dificultad de aplicación y/o viabilidad técnica (tabla 14), obteniendo un total de 14 puntos de historia para el desarrollo del MVITS. Además, se acordó con el Cliente que cada ciclo se conforma por 2 semanas y la jornada de trabajo fuera de 5 horas por día (horas reales de trabajo), ya que solo se contaba con 1 trabajador (estudiante).

Visión MVITS: La [Plataforma Open-Source Web de Ciencia de Datos] será usada por [empleados de Sitis De Aguascalientes S.C.] para [crear, eliminar, copiar y usar bases de datos; dar acceso, cargar datos, guardar datos, preprocesar datos y calcular métricas; generar gráficas básicas y avanzadas] para tener [servicios de bases de datos, servicios de análisis estadístico descriptivo y servicios de visualización de datos] con métricas de servicio previstas de (funcionalidad, normal); (capacidad, alto); (disponibilidad, normal); (continuidad, normal); y (seguridad, normal).										
# HISTORIAS DE USUARIO	ID EPIC(DIAGRAMA CASO DE USO)	HISTORIA DE USUARIO	VALOR DE APLICACIÓN PARA EL CLIENTE	ID HISTORIA DE USUARIO	PUNTOS DE HISTORIA	PLAN BACKLOG				
			1.VH, 2.H, OR 3.N			CICLO 1	CICLO 2	CICLO 3	CICLO 4	CICLO 5
10	USO.CASO.5: Instalación de componentes Open-Source	Como [Administrador] debo [instalar las versiones adecuadas de los componentes Open-Source] para [garantizar el funcionamiento adecuado de los servicios].	1.VH	1.VH.STO.10.instalación	3	3				
1	USO.CASO.1: Servicios de Base de Datos	Como [Administrador o Usuario] puedo [crear, eliminar y copiar bases de datos] para [mi análisis personal de datos].	1.VH	1.VH.STO.01.baseDatos	1	4				
4	USO.CASO.2: Servicios de Análisis Estadístico Descriptivo	Como [Administrador o Usuario] puedo [preprocesar datos] para [garantizar que los datos estén listos para su análisis y facilite la extracción de información valiosa].	1.VH	1.VH.STO.04.preprocesar	1	5				
5	USO.CASO.2: Servicios de Análisis Estadístico Descriptivo	Como [Administrador o Usuario] puedo [calcular métricas descriptivas] para [proveer una interpretación personal de la información].	1.VH	1.VH.STO.05.metricasDes	1	6				
8	USO.CASO.4: Administración de los Servicios	Como [Administrador o Usuario] puedo [gestionar la creación y administración de usuarios] para [llevar un control de usuarios que tienen acceso a los servicios y supervisar avances].	1.VH	1.VH.STO.08.usuarios	2	8				
9	USO.CASO.4: Administración de los Servicios	Como [Administrador o Usuario] puedo [gestionar los servicios de TI] para [controlar y mitigar fallas en lo servicios].	1.VH	1.VH.STO.09.gestionServicio	2		2			
2	USO.CASO.1: Servicios de Base de Datos	Como [Administrador o Usuario] puedo [usar (administrar tablas, consultar tablas, imprimir tablas, exportar, importar) bases de datos] para [mi análisis personal de datos].	2.H	2.H.STO.02.tablasDatos	1		3			
6	USO.CASO.3: Servicios de Visualización Exploratoria de Datos	Como [Administrador o Usuario] puedo [generar gráficas básicas] para [realizar una exploración inicial de manera visual].	2.H	2.H.STO.06.graficasBasicas	1		4			
7	USO.CASO.3: Servicios de Visualización Exploratorio de Datos	Como [Administrador o Usuario] puedo [generar gráficas avanzadas] para [la detección de patrones o diferencias, además de analizar su distribución].	2.H	2.H.STO.07.graficasAvanzadas	1		5			
3	USO.CASO.2: Servicios de Análisis Estadístico Descriptivo	Como [Administrador o Usuario] puedo [acceder, cargar y guardar paquetes de datos] para [crear, administrar y monitorear variables de interés].	3.N	3.N.STO.03.paqueteDatos	1		6			
TOTAL DE PUNTOS DE HISTORIA					14	8	6	0	0	0

Tabla 14: Hoja de backlog (reordenada) añadiendo los puntos de historia respectivos a cada historia de usuario, además de señalar las historias de usuario a realizar en cada ciclo.

El total de puntos de historia por ciclo debería de ser 6 acorde a lo calculado en la tabla 13, pero al obtener una carga de trabajo del 62.5% en adición de acordar 14 puntos de historia en total, se estipuló aumentar a 8 en el ciclo 1 para poder respetar el número de ciclos determinado por el Equipo de desarrollo y el Cliente.

4.2 Etapa 2. Ciclos MVITS. Diseño, transición y construcción

Esta segunda etapa se compone de seis tareas: T.2.1 Planificación de los ciclos, T.2.2 Reuniones rápidas, T.2.3 Aceptación del cliente y elaboración de pruebas técnicas, T.2.4 Incremento de la ingeniería y T.2.5 Revisión de los ciclos (figura 16), se inicia tomando el PKG.01 desarrollado en la etapa 1 para poder construir el paquete de ciclos (PKG.02), el cual se compone por los siguientes elementos: hoja de trabajo por ciclo, hoja de aceptación y pruebas técnicas, además de una hoja de revisión del ciclo. Cabe mencionar que la etapa tiene un tiempo límite de entrega de dos semanas y los roles involucrados son los siguientes: Cliente (empresa Sitis De Aguascalientes S.C., Mtro. Jorge Eduardo Macías Luevano), Experto (Dr. José Manuel Mora Tavarez), Equipo de desarrollo (Ing. Juan Daniel Gallegos Rodríguez) y Personal externo (Dra. Yuritzy Reyes Delgado).

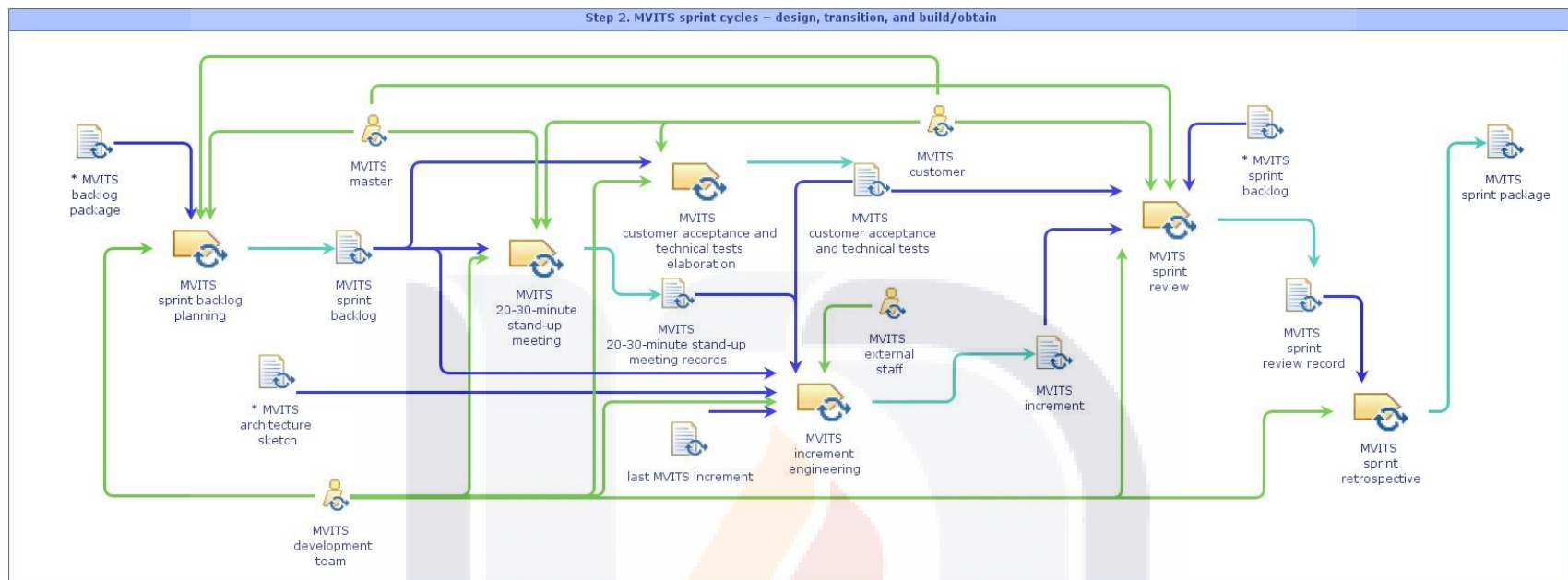


Figura 16: Etapa de Diseño, Transición y Construcción (tomada de “Metodología Ágil para el Desarrollo de un Nuevo Servicio de TI Mínimo Viable (AVS_FD_MVITS.v1-2022)”).

4.2.1 T.2.1 Planificación de los ciclos

T.2.1 Planificación de los ciclos	
Prácticas ágiles	Definición de lo “Hecho”. Matriz de mapeo de historias. Velocidad de trabajo del equipo de desarrollo.
Tiempo de entrega	2 días
Actividades: ○ El Cliente define la meta/objetivo del ciclo usando la plantilla (tabla 16 y 17): “Meta del ciclo [# número de ciclo] es implementar [funcionalidades] para obtener [beneficios] con métricas de servicio previstas de [(funcionalidad, nivel); (capacidad, nivel); (disponibilidad, nivel); (continuidad, nivel); y (seguridad, nivel)]”. ○ El Cliente proporciona la declaración de “Hecho” para cada uno de los ciclos con la siguiente plantilla (tabla 16 y 17): “Hecho [# ciclo] se logra cuando [(SI/NO) todas las historias de usuario del ciclo son aceptadas; (SI/NO) el incremento de la ingeniería del ciclo fue entregado; (SI/NO) el repositorio del proyecto MVITS fue actualizado]. ○ El Equipo de desarrollo elabora la hoja de trabajo por ciclo utilizando la práctica ágil de matriz de mapeo de historias (tabla 16 y 17). Dicha matriz se conforma por las historias de usuario seleccionadas por ciclo (columnas) y las cinco capas de componentes TI (filas). En cada una de las celdas se debe describir las tareas a realizar y estipular la cantidad de horas de trabajo asignadas.	

Tabla 15: Descripción de prácticas ágiles, tiempos de entrega y actividades involucradas en la tarea T.2.1 acorde a la metodología AVS_FD_MVITS.v1-2022.

META DEL CICLO [ciclo #1]		Meta del ciclo [1] es implementar [instalar componente TI; crear, eliminar y copiar bases de datos; procesar datos; calcular métricas descriptivas; y administrar usuarios] para obtener [entorno de trabajo optimo para albergar el servicio] con métricas de servicio previstas de [(funcionalidad, normal); (capacidad, alto); (disponibilidad, normal); (continuidad, normal); y (seguridad, normal)].					
DEFINICIÓN DE LO "HECHO" [ciclo # 1]		Hecho [ciclo 1] se logra cuando [(SI) todas las historias de usuario del ciclo son aceptadas; (SI) el incremento de la ingeniería del ciclo fue entregado; (SI) el repositorio del proyecto MVITS fue actualizado].					
		HOJA DE TRABAJO DEL CICLO 1 (TAREAS Y HORAS)					
		1.VH.STO.10.instalación	1.VH.STO.01.baseDatos	1.VH.STO.04.preprocesar	1.VH.STO.05.métricasDes	1.VH.STO.08.usuarios	TOTAL
PUNTOS DE HISTORIA		3	1	1	1	2	8
TOTAL DE HORAS DISPONIBLES		24	8	8	8	16	64
TOTAL DE HORAS REALES DE TRABAJO		15	5	5	5	10	40
CAPA APPSW-DATA	TAREAS	Instalar paquete R y Rstudio Server. Realizar pruebas de funcionamiento y compatibilidad en la terminal del sistema y servidor.	...	Importar bases de datos en Rstudio y realizar pruebas de preprocesamiento.	Realizar scrips en el entorno de Rstudio fáciles de modificar y confirmar su correcto funcionamiento.	Realizar pruebas en el servidor de Rstudio con diferentes usuarios para confirmar cambios en sus entornos de trabajo y administrar permisos.	
	HORAS	6		2	2	3	13
CAPA SYSSW-DBMS	TAREAS	Instalar CentOS Linux 7 distribución Red Hat 64 bits VM-labdc-uaa.	Instalar versión de escritorio (workbench) de MySQL 8.0.22 64 bits Community.	Realizar pruebas de depuración de datos en el workbench para eliminar información inútil o registros incompletos.	Realizar pruebas de importación y exportación de bases de datos para mantenerlas actualizadas.	Realizar pruebas de creación, modificación y eliminación de usuarios dentro de la máquina virtual.	
	HORAS	4	2	1	2	4	13
CAPA HW	TAREAS	Crear un máquina virtual para el servidor labdc-uaa [vmware 4.1].	...	...	...	Asegurarse que el servidor labdc-uaa actualice los cambios de usuarios.	
	HORAS	2				2	4
CAPA NW	TAREAS	Activar y aprobar la autorización del firewall para acceso remoto al servidor labdc- uaa, además del puerto 8787 para acceder al servicio de Rstudio server	Activar servidor MySQL para poder usarlo en workbench y establecer conexión con Rstudio.	Establecer conexión entre Rstudio server y MySQL workbench para importar bases de datos para su análisis.	...	...	
	HORAS	2	2	1			5
CAPA ENV-EXT	TAREAS	Realizar prueba de conexión remota para acceder al servidor labdc-uaa por medio de un protocolo de escritorio remoto (RDP).	Realizar prueba de conexión remota para acceder al servidor labdc-uaa por medio de un protocolo de escritorio remoto (RDP).	Realizar prueba de conexión remota para acceder al servidor labdc-uaa por medio de un protocolo de escritorio remoto (RDP).	Realizar prueba de conexión remota para acceder al servidor labdc-uaa por medio de un protocolo de escritorio remoto (RDP).	Realizar prueba de conexión remota para acceder al servidor labdc-uaa por medio de un protocolo de escritorio remoto (RDP).	
	HORAS	1	1	1	1	1	5
TOTAL DE HORAS ASIGNADAS		15	5	5	5	10	40

Tabla 16: Hoja de trabajo del ciclo #1.

META DEL CICLO [ciclo #2]		Meta del ciclo [2] es implementar [gestionar los servicios; administrar, consultar, imprimir, exportar e importar tablas; generar gráficas básicas; generar gráficas avanzadas; y administrar paquetes de datos] para obtener [máximo aprovechamiento de datos empresariales almacenados a través de los años] con métricas de servicio previstas de [(funcionalidad, normal); (capacidad, alto); (disponibilidad, normal); (continuidad, normal); y (seguridad, normal)].					
DEFINICIÓN DE LO "HECHO" [ciclo # 2]		Hecho [ciclo 2] se logra cuando [(SI) todas las historias de usuario del ciclo son aceptadas; (SI) el incremento de la ingeniería del ciclo fue entregado; (SI) el repositorio del proyecto MVITS fue actualizado].					
		HOJA DE TRABAJO DEL CICLO 2 (TAREAS Y HORAS)					
		1.VH.STO.09.gestionServicio	2.H.STO.02.tablasDatos	2.H.STO.06.graficasBasicas	2.H.STO.07.graficasAvanzadas	3.N.STO.03.paqueteDatos	TOTAL
PUNTOS DE HISTORIA		2	1	1	1	1	6
TOTAL DE HORAS DISPONIBLES		16	8	8	8	8	48
TOTAL DE HORAS REALES DE TRABAJO		10	5	5	5	5	30
CAPA APPSW-DATA	TAREAS	Mantener actualizadas las versiones de Rstudio y paquete R si lo permite la infraestructura.	Establecer conexión con el gestor de bases de datos MySQL.	Realizar pruebas con los demos de gráficas para comprobar el funcionamiento de paquetes.	Instalar paquetes/bibliotecas que permitan elaborar gráficos más avanzados/complexos.	Cargar scrips que interactuen con datos importados.	
	HORAS	2	1	4	4	1	12
CAPA SYSSW-DBMS	TAREAS	Revisar el estatus de los servidores de Rstudio y MySQL.	Mantener actualizadas las tablas que conforman las bases de datos con las que se esten trabajando.	...	...	Realizar pruebas dentro del entorno de trabajo MySQL cargando datos para su análisis	
	HORAS	1	1			3	5
CAPA HW	TAREAS	Dar mantenimiento al servidor labdc-uaa.	Dar mantenimiento a los componentes de almacenamiento y remplazarlos de ser necesario.	...	...	...	
	HORAS	5	2				7
CAPA NW	TAREAS	Monitorear conexión de red para preveer posibles caídas.	...	...	...	...	
	HORAS	1					1
CAPA ENV-EXT	TAREAS	Realizar prueba de conexión remota para acceder al servidor labdc-uaa por medio de un protocolo de esquitorio remoto (RDP).	Realizar prueba de conexión remota para acceder al servidor labdc-uaa por medio de un protocolo de esquitorio remoto (RDP).	Realizar prueba de conexión remota para acceder al servidor labdc-uaa por medio de un protocolo de esquitorio remoto (RDP).	Realizar prueba de conexión remota para acceder al servidor labdc-uaa por medio de un protocolo de esquitorio remoto (RDP).	Realizar prueba de conexión remota para acceder al servidor labdc-uaa por medio de un protocolo de esquitorio remoto (RDP).	
	HORAS	1	1	1	1	1	5
TOTAL DE HORAS ASIGNADAS		10	5	5	5	5	30

Tabla 17: Hoja de trabajo del ciclo #2.

4.2.2 T.2.2 Reuniones rápidas

T.2.1 Reuniones rápidas	
Prácticas ágiles	Reuniones diarias. Tres preguntas.
Tiempo de entrega	20-30 minutos
Actividades: ○ El Experto organiza reuniones de 20-30 minutos donde el Equipo de desarrollo contesta las siguientes preguntas al Cliente: 1) ¿Qué tareas se hicieron ayer? 2) ¿Qué tareas se harán hoy? 3) ¿Qué impedimentos hay el día de hoy?	

Tabla 18: Descripción de prácticas ágiles, tiempos de entrega y actividades involucradas en la tarea T.2.2 acorde a la metodología AVS_FD_MVITS.v1-2022.

Por temas de confidencialidad no se adjuntan todos los reportes de reuniones con el Cliente, ya que se tratan temas de infraestructura, protocolos y datos delicados para los procesos que rigen la compañía. Sin embargo, en la tabla 19 se muestra un ejemplo del reporte que se guarda dentro del repositorio digital.

		CICLO		1
		ID DE LA REUNION		1.1
		¿Qué tareas se hicieron ayer?	¿Qué tareas se harán hoy?	¿Qué impedimentos hay el día de hoy?
CAPA APPSW-DATA	TAREAS	...	Se instalará el paquete de R en la máquina virtual.	En contrar versión compatible con CentOS 7.
CAPA SYSSW-DBMS	TAREAS	Se instaló CentOS Linux 7 distribución Red Hat 64 bits dentro de una máquina virtual en el VM-labdc-uaa.	...	...
CAPA HW	TAREAS	...	...	...
CAPA NW	TAREAS	...	...	...
CAPA ENV-EXT	TAREAS	...	...	...

Tabla 19: Ejemplo de una hoja de reuniones rápidas acorde a la metodología AVS_FD_MVITS.v1-2022.

4.2.3 T.2.3 Aceptación del cliente y elaboración de pruebas técnicas

T.2.3 Aceptación del cliente y elaboración de pruebas técnicas	
Prácticas ágiles	Prueba de aceptación. Pruebas unitarias.
Tiempo de entrega	1 día
Actividades: ○ El Cliente estipula una prueba de aceptación para cada historia de usuario de la hoja de trabajo del ciclo actual implementando la siguiente plantilla (tabla 21 y 22): “[ID de la prueba de aceptación]: como [tipo de usuario] acepto la historia de usuario [ID de la historia de usuario] cuando pueda [conjunto de funcionalidades]”. ○ El Equipo de desarrollo desarrolla para cada prueba de aceptación dos o tres pruebas técnicas utilizando la siguiente plantilla (tabla 21 y 22): “[ID prueba técnica]: como [rol técnico] cuando logro [acción] entonces el servicio MVITS [respuesta]”. Los tipos de roles técnicos están relacionados con las cinco capas de componentes TI, no obstante, en este caso al solo contar con un integrante de desarrollo se usará [rol 1]. ○ El Experto actualiza el repositorio digital para incluir las pruebas técnicas.	

Tabla 20: Descripción de prácticas ágiles, tiempos de entrega y actividades involucradas en la tarea T.2.3 acorde a la metodología AVS_FD_MVITS.v1-2022.

CICLO						1
# HISTORIAS DE USUARIO	ID EPIC(DIAGRAMA CASO DE USO	HISTORIA DE USUARIO	ID HISTORIA DE USUARIO	PUNTOS DE HISTORIA	PRUEBAS DE ACEPTACIÓN	PRUEBAS TÉCNICAS
				1, 2, 3 OR 5		
10	USO.CASO.5: Instalación de componentes Open-Source	Como [Administrador] debo [instalar las versiones adecuadas de los componentes Open-Source] para [garantizar el funcionamiento adecuado de los servicios].	1.VH.STO.10.instalación	3	PA.1.VH.STO.10: como [administrador] acepto la historia de usuario [1.VH.STO.10.instalación] cuando pueda [acceder a las aplicaciones de código abierto].	PT.1.1.VH.STO.10: como [rol técnico 1] cuando logro [instalar el paquete R y Rstudio server] entoces el servicio MVITS [me da acceso a los comandos del lenguaje R y puedo crear scrips]. PT.2.1.VH.STO.10: como [rol técnico 1] cuando logro [activar la conexión remota] entoces el servicio MVITS [puedo acceder a los servicios MVITS desde cualquier dispositivo].
1	USO.CASO.1: Servicios de Base de Datos	Como [Administrador o Usuario] puedo [crear, eliminar y copiar bases de datos] para [mi análisis personal de datos].	1.VH.STO.01.baseDato0s	1	PA.1.VH.STO.01: como [administrador o usuario] acepto la historia de usuario [1.VH.STO.01.baseDato0s] cuando pueda [administrar mis bases de datos desde el servicio MVITS].	PT.1.1.VH.STO.01: como [rol técnico 1] cuando logro [instalar el workbench de MySQL] entoces el servicio MVITS [me da acceso a herramientas de administración de bases de datos]. PT.2.1.VH.STO.01: como [rol técnico 1] cuando logro [activar el servidor MySQL] entoces el servicio MVITS [dar permisos a usuarios para acceder al servidor de bases de datos].
4	USO.CASO.2: Servicios de Análisis Estadístico Descriptivo	Como [Administrador o Usuario] puedo [preprocesar datos] para [garantizar que los datos estén listos para su análisis y facilite la extracción de información valiosa].	1.VH.STO.04.preprocesar	1	PA.1.VH.STO.04: como [administrador o usuario] acepto la historia de usuario [1.VH.STO.04.preprocesar] cuando pueda [obtener bases de datos limpias].	PT.1.1.VH.STO.04: como [rol técnico 1] cuando logro [importar bases de datos en Rstudio server] entoces el servicio MVITS [me ofrece comandos para extraer información irrelevante y eliminar registros incompletos]. PT.2.1.VH.STO.04: como [rol técnico 1] cuando logro [exportar datos de Rstudio server] entoces el servicio MVITS [mantiene las bases de datos actualizadas].
5	USO.CASO.2: Servicios de Análisis Estadístico Descriptivo	Como [Administrador o Usuario] puedo [calcular métricas descriptivas] para [proveer una interpretación personal de la información].	1.VH.STO.05.metricasDes	1	PA.1.VH.STO.05: como [administrador o usuario] acepto la historia de usuario [1.VH.STO.05.metricasDes] cuando pueda [obtener métricas descriptivas de mis datos empresariales].	PT.1.1.VH.STO.05: como [rol técnico 1] cuando logro [descargar paquetes para análisis descriptivo] entoces el servicio MVITS [me ofrece comandos para extraer información irrelevante y eliminar registros incompletos]. PT.2.1.VH.STO.05: como [rol técnico 1] cuando logro [crear un script con comandos] entoces el servicio MVITS [puede guardar el archivo para poder implementarlo después].
8	USO.CASO.4: Administración de los Servicios	Como [Administrador o Usuario] puedo [gestionar la creación y administración de usuarios] para [llevar un control de usuarios que tienen acceso a los servicios y supervisar avances].	1.VH.STO.08.usuarios	2	PA.1.VH.STO.08: como [administrador o usuario] acepto la historia de usuario [1.VH.STO.08.usuarios] cuando pueda [obtener un registro de usuarios y de los permisos que disponen].	PT.1.1.VH.STO.08: como [rol técnico 1] cuando logro [instalar el sistema operativo CentOS] entoces el servicio MVITS [permite la creación, modificación y eliminación de usuarios]. PT.2.1.VH.STO.05: como [rol técnico 1] cuando logro [crear usuarios] entoces el servicio MVITS [me permite administrar credenciales de acceso y permisos de cada usuario].
		TOTAL DE PUNTOS DE HISTORIA		8		

Tabla 21: Hoja de pruebas de aceptación y pruebas técnicas de cada historia de usuario que conforma el ciclo #1.

CICLO					2	
# HISTORIAS DE USUARIO	ID EPIC(DIAGRAMA CASO DE USO	HISTORIA DE USUARIO	ID HISTORIA DE USUARIO	PUNTOS DE HISTORIA 1, 2, 3 OR 5	PRUEBAS DE ACEPTACIÓN	PRUEBAS TÉCNICAS
9	USO.CASO.4: Administración de los Servicios	Como [Administrador o Usuario] puedo [gestionar los servicios de TI] para [controlar y mitigar fallas en lo servicios].	1.VH.STO.09.gestionServicio	2	PA.1.VH.STO.09: como [administrador o usuario] acepto la historia de usuario [1.VH.STO.09.gestionServicio] cuando pueda [monitorear los servicios de la plataforma].	PT.1.1.VH.STO.09: como [rol técnico 1] cuando logro [acceder a los servidores de MySQL y Rstudio] entoces el servicio MVITS [me permite gestionar los sevicios de bases de datos y y análisis de datos]. PT.2.1.VH.STO.09: como [rol técnico 1] cuando logro [acceder al servidor labdc-uaa] entonces el servicio MVITS [puedo monitorear de forma remota la actividad de cada usuario].
2	USO.CASO.1: Servicios de Base de Datos	Como [Administrador o Usuario] puedo [usar [administrar tablas, consultar tablas, imprimir tablas, exportar, importar] bases de datos] para [mi análisis personal de datos].	2.H.STO.02.tablasDatos	1	PA.1.H.STO.02: como [administrador o usuario] acepto la historia de usuario [2.H.STO.02.tablasDatos] cuando pueda [administrar tablas de datos para su análisis o interpretación].	PT.1.1.H.STO.02: como [rol técnico 1] cuando logro [habilitar conexión entre Rstudio y MySQL] entoces el servicio MVITS [me permite importar y exportar tablas para mantener la información actualizada].
6	USO.CASO.3: Servicios de Visualización Exploratoria de Datos	Como [Administrador o Usuario] puedo [generar gráficas básicas] para [realizar una exploración inicial de manera visual].	2.H.STO.06.graficasBasicas	1	PA.1.H.STO.06: como [administrador o usuario] acepto la historia de usuario [2.H.STO.06.graficasBasicas] cuando pueda [guardar gráficas básicas para su uso en informes].	PT.1.1.H.STO.06: como [rol técnico 1] cuando logro [aplicar comandos de lenguaje R para gráficar] entoces el servicio MVITS [me permite generar gráficas sencillas para facilitar la interpretación de los datos].
7	USO.CASO.3: Servicios de Visualización Exploratorio de Datos	Como [Administrador o Usuario] puedo [generar gráficas avanzadas] para [la detección de patrones o diferencias, además de analizar su distribución].	2.H.STO.07.graficasAvanzadas	1	PA.1.H.STO.07: como [administrador o usuario] acepto la historia de usuario [2.H.STO.07.graficasAvanzadas] cuando pueda [generar gráficas con mayor complejidad].	PT.1.1.H.STO.07: como [rol técnico 1] cuando logro [instalar librerías para aumentar comandos gráficos] entoces el servicio MVITS [ofrece una lista amplia de opciones de gráficos para la representación idonea de cualquier tipo de dato].
3	USO.CASO.2: Servicios de Análisis Estadístico Descriptivo	Como [Administrador o Usuario] puedo [acceder, cargar y guardar paquetes de datos] para [crear, administrar y monitorear variables de interés].	3.N.STO.03.paqueteDatos	1	PA.1.N.STO.03: como [administrador o usuario] acepto la historia de usuario [3.N.STO.03.paqueteDatos] cuando pueda [generar paquete de datos con información, gráficas e informes].	PT.1.1.N.STO.08: como [rol técnico 1] cuando logro [exportar todo tipo de información de Rstudio a una aplicación de texto] entoces el servicio MVITS [permite generar informes extrañendo informacion relevante].
TOTAL DE PUNTOS DE HISTORIA				6		

Tabla 22: Hoja de pruebas de aceptación y pruebas técnicas de cada historia de usuario que conforma el ciclo #1.

4.2.4 T.2.4 Incremento de la ingeniería (tarea opcional)

T.2.4 Incremento de la ingeniería	
Prácticas ágiles	Pruebas de aceptación. Pruebas unitarias. Desarrollo de incremento. Definición de lo “Hecho.”
Tiempo de entrega	2 semanas
Actividades: ○ El Equipo de desarrollo implementa un conjunto de historias de usuario contenidas en el ciclo actual utilizando el boceto de la arquitectura como guía. La hoja de trabajo del ciclo se representa con la matriz de mapeado de historias. Esta implementación se realiza dentro de un entorno de desarrollo y estas nuevas implementaciones se integran ahí. ○ El Equipo de desarrollo aplica las pruebas de aceptación y pruebas técnicas a este nuevo incremento. En el caso inesperado de contar con una historia de usuario fallida que no puede ser implementada como se acordó, el Equipo de desarrollo se comunica con el Cliente y se realiza una negociación interna para: 1) definir una modificación a la historia de usuario. 2) Mover la historia de usuario a otro ciclo. 3) Eliminar la historia de usuario. ○ El Equipo de desarrollo realiza una prueba global en el nuevo incremento utilizando el objetivo del ciclo y la definición de “Hecho” una vez implementadas todas las historias de usuario del ciclo actual. ○ Por último, el Equipo de desarrollo libera el entorno de producción del nuevo incremento con la que se trabajará en la T.2.5.	

Tabla 23: Descripción de prácticas ágiles, tiempos de entrega y actividades involucradas en la tarea T.2.4 acorde a la metodología AVS_FD_MVITS.v1-2022.

4.2.5 T.2.5 Revisión de los ciclos

T.2.5 Revisión de los ciclos	
Prácticas ágiles	Definición de “Hecho”.
Tiempo de entrega	4 horas
Actividades: ○ El Experto organiza una reunión para la revisión de los ciclos. ○ El Cliente utiliza las pruebas de aceptación y pruebas técnicas para aceptar o rechazar cada historia de usuario acordada en cada ciclo (tabla 20 y 21). ○ El Cliente y el Equipo de desarrollo revisan juntos el objetivo del ciclo y la definición de “Hecho” (tabla 16 y 17). ○ El Experto registra la revisión de los ciclos (tabla 25 y 26).	

Tabla 24: Descripción de prácticas ágiles, tiempos de entrega y actividades involucradas en la tarea T.2.5 acorde a la metodología AVS_FD_MVITS.v1-2022.

META DEL CICLO [ciclo #1]	Meta del ciclo [1] es implementar [instalar componente TI; crear, eliminar y copiar bases de datos; procesar datos; calcular métricas descriptivas; y administrar usuarios] para obtener [entorno de trabajo optimo para albergar el servicio] con métricas de servicio previstas de [(funcionalidad, normal); (capacidad, alto); (disponibilidad, normal); (continuidad, normal); y (seguridad, normal)].
DEFINICIÓN DE LO "HECHO" [ciclo # 1]	Meta del ciclo [1] es implementar [instalar componente TI; crear, eliminar y copiar bases de datos; procesar datos; calcular métricas descriptivas; y administrar usuarios] para obtener [entorno de trabajo optimo para albergar el servicio] con métricas de servicio previstas de [(funcionalidad, normal); (capacidad, alto); (disponibilidad, normal); (continuidad, normal); y (seguridad, normal)].

CICLO				1	
# HISTORIAS DE USUARIO	ID EPIC(DIAGRAMA CASO DE USO	HISTORIA DE USUARIO	PRUEBAS DE ACEPTACIÓN	¿SE ACEPTA LA HISTORIA DE USUARIO? (● TOTALMENTE ACEPTADA; ⊕ PARCIALMENTE ACEPTADA; ⊖ NO ACEPTADA)	COMENTARIOS SOBRE LAS HISTORIAS DE USUARIO QUE FUERON: ⊕ PARCIALMENTE ACEPTADOS OR ⊖ NO ACEPTADOS
10	USO.CASO.5: Instalación de componentes Open-Source	Como [Administrador] debo [instalar las versiones adecuadas de los componentes Open-Source] para [garantizar el funcionamiento adecuado de los servicios].	PA.1.VH.STO.10: como [administrador] acepto la historia de usuario [1.VH.STO.10.instalación] cuando pueda [acceder a las aplicaciones de código abierto].	●	...
1	USO.CASO.1: Servicios de Base de Datos	Como [Administrador o Usuario] puedo [crear, eliminar y copiar bases de datos] para [mi análisis personal de datos].	PA.1.VH.STO.01: como [administrador o usuario] acepto la historia de usuario [1.VH.STO.01.baseDatos] cuando pueda [administrar mis bases de datos desde el servicio MVITS].	●	...
4	USO.CASO.2: Servicios de Análisis Estadístico Descriptivo	Como [Administrador o Usuario] puedo [preprocesar datos] para [garantizar que los datos estén listos para su análisis y facilite la extracción de información valiosa].	PA.1.VH.STO.04: como [administrador o usuario] acepto la historia de usuario [1.VH.STO.04.preprocesar] cuando pueda [obtener bases de datos limpias].	●	...
5	USO.CASO.2: Servicios de Análisis Estadístico Descriptivo	Como [Administrador o Usuario] puedo [calcular métricas descriptivas] para [proveer una interpretación personal de la información].	PA.1.VH.STO.05: como [administrador o usuario] acepto la historia de usuario [1.VH.STO.05.metricasDes] cuando pueda [obtener métricas descriptivas de mis datos empresariales].	●	...
8	USO.CASO.4: Administración de los Servicios	Como [Administrador o Usuario] puedo [gestionar la creación y administración de usuarios] para [llevar un control de usuarios que tienen acceso a los servicios y supervisar avances].	PA.1.VH.STO.08: como [administrador o usuario] acepto la historia de usuario [1.VH.STO.08.usuarios] cuando pueda [obtener un registro de usuarios y de los permisos que disponen].	●	...

Tabla 25: Hoja de registro sobre la revisión del ciclo #1 acorde a la metodología AVS_FD_MVITS.v1-2022.

META DEL CICLO [ciclo #2]	Meta del ciclo [2] es implementar [gestionar los servicios; administrar, consultar, imprimir, exportar e importar tablas; generar gráficas básicas; generar gráficas avanzadas; y administrar paquetes de datos] para obtener [máximo aprovechamiento de datos empresariales almacenados a través de los años] con métricas de servicio previstas de [(funcionalidad, normal); (capacidad, alto); (disponibilidad, normal); (continuidad, normal); y (seguridad, normal)].
DEFINICIÓN DE LO "HECHO" [ciclo # 2]	Hecho [ciclo 2] se logra cuando [(SI) todas las historias de usuario del ciclo son aceptadas; (SI) el incremento de la ingeniería del ciclo fue entregado; (SI) el repositorio del proyecto MVITS fue actualizado].

CICLO				2	
# HISTORIAS DE USUARIO	ID EPIC(DIAGRAMA CASO DE USO	HISTORIA DE USUARIO	PRUEBAS DE ACEPTACIÓN	¿SE ACEPTA LA HISTORIA DE USUARIO? (● TOTALMENTE ACEPTADA; ⊙ PARCIALMENTE ACEPTADA; ⊖ NO ACEPTADA)	COMENTARIOS SOBRE LAS HISTORIAS DE USUARIO QUE FUERON: ⊙ PARCIALMENTE ACEPTADOS OR ⊖ NO ACEPTADOS
9	USO.CASO.5: Instalación de componentes Open-Source	Como [Administrador o Usuario] puedo [gestionar los servicios de TI] para [controlar y mitigar fallas en lo servicios].	PA.1.VH.STO.09: como [administrador o usuario] acepto la historia de usuario [1.VH.STO.09.gestionServicio] cuando pueda [monitorear los servicios de la plataforma].	●	...
2	USO.CASO.1: Servicios de Base de Datos	Como [Administrador o Usuario] puedo [usar [administrar tablas, consultar tablas, imprimir tablas, exportar, importar] bases de datos] para [mi análisis personal de datos].	PA.1.H.STO.02: como [administrador o usuario] acepto la historia de usuario [2.H.STO.02.tablasDatos] cuando pueda [administrar tablas de datos para su análisis o interpretación].	●	...
6	USO.CASO.2: Servicios de Análisis Estadístico Descriptivo	Como [Administrador o Usuario] puedo [generar gráficas básicas] para [realizar una exploración inicial de manera visual].	PA.1.H.STO.06: como [administrador o usuario] acepto la historia de usuario [2.H.STO.06.graficasBasicas] cuando pueda [guardar gráficas básicas para su uso en informes].	●	...
7	USO.CASO.2: Servicios de Análisis Estadístico Descriptivo	Como [Administrador o Usuario] puedo [generar gráficas avanzadas] para [la detección de patrones o diferencias, además de analizar su distribución].	PA.1.H.STO.07: como [administrador o usuario] acepto la historia de usuario [2.H.STO.07.graficasAvanzadas] cuando pueda [generar gráficas con mayor complejidad].	●	...
3	USO.CASO.4: Administración de los Servicios	Como [Administrador o Usuario] puedo [acceder, cargar y guardar paquetes de datos] para [crear, administrar y monitorear variables de interés].	PA.1.N.STO.03: como [administrador o usuario] acepto la historia de usuario [3.N.STO.03.paqueteDatos] cuando pueda [generar paquete de datos con información, gráficas e informes].	●	...

Tabla 26: Hoja de registro sobre la revisión del ciclo #1 acorde a la metodología AVS_FD_MVITS.v1-2022.

4.3 Etapa 3. Versión final. Entrega y soporte.

Esta etapa final se compone de dos tareas: T.3.1 Soporte de ingeniería para la versión final y T.3.2 Reunión de entrega de la versión final (figura 17), se toma el PKG.01 y el PKG.02 para poder construir paquete de la versión final (PKG.03) el cual se compone por los elementos: Acuerdo de soporte para la versión final y Documento de aceptación de la versión final, además del repositorio digital con toda la documentación vista en las etapas anteriores y obviamente se hace entrega de la versión final de la plataforma de código abierto para ciencia de datos.

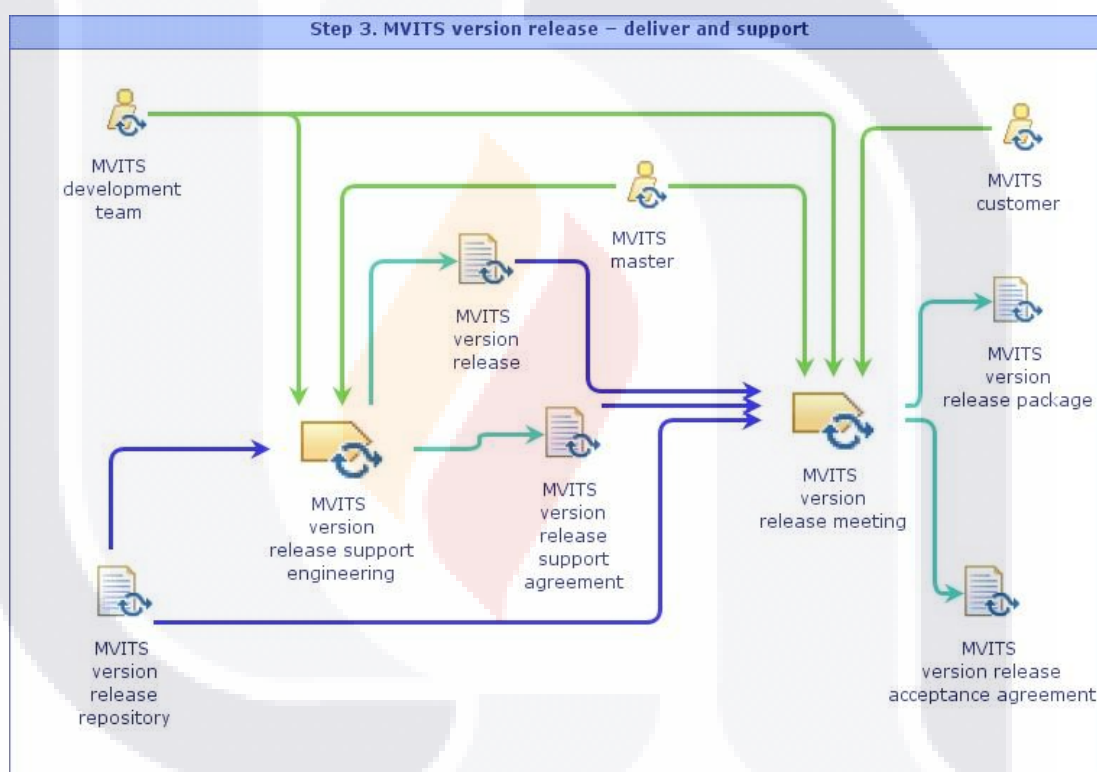


Figura 17: Etapa Entrega y Soporte (tomada de “Metodología Ágil para el Desarrollo de un Nuevo Servicio de TI Mínimo Viable (AVS_FD_MVITS.v1-2022)”).

4.3.1 T.3.1 Soporte de ingeniería para la versión final

T.3.1 Soporte de ingeniería para la versión final	
Prácticas ágiles	Entregas frecuentes.
Tiempo de entrega	1 día
Actividades: ○ El Equipo de desarrollo elabora el acuerdo de soporte para la versión final. Este documento representa una versión ágil del Acuerdo de Nivel de Servicio que se usa normalmente (tabla 28). ○ Este acuerdo de soporte para la versión final incluye: ▪ Descripción. ▪ Horario. ▪ Métricas de evaluación de servicio (funcionalidad, capacidad, disponibilidad, continuidad y seguridad). ▪ Datos del cliente (número de contacto, email). ▪ Recomendaciones de uso.	

Tabla 27: Descripción de prácticas ágiles, tiempos de entrega y actividades involucradas en la tarea T.3.1 acorde a la metodología AVS_FD_MVITS.v1-2022.

DESCRIPCIÓN	La [Plataforma Open-Source Web de Ciencia de Datos] será usada por [empleados de Sitis De Aguascalientes S.C.] para [crear, eliminar, copiar y usar bases de datos; dar acceso, cargar datos, guardar datos, preprocesar datos y calcular métricas; generar gráficas básicas y avanzadas] para tener [servicios de bases de datos, servicios de análisis estadístico descriptivo y servicios de visualización de datos] con métricas de servicio previstas de (funcionalidad, normal); (capacidad, alto); (disponibilidad, normal); (continuidad, normal); y (seguridad, normal).	
HORARIO	DÍAS DE LA SEMANA	HORAS DEL DÍA
	LUN-MAR-MIE-JUE-VIE	9:00 AM - 16:00 PM
MÉTRICAS	NIVEL DE FUNCIONALIDAD (NORMAL)	
	NORMAL (N)	
	NIVEL DE CAPACIDAD (VH,H,N)	NIVEL DE DISPONIBILIDAD (VH, H, N)
	ALTO (H)	NORMAL (N)
	NIVEL DE CONTINUIDAD (VH,H,N)	NIVEL DE SEGURIDAD (VH, H, N)
	NORMAL (N)	NORMAL (N)
DATOS DEL SOPORTE	CONTACTO	EMAIL / TEL
	ING. JUAN DANIEL GALLEGOS RODRIGUEZ	al161692@edu.uaa.mx
	DÍAS DE LA SEMANA	HORAS DEL DÍA
	LUN-MAR-MIE-JUE-VIE	9:00 AM - 16:00 PM
DATOS DEL CLIENTE	CLIENTE	EMAIL / TEL
	MITC. JORGE EDUARDO MACIAS LUEVANO	eduardo.maciasl@edu.uaa.mx
RECOMENDACIONES DE USO	1. USAR NAVEGADOR FIREFOX Y LA ULTIMA VERSIÓN DISPONIBLE COMPATIBLE CON EL SISTEMA OPERATIVO.	
	2. GESTIONAR LAS CONTRASEÑAS DE USUARIOS ADECUADAMENTE PARA EVITAR FILTRACIONES Y CAMBIAR CONTRASEÑAS PERIODICAMENTE.	
	3. REVISAR FIREWALL PARA COMPROBAR PUERTOS ACTIVOS.	
	4. NO MANEJAR INFORMACIÓN DELICADA POR CORREO ELECTRONICO, USAR LA PLATAFORMA.	
	...	

Tabla 28: Documento de acuerdo de soporte para la versión final acorde a la metodología AVS_FD_MVITS.v1-2022.

4.3.2 T.3.2 Reunión de entrega de la versión final

T.3.2 Reunión de entrega de la versión final	
Prácticas ágiles	Entregas frecuentes.
Tiempo de entrega	2 horas
Actividades: ○ El Cliente presenta el grupo de usuarios finales. ○ El Experto entrega el paquete de la versión final (PKG.03) al Cliente el cual contiene: ▪ Acuerdo de soporte para la versión final. ▪ Documento de aceptación de la versión final. ▪ Repositorio digital de la versión final. ▪ Versión final de la plataforma. ○ El Cliente y el Experto cierran los ciclos de trabajo de la versión final con el acuerdo de aceptación de la versión final firmado (tabla 30).	

Tabla 29: Descripción de prácticas ágiles, tiempos de entrega y actividades involucradas en la tarea T.3.2 acorde a la metodología AVS_FD_MVITS.v1-2022.

LUGAR Y FECHA	LUGAR	Date
	AGUASCALIENTES, MÉXICO	17 DE JUNIO DE 2024
DESCRIPCIÓN DEL MVITS	VISIÓN DEL MVITS	
	La [Plataforma Open-Source Web de Ciencia de Datos] será usada por [empleados de Sitis De Aguascalientes S.C.] para [crear, eliminar, copiar y usar bases de datos; dar acceso, cargar datos, guardar datos, preprocesar datos y calcular métricas; generar gráficas básicas y avanzadas] para tener [servicios de bases de datos, servicios de análisis estadístico descriptivo y servicios de visualización de datos] con métricas de servicio previstas de (funcionalidad, normal); (capacidad, alto); (disponibilidad, normal); (continuidad, normal); y (seguridad, normal).	
PAQUETE DE LA VERSIÓN FINAL DEL MVITS	¿LA VERSIÓN FINAL DEL MVITS ESTA LISTA? (SI O NO)	¿EL ACUERDO DE SOPORTE DE LA VERSIÓN FINAL ESTA LISTO? (SI O NO)
	SI	SI
DATOS DEL CLIENTE Y FIRMA DE ACEPTACIÓN DE LA VERSIÓN FINAL DEL MVITS	NOMBRE	FIRMA
	MITC. JORGE EDUARDO MACIAS LUEVANO	
DATOS DEL EXPERTO Y FIRMA DE ACEPTACIÓN DE LA VERSION FINAL DEL MVITS	NOMBRE	FIRMA
	DR. JOSÉ MANUEL MORA TAVAREZ	
DATOS DEL EQUIPO DE DESARROLLO Y FIRMA DE ACEPTACIÓN DE LA VERSIÓN FINAL DEL MVITS	NOMBRE	FIRMA
	ING. JUAN DANIEL GALLEGO RODRÍGUEZ	

Tabla 30: Documento de aceptación de la versión final del MVITS acorde a la metodología AVS_FD_MVITS.v1-2022.

5. Resultados de la Intervención.

Para la evaluación de la plataforma de código abierto (open-source) web de ciencia de datos desarrollada en las etapas anteriores, se implementó un instrumento capaz de medir el grado de adopción de este nuevo tipo de servicios de TI. Por otro lado, también se aplicó una encuesta demográfica con el objetivo de recopilar información de la muestra de expertos que tuvieron acercamiento a la versión final del MVITS.

El anexo 1 “Encuesta Demográfica” muestra las seis preguntas que conforman el instrumento encargado de extraer datos al grupo entrevistado y así conocer sus características.

El anexo 2 “Instrumento Conceptual de Métricas de Usabilidad de Tecnologías de Información” muestra los cuatro constructos encargados de medir la aceptación del servicio diseñado: utilidad, facilidad de uso, compatibilidad y valor. Para poder medir dichos conceptos se escogieron preguntas específicas, las cuales se seleccionaron por medio de la literatura científica (trabajos y estudios previos) usando una escala de Likert que va del 1 (Totalmente desacuerdo) al 5 (Totalmente de acuerdo).

6. Evaluación de la Intervención.

6.1 Datos Demográficos

Para evaluar la plataforma diseñada, se implementó una encuesta demográfica encargada de recopilar información sobre diferentes características de la población encuestada. Esto confirmó que el instrumento empleado para el análisis del grupo arrojó resultados interesantes y útiles para determinar el valor de la plataforma para dicho sector.

Número de pregunta	Descripción	Resultado
1	Tipo de organización de trabajo.	Empresa PYME: 37.5% Institución Educativa: 62.5% Unidad de Gobierno: 0%
2	Nivel del puesto laboral.	Posición técnica de TI: 37.5% Puesto administrativo: 0% Académico: 37.5% Estudiante de Posgrado: 25%
3	Máximo nivel de escolaridad	Licenciatura: 0% Especialidad Profesional: 0% Maestría: 37.5% Doctorado: 62.5%
4	Rango de edad	Menos de 24 años: 0% 25-34 años: 25% 35-44 años: 37.5% 45-54 años: 25% Mas de 55 años: 12.5%
5	Experiencia en gestión de servicios de TI	Novato: 0% Inicial: 0% Normal: 37.5% Avanzado: 37.5% Experto: 25%
6	Experiencia con herramientas de ciencia de datos y analítica	Novato: 0% Inicial: 0% Normal: 62.5% Avanzado: 37.5% Experto: 0%

Tabla 31: Distribución de resultados de la encuesta demográfica.

Observando la tabla 31, se infiere de la media de los datos de la encuesta demográfica que gran parte de la población encuestada trabaja en una institución educativa, cuentan con un puesto de trabajo de posición técnica de TI o es un académico, cuentan con doctorado que es su máximo grado de estudios y oscilan entre los [35-44] años de edad.

Los profesionistas que conforman la población encuestada cuentan con un grado de experiencia en gestión de servicios de TI entre normal (37.5%) y avanzado (37.5%), lo que nos indica que la mayoría de ellos han tenido alguna experiencia previa o han diseñado un servicio por lo que conocen los marcos de metodología, procesos, conceptos y/o etapas que se pueden encontrar en este tipo de trabajos de estudio. Además, la mayoría de los expertos que conforman la población consideran que su nivel de experiencia con herramientas de ciencia de datos y analítica va desde un 40% hasta máximo 60% respecto a su comprensión y entendimiento de estas, por lo cual es una muestra de la importancia de aumentar la difusión de conocimiento sobre este tipo de herramientas, ya que pueden ser de gran utilidad para analizar datos almacenados sin ningún uso.

Por último, se escogieron las métricas (constructos) que actuarán como método de evaluación para determinar el valor de la plataforma web de código abierto para la ciencia de datos diseñada en este trabajo práctico, las cuales fueron las siguientes:

- Utilidad.
- Facilidad de uso.
- Compatibilidad.
- Valor.

6.2 Estadística Descriptiva

Para poder calcular las siguientes estadísticas descriptivas se aprovechó el software IBM SPSS Statistics versión 22.0.0.0, dado que ya se tenían bases para su uso y aplicación trabajando con diversos tipos de datos, dicho conocimiento fue adquirido en el módulo “Herramientas para la obtención y validación de datos” el cual forma parte del programa de estudio de la maestría: informática y tecnologías computacionales impartida por la Universidad Autónoma de Aguascalientes.

Como primer paso se capturaron los datos recopilados de cada encuesta, dividiéndolos por constructo (métrica de evaluación) por medio del id de la pregunta como se muestra en la figura 18, ordenándolos respecto al primer profesional encuestado (1) hasta el último (8). Cabe mencionar que se tuvieron que estipular los siguientes campos de cada variable (pregunta): nombre, tipo, anchura, decimales, valores, alineación, medida y rol como se muestra en la figura 19.

Registro	UT.1	UT.2	UT.3	UT.4	FU.1	FU.2	FU.3	CO.1	CO.2	CO.3	VA.1	VA.2	VA.3	VA.4
1	4	4	3	3	5	5	1	5	4	4	3	3	5	5
2	4	4	4	5	5	5	1	5	5	4	4	5	5	5
3	4	5	5	4	5	5	5	3	3	4	4	5	5	5
4	4	4	4	3	5	5	1	4	4	5	5	4	4	4
5	4	4	3	3	4	5	1	4	4	3	3	4	3	3
6	5	5	5	5	5	5	1	5	5	4	5	5	5	5
7	4	5	4	4	4	4	1	3	4	4	4	4	4	4
8	5	5	5	5	5	4	5	5	5	5	5	5	5	5
9														
10														
11														
12														
13														
14														
15														
16														
17														
18														
19														
20														
21														
22														
23														

Figura 18: Base de datos acorde a la información recopilada con ayuda artefacto descrito en el anexo 2.

Analisis descriptivo_trabajo practico.sav [Conjunto_de_datos1] - IBM SPSS Statistics Editor de datos

Archivo Editar Ver Datos Transformar Analizar Marketing directo Gráficos Utilidades Ventana Ayuda

	Nombre	Tipo	Anchura	Decimales	Etiqueta	Valores	Perdidos	Columnas	Alineación	Medida	Rol
1	Registro	Numérico	1	0		Ninguna	Ninguna	8	Centrado	Ordinal	Entrada
2	UT.1	Numérico re...	1	0		{1, Total de...	Ninguna	8	Centrado	Escala	Entrada
3	UT.2	Numérico	1	0		{1, Total de...	Ninguna	8	Centrado	Escala	Entrada
4	UT.3	Numérico	1	0		{1, Total de...	Ninguna	8	Centrado	Escala	Entrada
5	UT.4	Numérico	1	0		{1, Total de...	Ninguna	8	Centrado	Escala	Entrada
6	FU.1	Numérico	1	0		{1, Total de...	Ninguna	8	Centrado	Escala	Entrada
7	FU.2	Numérico	1	0		{1, Total de...	Ninguna	8	Centrado	Escala	Entrada
8	FU.3	Numérico	1	0		{1, Total de...	Ninguna	8	Centrado	Escala	Entrada
9	CO.1	Numérico	1	0		{1, Total de...	Ninguna	8	Centrado	Escala	Entrada
10	CO.2	Numérico	1	0		{1, Total de...	Ninguna	8	Centrado	Escala	Entrada
11	CO.3	Numérico	1	0		{1, Total de...	Ninguna	8	Centrado	Escala	Entrada
12	VA.1	Numérico	1	0		{1, Total de...	Ninguna	8	Centrado	Escala	Entrada
13	VA.2	Numérico	1	0		{1, Total de...	Ninguna	8	Centrado	Escala	Entrada
14	VA.3	Numérico	1	0		{1, Total de...	Ninguna	8	Centrado	Escala	Entrada
15	VA.4	Numérico	1	0		{1, Total de...	Ninguna	8	Centrado	Escala	Entrada
16											
17											
18											
19											
20											
21											
22											
23											
24											

Figura 19: Características definidas para cada variable de la base de datos.

Las tablas 20 a la 23 muestran los valores obtenidos del análisis estadístico descriptivo (mínimo, máximo, media, desviación estándar y asimetría) de la base de datos, divididos conforme al id de la pregunta y al constructo al que pertenece.

Estadísticos descriptivos

	N	Mínimo	Máximo	Media	Desviación estándar	Asimetría	
	Estadístico	Estadístico	Estadístico	Estadístico	Estadístico	Estadístico	Error estándar
UT.1	8	4	5	4.25	.463	1.440	.752
UT.2	8	4	5	4.50	.535	.000	.752
UT.3	8	3	5	4.13	.835	-.277	.752
UT.4	8	3	5	4.00	.926	.000	.752
N válido (por lista)	8						

Tabla 32: Resultados del análisis descriptivo del constructo utilidad.

Estadísticos descriptivos							
	N	Mínimo	Máximo	Media	Desviación estándar	Asimetría	
	Estadístico	Estadístico	Estadístico	Estadístico	Estadístico	Estadístico	Error estándar
FU.1	8	4	5	4.75	.463	-1.440	.752
FU.2	8	4	5	4.75	.463	-1.440	.752
FU.3	8	1	5	2.00	1.852	1.440	.752
N válido (por lista)	8						

Tabla 33: Resultados del análisis descriptivo del constructo facilidad de uso.

Estadísticos descriptivos							
	N	Mínimo	Máximo	Media	Desviación estándar	Asimetría	
	Estadístico	Estadístico	Estadístico	Estadístico	Estadístico	Estadístico	Error estándar
CO.1	8	3	5	4.25	.886	-.615	.752
CO.2	8	3	5	4.25	.707	-.404	.752
CO.3	8	3	5	4.13	.641	-.068	.752
N válido (por lista)	8						

Tabla 34: Resultados del análisis descriptivo del constructo compatibilidad.

Estadísticos descriptivos							
	N	Mínimo	Máximo	Media	Desviación estándar	Asimetría	
	Estadístico	Estadístico	Estadístico	Estadístico	Estadístico	Estadístico	Error estándar
VA.1	8	3	5	4.13	.835	-.277	.752
VA.2	8	3	5	4.38	.744	-.824	.752
VA.3	8	3	5	4.50	.756	-1.323	.752
VA.4	8	3	5	4.50	.756	-1.323	.752
N válido (por lista)	8						

Tabla 35: Resultados del análisis descriptivo del constructo valor.

Para poder interpretar los datos obtenidos de cada métrica de evaluación se recurre a la literatura científica sobre conceptos estadísticos. Se formularon los siguientes enunciados para ayudar en la interpretación adecuada de la información:

- Los apartados mínimo y máximo se refieren a los valores más bajos y altos respectivamente que se pueden encontrar en dicha variable. Por ejemplo, en la pregunta UT.1 el menor valor que se puede encontrar en la respuesta de todos los profesionistas es de 4 (Alto) y el mayor valor que se puede encontrar es de 5 (Total Acuerdo).
- El valor de la media nos indica el valor central de los datos o también conocida como el promedio general de los datos. Por ejemplo, en la pregunta VA.1 el

valor de la media es de 4.13 lo que significa que en promedio la respuesta a dicha pregunta sería 4 (Acuerdo).

- c) El apartado de la desviación estándar nos indica que tan dispersos o juntos están los datos de la variable. Por lo que se infiere que un valor de desviación estándar pequeño indica que los valores se agrupan respecto a la media, o si se cuenta con un valor de desviación estándar grande significa que los datos están alejados de la media. Por ejemplo, en la pregunta CO.2 el valor de la desviación estándar es de 0.707 lo que indica que los datos que integran dicha pregunta se alejan de la media ± 0.707 , en otras palabras el valor de respuesta esperado puede estar 0.707 arriba o 0.707 abajo.
- d) El valor del apartado de asimetría (estadístico) indica que tan simétrica es la distribución de los datos respecto a la media, es decir, si se cuenta con un valor de asimetría positivo ($x > 0$) la distribución tiende irse a la izquierda respecto a la media lo que puede indicar valores altos atípicos, o si el valor de la asimetría es negativo ($x < 0$) la distribución tiende irse a la derecha respecto a la media y puede indicar valores bajos atípicos. Por ejemplo, en la pregunta FU.3 el valor de la asimetría es de 1.440, lo que indica que los datos tienden a encontrarse del lado izquierdo respecto a la media, en este caso la media tiene valor de 2, por ende la mayoría de los profesionistas respondieron 1 (Totalmente desacuerdo) en dicha pregunta y también se puede inferir que existen valores atípicos altos, ya que solo dos profesionistas respondieron 5 (Total acuerdo) en este apartado.

7. Conclusiones.

El objetivo general del trabajo práctico (descrito en el apartado 1.2 de acuerdo con el índice de este documento) se logró satisfactoriamente, mediante al cumplimiento de los objetivos específicos:

1. Elaborar una revisión comparativa sumariada entre los procesos de diseño de Servicios de TI principales de Rigor (ITIL v3.v2011) y de Agilidad (ITIL v4, AVS_FD_MVITS).
2. Revisar las funcionalidades y requerimientos de utilización de las principales herramientas de tipo Open Source para las diversas fases de la Ciencia de Datos, y seleccionar las herramientas a utilizar en este Trabajo Práctico.
3. Implementar un Servicio de TI (prototipo) de Plataforma Open-Source Web de Ciencia de Datos para PYMES, usando la metodología ágil AVS_FD_MVITS.
4. Obtener una evaluación piloto del Servicio de TI (prototipo) implementado respecto a la utilidad, facilidad de uso, compatibilidad y valor, por parte de un grupo de usuarios de la Organización Cliente contactada.

7.1 Conclusiones del objetivo específico 1

Para poder lograr dicho objetivo, se consultaron manuales de orientación para la implementación de estos procesos de diseño de servicios de TI de rigor, para el caso de ITILv3, y metodologías de gestión de procesos de diseño ágil, tanto para ITILv4 como AVS_FD_MVITS. Se identificaron roles, actividades y artefactos claves para todo el proceso de diseño de un nuevo servicio TI, para posteriormente reflejar estos puntos clave en una tabla comparativa.

7.2 Conclusiones del objetivo específico 2

Para alcanzar este objetivo, se visitaron varias páginas de proveedores de diferentes herramientas de código abierto, las cuales pueden ser implementadas en alguna de las etapas que conformar las fases genéricas del proceso de la ciencia de datos, ya se para identificar o adquirir, preprocesamiento, almacenamiento y procesamiento, modelado y análisis, o para acceso y uso dichas herramientas ofrecen ventajas que pueden ser

útiles para empresas de recursos limitados como las PYMES. Se identificaron ventajas de todas las herramientas para ayudar en la selección de componentes TI que integrarían la plataforma web de ciencia de datos.

Los elementos seleccionados para poder desempeñar tareas de las fases de ciencia de datos fueron los siguientes:

- MySQL + CentOS 7: para desempeñar tareas de preprocesamiento, procesamiento y proporcionar un gestor de bases de datos para administrar el almacenamiento de estos.
- Paquete R de estadística + Servidor Rstudio: para poder desempeñar tareas de modelado y análisis para presentar hallazgos obtenidos y presentarlos por medio de graficas sencillas o avanzadas. También se pensó en implementar Jupyter Notebook para usuarios expertos en algoritmos o con conocimientos en dicho entorno de desarrollo pero se descartó por falta de tiempo.

7.3 Conclusiones del objetivo específico 3

Para cumplir con el este objetivo específico, se tuvo que entender todos los componentes que integran la metodología ágil AVS_FD_MVITS_v1-2022 para poder implementar sus etapas y desarrollar cada tarea de forma adecuad, para justificar cada artefacto y práctica ágil o rol involucrado en ellas. Se reconocen las etapas 1 y 2 de la metodología, ya que estipulan las bases de diseño del nuevo servicio de TI que se toman en cuenta hasta la última reunión de entre de la versión final.

7.4 Conclusiones del objetivo específico 4

Para alcanzar el objetivo específico 4, se realizaron encuestas de evaluación a la plataforma web open-source de ciencia de datos, las cuales permitieron medir la percepción que tiene una población de profesionistas para adoptar un nuevo servicio de TI y proporcionaron resultados alentadores. Las métricas de evaluación de servicio fueron las siguientes:

- Utilidad.
- Facilidad de uso.

- Compatibilidad.
- Valor.

Los artefactos de evaluación fueron aplicados a una población de 8 potenciales usuarios finales, dicha población se constituía de profesionales con conocimiento especializado en algunas de las áreas de las tecnologías de información. Los usuarios fueron contactados por medio de correo electrónico y se les proporciono un curso de capacitación por medio de un video con duración de 20 minutos.

Los resultados plasmados en el apartado seis de acuerdo con el índice, fueron analizados por medio de pruebas estadísticas con el objetivo de determinar como perciben los usuarios potenciales a esta nueva plataforma.

7.5 Beneficios obtenidos

La plataforma web open-source de ciencia de datos para PYMES, ofrece a las empresas de bajos recursos una opción para incursionar en el análisis de sus bases de datos generadas a través de los años, para poder extraer nueva información que les permita tomar buenas decisiones empresariales u obtener nuevas ventajas competitivas que les ayude a prosperar.

La plataforma fue percibida como útil por los usuarios finales, esto se puede inferir por medio del anexo 1 y 2, los cuales conforman los dos cuestionarios donde se evalúan criterios demográficos y métricas de evaluación de servicio (utilidad, facilidad de uso, compatibilidad y valor).

7.6 Problemas adicionales encontrados

Uno de los principales problemas que se presentó en la etapa de diseño de la plataforma fue el poder encontrar las versiones de los componentes TI compatibles entre sí, ya que al instalar alguna de las versiones disponibles par el sistema operativo CentOS Linux 7 no garantizaba el contar con todos sus paquetes o funciones, además de no permitir la conexión con otro de las herramientas seleccionadas.

Al trabajar con comandos de terminal de sistema para instalar algunos componentes TI se pueden dañar los repositorios lo que causa que la plataforma pierda funcionalidad o incluso llegar a dañar la interfaz gráfica de usuario (GUI) y por ende se tendría que volver a crear la máquina virtual si se está trabajando con virtualización, si se contara con un servidor como Windows server 2012 que cuenta con opciones de descargas directas de versiones compatibles y se evita trabajar con comando como en el caso de Linux que dependiendo de la distribución no se cuenta con dicha función.

Por último, si se decide el implementar un servicio de TI como este y se piensa aprovechar la virtualización con algún gestor de máquinas virtuales como VirtualBox de Oracle, es importante que la computadora física donde se almacene el proyecto cuente con algún sistema operativo que permita habilitar la opción de escritorio remoto, ya que no importa que se active la opción dentro del gestor de máquinas virtuales por el simple hecho de que se utiliza la dirección IP de la computadora física y si esta misma opción no se activa el acceso remoto será imposible por medio de un protocolo de escritorio remoto (RDP), a menos que se adquiriera versiones empresariales de estos sistemas operativos generando un costo de inversión. Sin embargo, se puede recurrir a otra opción como Ngrok el cuales es un servicio que permite subir a un servidor en línea alguna aplicación que estemos corriendo en nuestra máquina física o virtual, lo que es perfecto para el servidor de Rstudio, no obstante, se tienen que estudiar los manuales de usuario para detectar posibles riesgos de seguridad que se puedan generar al agregar este tipo de alternativas, para crear backdoors y vulnerar nuestra infraestructura empresarial.

7.7 Recomendaciones finales para el cliente y usuarios

Este trabajo práctico puede tomarse como base para crear un nuevo servicio de TI con características mínimas y que sea viable, para cualquier rama que se dese explorar dentro de empresas PYMES aprovechando paquetes de código abierto y utilizando la infraestructura actual para ahorrar costos y generar hallazgos que puedan servir para tomar decisiones con sustento científico.

Este tipo de plataformas de ciencia de datos puede ayudar para el análisis de información empresarial y darles algún uso a los miles de gigabytes de datos que puedan estar siendo almacenados sin ningún propósito, causando simplemente gastos de almacenamiento. Es por lo anterior, que se hace la invitación a intentar aplicar lo descrito en este trabajo práctico para compartir los resultados y perfeccionar la metodología ágil AVS_FD_MVITS_v1-2022.

7.8 Conclusiones finales

Los marcos de metodología de gestión de servicios de TI ágiles o rigurosos, son un conjunto de mejores prácticas idóneas para administrar los procesos del negocio de manera eficiente, sin embargo, estos requieren certificaciones o contratación de personal experto que cuente con las credenciales que otorgan este tipo de certificaciones, lo que puede ser abrumador para las empresas pequeñas y/o medianas al suponer el costo de inversión sin contar con la garantía de obtener los resultados deseados. Es por lo anterior, que en mi opinión personal se deben buscar alternativas diseñadas específicamente para el tipo de empresa al que se pertenezca, ya que pueden ofrecer mejor diseño de proceso u ofrezcan plantillas que pueden ser modificadas con datos empresariales propios para analizar los posibles beneficios que puedan ofrecer respaldados con resultados estadísticos favorables.

Por último, la elaboración de este trabajo práctico me ayudó a comprender que el implementar este tipo de soluciones siguiendo alguna metodología de diseño de nuevos servicios TI o actualización de procesos de servicios actuales, puede ser beneficioso tanto para las empresas a quien ofrezcamos el proyecto, como para nosotros (estudiantes) contar con conocimiento previo de estos temas para aplicarlo en posibles cargos futuros.

Glosario

ITIL: Conjunto de publicaciones de buenas prácticas para la gestión de servicios de TI. Propiedad de la Oficina del Gabinete (parte del Gobierno de Su Majestad), ITIL ofrece orientación sobre la prestación de servicios de TI de calidad y los procesos, funciones y otras capacidades necesarias para apoyarlos. El marco ITIL se basa en un ciclo de vida del servicio y consta de cinco etapas (estrategia del servicio, diseño del servicio, transición del servicio, funcionamiento del servicio y mejora continua del servicio), cada una de las cuales tiene su propia publicación de apoyo (ITIL.V.2011.GLOSSARY.ACRONYMS, 2011).

Tecnología de Información (TI): Uso de la tecnología para almacenar, comunicar o procesar información. La tecnología suele incluir ordenadores, telecomunicaciones, aplicaciones y otros programas informáticos. La información puede incluir datos empresariales, voz, imágenes, vídeo, etc. La tecnología de la información se utiliza a menudo para apoyar los procesos empresariales a través de los servicios de TI (ITIL.V.2011.GLOSSARY.ACRONYMS, 2011).

Gestión de Servicios: Conjunto de capacidades organizativas especializadas para proporcionar valor a los clientes en forma de servicios.

Gestión de Servicios Informáticos: Implantación y gestión de servicios informáticos de calidad que satisfagan las necesidades de la empresa. La gestión de servicios de TI la realizan los proveedores de servicios de TI mediante una combinación adecuada de personas, procesos y tecnologías de la información. Implantación y gestión de servicios informáticos de calidad que satisfagan las necesidades de la empresa. Los proveedores de servicios de TI llevan a cabo la gestión de los servicios de TI mediante una combinación adecuada de personas, procesos y tecnologías de la información.

Paquete de Diseño de Servicios: Documento(s) que definen todos los aspectos de un servicio de TI y sus requisitos a través de cada etapa de su ciclo de vida. Se elabora un paquete de diseño del servicio para cada nuevo servicio de TI, cambio importante o retirada de un servicio de TI.

Roles: Un rol es un conjunto de responsabilidades, actividades y autoridades otorgadas a una persona o equipo. Un rol se define en un proceso o función. Una persona o equipo puede tener varios roles. ITIL no describe todos los roles que pueden existir en una organización, sino que proporciona ejemplos representativos para ayudar a las organizaciones a definir sus propios roles.

Métricas de calidad de servicio: son indicadores para medir el rendimiento, la calidad y el cumplimiento de los acuerdos en la prestación de un servicio. Estas métricas permiten evaluar si un servicio cumple con los objetivos establecidos, identificar áreas de mejora y garantizar la alineación con las necesidades del cliente o del negocio.

Clúster de computadoras: conjunto de computadoras independientes conectadas entre ellas, las cuales trabajan de tal forma como si fueran un único sistema y poder realizar tareas complejas distribuyendo la carga de trabajo.

Backlog: lista de tareas y requerimientos necesarios para apoyar una planificación estratégica.

MVITS: siglas en ingles de “Servicios de TI Mínimos Viables”.

Bibliografía

- Aguirre, R. (2015). Factores Determinantes en la adopción de tecnologías de información (TI) en las PYMES. *Universidad Autónoma de Nuevo León*, 1–23.
- Alcaraz Rodríguez, R. (2011). *El emprendedor de éxito*.
- Apache Software Foundation. (2023a). *Apache Storm*.
- Apache Software Foundation. (2023b, November 21). *Apache Hive*. <https://Hive.Apache.Org/>.
- Apache Software Foundation. (2023c, November 21). *Apache Kylin: Analytical Data Warehouse for Big Data*. <https://Kylin.Apache.Org/>.
- Boone, L. (2022). Industry 4.0 (Fourth industrial revolution). In *Salem Press Encyclopedia*. Salem Press. <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=ers&AN=119214086&lang=es&site=eds-live&scope=site>
- Bouwman, H., Nikou, S., Molina-Castillo, F. J., & de Reuver, M. (2018). The impact of digitalization on business models. *Digital Policy, Regulation and Governance*, 20(2), 105–124. <https://doi.org/10.1108/DPRG-07-2017-0039>
- Cartlidge, A., Ashley, X.-S., Hp, H., Rudd, C., Macfarlane, I., Windebank, J., Stuart, S., & Hp, R. (n.d.). *The IT Infrastructure Library An Introductory Overview of ITIL® V3*.
- Databricks. (2023, November 21). *Data Lakehouse*. <https://Www.Databricks.Com/Glossary/Data-Lakehouse>.
- Delen, D., & Demirkan, H. (2013). Data, information, and analytics as services. In *Decision Support Systems* (Vol. 55, Issue 1, pp. 359–363). Elsevier B.V. <https://doi.org/10.1016/j.dss.2012.05.044>
- Heredia, E. Á. (2014). *LAS PYMES EN MÉXICO: DESARROLLO Y COMPETITIVIDAD*.

Hindman, B., Konwinski, A., Zaharia, M., Ghodsi, A., Joseph, A. D., Katz, R., & Shenker, S. (n.d.). *Mesos: A Platform for Fine-Grained Resource Sharing in the Data Center*.

Hunnebeck, L., Rudd, C., Lacy, S., & Hanna, A. (2011). *ITIL service design* (2011th ed.). TSO (The Stationery Office).

International Business Machines (IBM). (2023, October 23). *HDFS*. <https://www.ibm.com/es-es/topics/hdfs>.

ITIL.V.2011.GLOSSARY. ACRONYMS. (n.d.).

Jupyter Team. (2023, November 29). *Project Jupyter Documentation*. <https://docs.jupyter.org/en/latest/>.

Medvedev, D., Lemson, G., & Rippin, M. (2016). SciServer compute: Bringing analysis close to the data. *ACM International Conference Proceeding Series, 18-20-July-2016*. <https://doi.org/10.1145/2949689.2949700>

Melendez, K., Dávila, A., & Pessoa, M. (2016). Information technology service management models applied to medium and small organizations: A systematic literature review. In *Computer Standards and Interfaces* (Vol. 47, pp. 120–127). Elsevier B.V. <https://doi.org/10.1016/j.csi.2015.10.001>

Mora, M., Adelakun, O., & Reyes-Delgado, Y. (2019). *Agile Value Stream for Developing a new Minimum Viable IT Service (based on Scrum, XP, ITIL v4 Agile Practices, and SysML v.1.6-2019)*.

Mora, M., Reyes-Delgado, P., Galvan-Cruz, S., Solano-Romo, L., Mora, M., Yuritzky Reyes-Delgado, P., Galvan-Cruz, S., & Solano-Romo, L. I. (2023). *Open-Source IT for Delivering Big Data Analytics Systems as Services-a Selective Review*.

Ofe, H. A., De Reuver, M., Nederstigt, B., & Janssen, M. (2023). *Data Analytics Platforms: Value Propositions and Adoption Challenges for Small Hospitality Businesses*. University of Hawaii at Manoa.

Olmera, M., & Rosas, X. (2011). Fundamentos de Tecnologías de Información: Viviendo en una Sociedad Tecnológica. *Editorial Digital: Tecnológico de Monterrey, Volumen I*.

Oracle Corporation. (2023, November 21). *MySQL 8.2 Reference Manual*. <https://dev.mysql.com/doc/refman/8.2/en/>.

Pantano, J. C., & Romagnano, M. (2022). A Pool of Free Software Tools to Assist Business Intelligence and Analytics. *International Journal of Combinatorial Optimization Problems and Informatics*, 13(2), 14–32. www.editada.org

Reyes-Delgado, P. Y., Mora, M., Wang, F., & Gómez, J. M. (2023). AHP evaluation of rigorous and agile IT service design-building phases-workflows in data centers. *Journal of Supercomputing*. <https://doi.org/10.1007/s11227-023-05219-x>

Rose, K., Eldridge, S., & Chapin, L. (2015). *The Internet of Things: An Overview Understanding the Issues and Challenges of a More Connected World*.

Sarker, I. H. (2021). Data Science and Analytics: An Overview from Data-Driven Smart Computing, Decision-Making and Applications Perspective. In *SN Computer Science* (Vol. 2, Issue 5). Springer. <https://doi.org/10.1007/s42979-021-00765-8>

Sharma, L., & Gupta, V. (2019). A comparative study of data analytics open-source tools. In *National Journal of Multidisciplinary Research and Development* www.nationaljournals.com (Vol. 4, Issue 2). www.nationaljournals.com

Stationery Office, T. (2012). *ITIL ® Foundation Handbook ITIL ® Foundation Handbook Best Management Practice Portfolio Product*. www.tso.co.uk

The Apache Software Foundation. (2020). *Apache Cloud Stack™ (Open-Source Cloud Computing™)*. <https://cloudstack.apache.org/about.html>.

The Apache Software Foundation. (2023a, November 21). *Apache HBase™*. <https://hbase.apache.org/>.

TESIS TESIS TESIS TESIS TESIS

The Apache Software Foundation. (2023b, November 21). *Apache Impala*.
<https://impala.apache.org/Overview.html>.

The Apache Software Foundation. (2023c, November 21). *Developer's Guide for Apache Mahout*. <https://mahout.apache.org/documentation/developers/>.

Thomas H. Davenport, & DJ Patil. (2012). Data Scientist_ The Sexiest Job of the 21st Century. *Harvard Business Review*.

Yang, F., Merlino, G., Ray, N., Léauté, X., Gupta, H., & Tschetter, E. (2017). *The RADStack: Open-Source Lambda Architecture for Interactive Analytics*.
<http://hdl.handle.net/10125/41359>

Zhao, L., & Deek, F. P. (2005). *Association for Information Systems AIS Electronic Library (AISeL) Improving Open-Source Software Usability Recommended Citation Improving Open-Source Software Usability*.
<http://aisel.aisnet.org/amcis2005/430>

Anexos

ENCUESTA DEMOGRÁFICA

EVALUADORES DE PLATAFORMA DE SERVICIOS DE TI: Open-Source Web de Ciencia de Datos para PYMES

INSTRUCCIONES. Por favor, agradecemos el llenar el siguiente cuestionario respecto a datos demográficos:

1 - Marque solamente una respuesta que mejor describa el tipo de la organización de su trabajo:

- ☐ Empresa PYME
- ☐ Institución Educativa.
- ☐ Unidad de Gobierno

2- Marque solamente una respuesta que mejor describa el nivel de su puesto laboral en su organización:

- ☐ Una posición técnica de TI
- ☐ Un puesto administrativo
- ☐ Un académico
- ☐ Un estudiante de Posgrado

3- Marque solamente una respuesta que mejor describa su máximo nivel escolar alcanzado:

- ☐ Licenciatura
- ☐ Especialidad Profesional (después de una Licenciatura)
- ☐ Graduado de Nivel I (Maestría o Pasante Maestría)
- ☐ Graduado de Nivel II (Doctorado o Candidato a Doctor)

4- Marque solamente una respuesta que mejor describa su rango de edad:

- ☐ x-24
- ☐ 25-34
- ☐ 35-44
- ☐ 45-54
- ☐ 55- o más

5.- Marque solamente una respuesta que mejor describa su auto-evaluación actual sobre su experiencia en la comprensión y entendimiento del enfoque de Gestión de Servicios de TI:

- ☐ Novato (hasta 20%)
- ☐ Inicial (más de 20% y hasta 40%)
- ☐ Normal (más de 40% y hasta 60%)
- ☐ Avanzado (más de 60% y hasta 80%)
- ☐ Experto (más de 80%)

6.- Marque solamente una respuesta que mejor describa su auto-evaluación actual sobre su experiencia en la comprensión y entendimiento del enfoque y herramientas de la Ciencia de Datos y Analítica :

- ☐ Novato (hasta 20%)
- ☐ Inicial (más de 20% y hasta 40%)
- ☐ Normal (más de 40% y hasta 60%)
- ☐ Avanzado (más de 60% y hasta 80%)
- ☐ Experto (más de 80%)

¡Muchas gracias por su valiosa participación!

Anexo 1: Encuesta demográfica.

INSTRUMENTO CONCEPTUAL DE MÉTRICAS DE USABILIDAD DE TECNOLOGÍAS DE INFORMACIÓN

EVALUADORES DE PLATAFORMA DE SERVICIOS DE TI: Open-Source Web de Ciencia de Datos para PYMES

INSTRUCCIONES. Por favor, agradecemos el asignar de manera personal a cada estatuto el grado de acuerdo o desacuerdo que percibo sobre la Tecnología de Información a evaluar de [PLATAFORMA DE SERVICIOS DE TI OPEN-SOURCE WEB DE CIENCIA DE DATOS PARA PYME: [OSWebCienciaDatosPymes] en base su experiencia en TIs, Ciencia de Datos y el video-demo de 20 minutos revisado. Favor de no dejar algún reactivo sin contestar. Muchas gracias!

Constructo	Total desacuerdo	Desacuerdo	Parcialmente Acuerdo-Desacuerdo	Acuerdo	Total Acuerdo
UTILIDAD	1	2	3	4	5
UT.1 Utilizar la Plataforma de Servicios de TI [OSWebCienciaDatosPymes] me habilitaría a cumplir mis tareas de [Análisis de Datos en mi organización] más rápidamente.					
UT.2 Utilizar la Plataforma de Servicios de TI [OSWebCienciaDatosPymes] mejoraría la calidad de mis tareas de [Análisis de Datos en mi organización].					
UT.3 Usar la Plataforma de Servicios de TI [OSWebCienciaDatosPymes] realzaría la efectividad de mi [Análisis de Datos en mi Organización].					
UT.4 Usar la Plataforma de Servicios de TI [OSWebCienciaDatosPymes] me daría mayor control sobre el proceso de [Análisis de Datos en mi Organización].					

Construceto	Total desacuerdo	Desacuerdo	Parcialmente Acuerdo-Desacuerdo	Acuerdo	Total Acuerdo
FACILIDAD DE USO	1	2	3	4	5
FU.1 Aprender a utilizar/operar la Plataforma de Servicios de TI [OSWebCienciaDatosPymes] sería fácil para mí.					
FU.2 En caso de obligación de usar la Plataforma de Servicios de TI [OSWebCienciaDatosPymes], sería fácil para mí.					
FU.3 En caso de obligación de usar la Plataforma de Servicios de TI [OSWebCienciaDatosPymes], sería difícil para mí.					

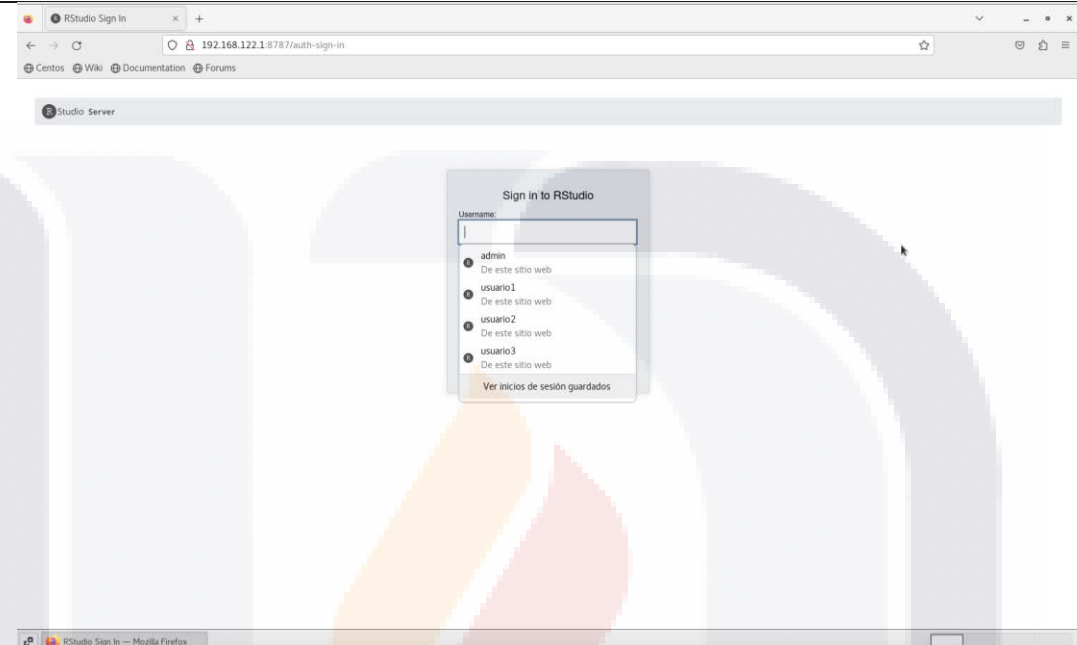
Construtto	Total desacuerdo	Desacuerdo	Parcialmente Acuerdo-Desacuerdo	Acuerdo	Total Acuerdo
COMPATIBILIDAD	1	2	3	4	5
CO.1 Utilizar la Plataforma de Servicios de TI [OSWebCienciaDatosPymes] para realizar la [Análisis de Datos en mi Organización] sería compatible culturalmente con todos los aspectos de mi trabajo.					
CO.2 Utilizar la Plataforma de Servicios de TI [OSWebCienciaDatosPymes] para realizar la [Análisis de Datos en mi Organización] encajaría con mi estilo de trabajo.					
CO.3 Utilizar la Plataforma de Servicios de TI [OSWebCienciaDatosPymes] para realizar la [Análisis de Datos en mi Organización] encajaría muy bien con la manera que me gusta analizar datos.					

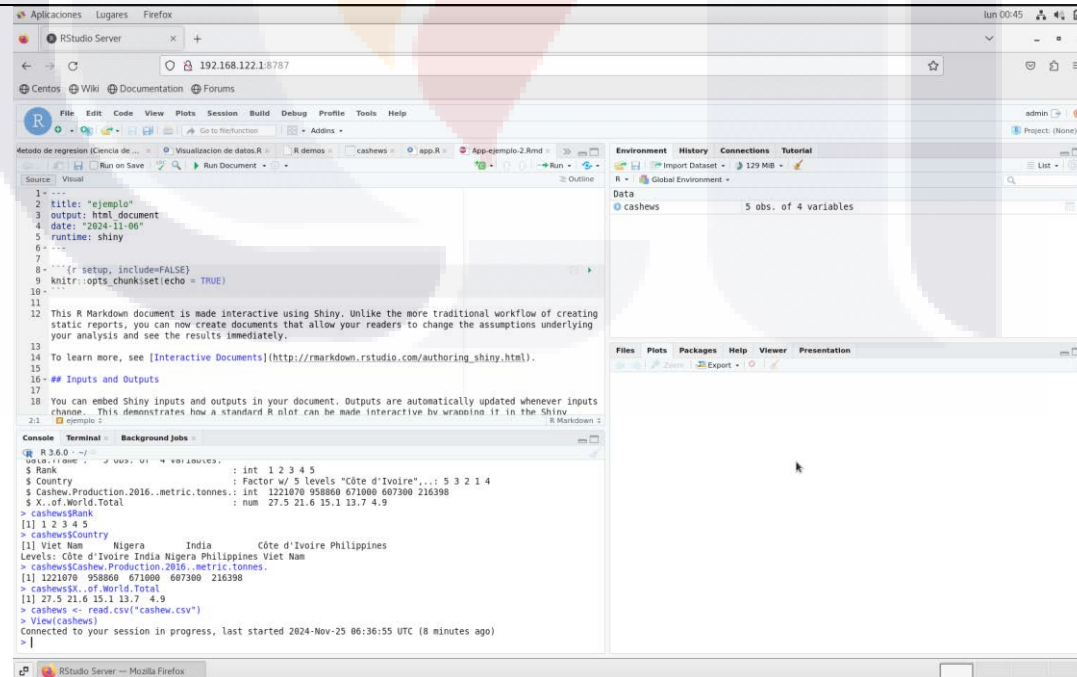
Constructo	Muy Bajo	Bajo	Moderado	Alto	Muy Alto
VALOR	1	2	3	4	5
VA.01 El valor por ahorrar dinero utilizando la Plataforma de Servicios de TI [OSWebCienciaDatosPymes] para la [Análisis de Datos en mi Organización] es:					
VA.02 El valor por ahorrar tiempo valioso utilizando la Plataforma de Servicios de TI [OSWebCienciaDatosPymes] para la [Análisis de Datos en mi Organización] es:					
VA.03. El valor para poder localizar información al usar la Plataforma de Servicios de TI [OSWebCienciaDatosPymes] para la [Análisis de Datos en mi Organización] es:					
VA.04 En general, el valor de uso de la Plataforma de Servicios de TI [OSWebCienciaDatosPymes] para la [Análisis de Datos en mi Organización] es:					

Anexo 2: Instrumento conceptual de métricas de usabilidad de tecnologías de información.

Apéndices

RSTUDIO SERVER COMMUNITY	
CAPA	APPSW-DATA
VERSIÓN	2023.06.0

EVIDENCIA 1

EVIDENCIA 2

Apéndice 1: Evidencia de funcionamiento del Rstudio Server instalado en el MVITS.

R PARA CENTOS 7

CAPA

VERSION

APPSW-DATA

3.6.0 (2019-04-26)

EVIDENCIA 1

The screenshot shows a CentOS 7 desktop environment. A terminal window titled 'admin@localhost' is open, displaying the output of the 'R' command. The output shows the R version (3.6.0) and the installation path. The desktop background is the standard CentOS 7 logo.

EVIDENCIA 2

The screenshot shows the same CentOS 7 desktop environment. The terminal window now displays R code for creating a plot. To the right of the terminal, an R plot window titled 'R Graphics: Device 2 (ACTIVE)' is open, showing a line plot with points. The plot has a title 'Simple Use of Color In a Plot' and a subtitle 'Just a Whisper of a Label'. The plot shows a series of points connected by lines, with a color gradient from blue to red.

Apéndice 2: Evidencia de funcionamiento del paquete R para CentOS 7.

CENTOS 7	
CAPA	SYSSW-DBMS
VERSIÓN	64 bits
EVIDENCIA 1	

The screenshot displays the CentOS 7 desktop environment. A window titled 'Acerca de' (About) is open, showing system information:

- Nombre del dispositivo: localhost.localdomain
- Memoria: 3.7 GB
- Procesador: 11th Gen Intel® Core™ i5-11400H @ 2.70GHz x 2
- Gráficos: i965pipe (LLVM 7.0, 128 bits)
- GNOME: Versión 3.28.2
- Tipo de SO: 64 bits
- Virtualización: KVM
- Disco: 29.0 GB

 The desktop background is dark blue with the CentOS logo and the number '7' visible. The taskbar at the bottom shows 'Aplicaciones', 'Lugares', and 'Configuración'.

Apéndice 3: Evidencia de instalación del sistema operativo CentOS 7 en la máquina virtual.

8.0.22 Community (64 bits)

MySQL Workbench

File Edit View Database Tools Scripting Help

Welcome to MySQL Workbench

MySQL Workbench is the official graphical user interface (GUI) tool for MySQL. It allows you to design, create and browse your database schemas, work with database objects and insert data as well as design and run SQL queries to work with stored data. You can also migrate schemas and data from other database vendors to your MySQL database.

[Browse Documentation](#) [Discuss on the Forums >](#)

MySQL Connections @@

- Local instance 3306 (auto sav...)
 - ⌘ root
 - ⌘ localhost 3306

MySQL Workbench 8.0
Version 8.0.22 build 107600 CE (64 bits) Community

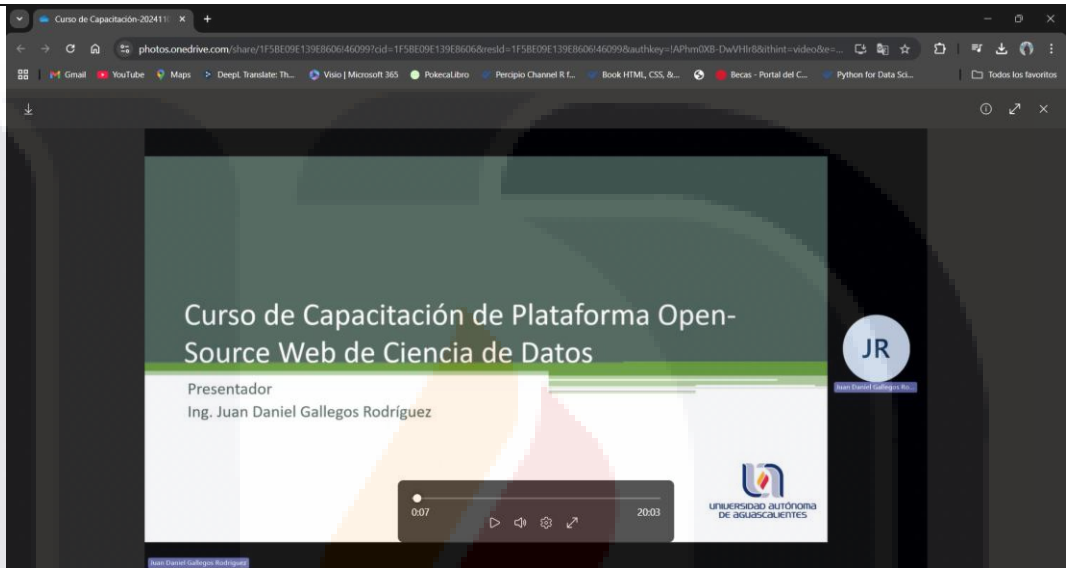
Copyright © 2008, 2020 Oracle and/or its affiliates. All Rights Reserved.
Oracle is a registered trademark of Oracle Corporation and/or its affiliates.
Other names may be trademarks of their respective owners.

ORACLE

Filter connections

Ready

105

CURSO DE CAPACITACIÓN	
OBJETIVO	Instruir al usuario final sobre el lenguaje de programación R, además de mostrar el potencial de la plataforma de ciencia de datos.
ENLACE	Curso de Capacitación-20241108_085057-Grabación de la reunión.mp4 - OneDrive
EVIDENCIA 1	
	

Apéndice 5: Evidencia del curso de capacitación destinado para usuarios finales.