Informe de análisis
Compilatio Magister+ | ZEGEL-PE

Grupo_01_PIVDE_VS1_LUN26ENE2026

ID : b96e64a115c08b5445ba068053d635b16bb4a5f2

2%
Textos sospechosos

Nombre del fichero : Grupo_01_PIVDE_VS1_LUN26ENE2026.txt
Tamaño del archivo original : 2,66 MB
Número de palabras : 10.540

Depositante : Jerson Sanchez Ramos
Fecha de depósito : 18 de septiembre de 2026
Tipo de carga : interface
fecha de fin de análisis : 18 de septiembre de 2026

Resumen (sección 1/3)

Localización de los textos sospechosos en el documento :



Incluido en el porcentaje de textos sospechosos :

Similitudes 2%

Sintáctica 1% | Semántica <1%

Pasajes con similitudes a fuentes encontradas en diferentes colecciones.



Detección de IA 37%

Textos estilísticamente próximos a un texto generado por una IA.
Este índice es un indicador y no una prueba. Comprueba con el autor si domina los conocimientos mencionados en el documento.



Idiomas no reconocidos 0%

Pasajes en los que parte del vocabulario utilizado no forma parte del diccionario de la lengua.
Puede tratarse de un intento del autor de modificar el texto para evitar ser detectado.

No incluido en el porcentaje de textos sospechosos :

Textos entre comillas <1%

Pasajes entre comillas, a menudo indicativos de una cita.

Similitudes

2%

Pasajes con similitudes a fuentes encontradas en diferentes colecciones.



Fuente principal detectada

N°	Descripciones	Similitudes	Ubicaciones
1	E6 Tasilla Sistema Web para mejorar la... #5a7e60 Detección semántica Viene de de mi grupo	<1%	
2	CRAI (1) #253e03 Detección semántica Viene de de mi grupo	<1%	
3	Documento de otro usuario #ea52e1 Detección semántica Viene de de otro grupo	<1%	
5	Las bibliotecas de La Rioja apuestan por AbsysClou... Detección semántica www.baratz.es/las-bibliotecas-de-la-rioja-apuestan-por-absyscloud-para-...	<1%	

Fuente con similitudes fortuitas

N°	Descripciones	Similitudes	Ubicaciones
4	Documento de otro usuario #7a1e4f Detección semántica Viene de de otro grupo	<1%	
6	repositorio.idat.edu.pe Detección semántica repositorio.idat.edu.pe/bitstream/handle/20.500.14364/58/Desarrollo%20...	<1%	
7	Documento de otro usuario #4615e2 Detección semántica Viene de de otro grupo	<1%	
8	repositorio.puce.edu.ec Detección semántica repositorio.puce.edu.ec/bitstreams/b896d6d3-8036-4e9a-bb0e-445915e72...	<1%	
9	Documento de otro usuario #1a33b8 Detección semántica Viene de de otro grupo	<1%	

Fuente mencionada (sin similitudes detectadas)

N°	Descripciones
1	https://maps.app.goo.gl/kBVUdU39VquLSLk68
2	https://www.gob.pe/minedu
3	https://scrumguides.org/
4	https://docs.streamlit.io/

1

INSTITUTO DE EDUCACIÓN SUPERIOR PRIVADO IDAT

PROGRAMA DE ESTUDIOS EN DESARROLLO DE SISTEMAS DE INFORMACIÓN

PLATAFORMA DE INTEGRACIÓN Y VISUALIZACIÓN DE DATOS EDUCATIVOS PROVENIENTES DE MÚLTIPLES FUENTES DE UNA ORGANIZACIÓN PÚBLICA PERÚ-EDUCACIÓN DEL SECTOR EDUCACIÓN (2019–2025)

Trabajo de aplicación profesional para obtener el título de Desarrollo de Sistemas de Información

DAVID RICARDO HUAMANI LOAYZA (0009-0002-4056-4903)

Lima – Perú

Año 2026

2

2

Dedico este trabajo con gratitud y aprecio a toda mi familia, por su apoyo incondicional,

comprensión y motivación constante durante todo el proceso de formación.

3 **11 12 13** 13 *13 14 14 17 19 21 22* 24 *25 26 26* 29 *29 30 31 32 33*

Índice General

Pág.

Resumen Ejecutivo.....

Introducción .....

Capítulo I: Análisis del Negocio .....

Descripción de la Organización .....

Antecedentes .....

Visión Organización .....

Misión Organización .....

Organigrama y Funciones de las Principales Áreas .....

Mapa de Procesos y Proceso(s) a Automatizar.....

Análisis FODA .....

Elicitación de Requisitos .....

Identificación de la Necesidad (Problema) .....

Análisis del Problema .....

Propuesta de Solución.....

Alternativas de Solución .....

Factibilidad del Proyecto .....

Factibilidad Técnica.....

Recursos Tecnológicos .....

Factibilidad Operativa.....

Recursos Humanos .....

Factibilidad Económica .....

4 *35 38* **40** 40 *40 40* 41 *42* 43 *43 44 45 45* 46 46 46 47 47 48 **49** 49 *49 51* 54

Análisis Costo-Beneficio .....

Acta de Constitución del Proyecto .....

Capítulo II: Planificación del Proyecto .....

Enfoque de Desarrollo y del Ciclo de Vida .....

Enfoque a aplicar .....

Justificación del enfoque a aplicar .....

Arquitectura de Software a Utilizar .....

Modelos, métodos y artefactos a aplicar.....

Planificación del Proyecto .....

Plan del Proyecto .....

Declaración de Alcance del Proyecto .....

Objetivos del Proyecto (General y Específicos) .....

Beneficios del Proyecto .....

Cronograma.....

Interesados .....

Supuestos .....

Factores Críticos de Éxito (FCE) .....

Riesgos .....

Comunicaciones .....

Capítulo III: Ejecución del Proyecto.....

Fase Inicio .....

Presentación del Lanzamiento del Proyecto (Kick Off) .....

Requerimientos/Historias de Usuarios (HU) .....

Fase Planificación .....

5 *54 55 56 57 58* 59 *59 60 60* 61 *61 62 62 63 64* 64 *64* **65** 65 68 70 72 74 77

Historias de Usuarios Priorizados .....

Estimación de Tareas de las HU .....

Product Backlog Priorizado y Estimado .....

Cronograma/Sprint Planning .....

Descomposición de Tareas de las HU (Plantilla Scrum) .....

Fase de Implementación .....

Incrementos por Sprint.....

Daily Scrum Meeting .....

Desarrollo del Sprint Backlog .....

Fase de Revisión y Retrospectiva .....

Scrum Board .....

Sprint Review .....

Burn Down Chart.....

Impediment Log .....

Retrospectiva del Sprint .....

Fase de Lanzamiento .....

Acta de Conformidad de Entregables .....

Capítulo IV: Programación .....

Implementación de la Arquitectura de Software .....

Configuración del Entorno de Desarrollo y Dependencias .....

Implementación de la Base de Datos .....

Procesos de Integración y Carga de Datos (ETL) .....

Gestión de Conexiones y Fuentes de Datos Institucionales .....

Implementación del Backend de la Plataforma .....

6 81 85 **89** 89 89 90 94 99 99 **101 102 103 106** 106 108 110 111 113

Implementación del Frontend – Dashboard .....

Consultas, Reportes e Indicadores .....

Capítulo V: Pruebas de Calidad de Software.....

Objetivo de las Pruebas de Calidad .....

Estrategia de Pruebas Aplicada .....

Pruebas Unitarias .....

Pruebas Integrales .....

Registro de Incidencias y Correcciones .....

Validación por el Usuario .....

Conclusiones .....

Recomendaciones .....

Referencias Bibliográficas .....

Anexos .....

Anexo 01. Arquitectura general de la plataforma PIVDE .....

Anexo 02. Modelo de la base de datos perueduca4 .....

Anexo 03. Script de integración y carga de datos (ETL) .....

Anexo 04. Configuración del entorno de desarrollo .....

Anexo 05. Capturas de funcionamiento de la plataforma PIVDE .....

7 18 21 22 23 25 27 30 32 33 34 34 35 36 36 38 46 47 52 53

Índice de Tablas

Pág.

Tabla 1. Funciones de la organización .....

Tabla 2. Análisis FODA de la organización pública PERÚ-EDUCACIÓN .....

Tabla 3. Requisitos Funcionales del Sistema .....

Tabla 4. Requisitos No Funcionales del Sistema .....

Tabla 5. Técnica de Análisis de Causa Raíz: Los 5 Porqué .....

Tabla 6. Alternativas de solución.....

Tabla 7. Recursos tecnológicos disponibles y requeridos .....

Tabla 8. Recursos humanos involucrados .....

Tabla 9. Costo de recursos humanos.....

Tabla 10. Costo de recursos tecnológicos .....

Tabla 11. Costo total del proyecto .....

Tabla 12. Costo mensual del proceso sin sistema .....

Tabla 13. Costo mensual del proceso con sistema .....

Tabla 14. Ahorro anual .....

Tabla 15. Acta de Constitución del Proyecto.....

Tabla 16. Interesados .....

Tabla 17. Matriz de Riesgos del Proyecto .....

Tabla 18. Historias de Usuario.....

Tabla 19. Product Backlog Inicial .....

8 54 55 56 58 61 68 75

Tabla 20. Historias Priorizadas (MoSCoW) .....

Tabla 21. Estimación de Historias .....

Tabla 22. Product Backlog Priorizado .....

Tabla 23. Descomposición de actividades del proyecto (alineada a Scrum) .....

Tabla 24. Scrum Board del Sprint .....

Tabla 25. Principales librerías utilizadas en el proyecto .....

Tabla 26. Fuentes de datos integradas en la plataforma PIVDE .....

9 14 15 16 17 19 20 50 63 67 69 69 71 73 74 78 79 80 82 83

Índice de Figuras

Pág.

Figura 1. Logo de la organización .....

Figura 2. Ubicación geográfica .....

Figura 3. Frontis de la entrada al Ministerio de Educación .....

Figura 4. Organigrama de la organización .....

Figura 5. Mapa de procesos de la organización pública PERÚ-EDUCACIÓN .....

Figura 6. Proceso actual (As-Is) de consolidación y generación de reportes educativos .....

Figura 7. Presentación del Lanzamiento del Proyecto (Kick Off) .....

Figura 8. Burn Down Chart – Sprint 2 (Visualización y Análisis) .....

Figura 9. Arquitectura general de la plataforma PIVDE .....

Figura 10. Archivo requirements del Backend .....

Figura 11. Archivo requirements del Frontend .....

Figura 12. Estructura de la base de datos perueduca4 .....

Figura 13. Script ETL .....

Figura 14. Flujo ETL .....

Figura 15. Estructura básica del backend .....

Figura 16. Muestra del archivo main.py .....

Figura 17. Muestra del archivo db.py .....

Figura 18. Estructura básica del Frontend .....

Figura 19. Muestra del archivo app.py .....

10 84 85 86 87 88 91 92 93 95 96 97 98 107 109 110 111 112 112 113 114 115

Figura 20. Muestra del archivo cursos.py .....

Figura 21. Modelo de consulta de cursos .....

Figura 22. Informe gráfico de conexiones por día .....

Figura 23. Consulta de matrículas por usuarios en todos los campus .....

Figura 24. Reporte exportado en formato Excel .....

Figura 25. Ejecución del proceso ETL Cargar conexiones diarias .....

Figura 26. Consulta SQL validada .....

Figura 27. Generación de gráfico unitario .....

Figura 28. Dashboard principal .....

Figura 29. Aplicación de filtros .....

Figura 30. Muestra de catálogo de cursos .....

Figura 31. Exportación de reporte .....

Figura 32. Arquitectura general de la plataforma PIVDE .....

Figura 33. Esquema de la base de datos perueduca4 .....

Figura 34. Ejecución del proceso ETL MongoDB – MySQL .....

Figura 35. Archivo .env (sin credenciales reales) .....

Figura 36. Archivo requirements.txt del Backend .....

Figura 37. Archivo requirements.txt del Frontend .....

Figura 38. Pantalla de acceso .....

Figura 39. Ejecución del Backend – API de Integración de Datos (FastAPI / Uvicorn) .....

Figura 40. Ejecución del Frontend – Dashboard en Streamlit .....

Resumen Ejecutivo

La organización pública PERÚ-EDUCACIÓN, perteneciente al sector educación, tiene como finalidad la gestión, administración y ejecución de programas educativos orientados al desarrollo integral de la comunidad educativa a nivel nacional. En el marco del proceso de gestión de matrícula, se identificó una problemática relacionada con la administración y el tratamiento de información proveniente de diversas plataformas digitales y bases de datos internas que operaban de manera aislada. Esta falta de integración generaba retrasos en la elaboración de reportes, duplicidad de registros e inconsistencias en los datos, lo que dificultaba la disponibilidad de información confiable y oportuna para la toma de decisiones. Dicha situación afectaba directamente al personal encargado de la gestión de matrícula, la atención en mesa de ayuda y la consulta de información educativa, quienes requerían acceder a datos actualizados y precisos en tiempo real. Ante este escenario, se diseñó e implementó una solución tecnológica basada en una plataforma de integración y visualización de datos educativos, orientada a consolidar, depurar y presentar de manera dinámica la información generada entre los años 2019 y 2025. La implementación de esta solución permitió optimizar el acceso y análisis de la información, reducir los tiempos de respuesta y fortalecer la transparencia institucional, contribuyendo a una mayor eficiencia operativa y a la mejora continua de la gestión educativa.

Palabras clave: Gestión de matrícula, integración de datos, sistemas de información educativa,

visualización de datos, toma de decisiones, gestión educativa.

12

Introducción

El presente documento presenta lo que fue el desarrollo de una solución tecnológica orientada a la integración, procesamiento y visualización de datos educativos provenientes de diversas fuentes institucionales. El sistema fue desarrollado empleando el lenguaje de programación Python, en conjunto con el framework Streamlit. Asimismo, se utilizaron librerías especializadas como Pandas y NumPy para la manipulación y análisis de grandes volúmenes de datos, mientras que Plotly y Matplotlib. Para la conexión, consulta e integración con fuentes de datos externas, se empleó SQLAlchemy, lo que garantizó una comunicación eficiente y segura entre sistemas heterogéneos.

El contenido del presente documento se encuentra estructurado en cinco capítulos. En el Capítulo I se aborda el análisis del negocio, presentando la información institucional y organizacional, la identificación de necesidades, el análisis FODA y la propuesta de solución. En el Capítulo II se describe la planificación del proyecto, determinando los objetivos, el alcance, el cronograma y los riesgos. En el Capítulo III se expone la ejecución del proyecto, detallando las fases de desarrollo, la gestión de historias de usuarios, las pruebas y el control de calidad.

En el Capítulo IV se explica la programación del sistema, incluyendo la implementación de la arquitectura de software definida, la creación de la base de datos, el desarrollo del backend, frontend, reportes y mantenedores. Finalmente, en el Capítulo V se presentan las pruebas de calidad de software, junto con las conclusiones, recomendaciones y los anexos.

13

Capítulo I

Análisis del Negocio

Descripción de la Organización

La organización pública PERÚ-EDUCACIÓN del sector educación en la que se ejecutó el proyecto se dedica a la gestión, promoción y desarrollo de programas de formación y capacitación en el ámbito educativo nacional. Su principal objetivo es fortalecer las competencias de los actores del sistema educativo mediante la oferta de servicios, recursos y cursos virtuales orientados a la mejora continua de la enseñanza y el

aprendizaje.

NOTA: El nombre de la organización pública PERÚ-EDUCACIÓN es ficticio, sin embargo, el producto del proyecto sí es real.

Antecedentes

Durante los años 2019 al 2025, la organización pública PERÚ-EDUCACIÓN implementó diversas plataformas tecnológicas para administrar cursos virtuales, gestionar registros académicos y monitorear la participación de los usuarios. Sin embargo, la información generada permaneció distribuida en distintos sistemas, sin un repositorio centralizado ni mecanismos de análisis unificado, lo que dificulta la generación oportuna de reportes y la toma de decisiones basada en datos.

Visión Organización

Ser una organización pública PERÚ-EDUCACIÓN líder en la gestión educativa digital, reconocida por la calidad, accesibilidad y transparencia en el uso de la información para la mejora del sistema educativo nacional (Ministerio de Educación del Perú, 2024).

Misión Organización

Promover el desarrollo de la educación mediante la gestión eficiente de recursos digitales, la integración tecnológica y la disponibilidad de información educativa confiable que fortalezca las decisiones estratégicas y operativas del sector (Ministerio de Educación del Perú, 2024).

Figura 1.

Logo de la organización

Figura 2.

Ubicación geográfica

Calle del Comercio N° 193, distrito de San Borja, Lima, Perú

Nota: Google Maps 2025 [Captura de pantalla de Google Maps]. Recuperado de:

https://maps.app.goo.gl/kBVUdU39VquLSLk68

16

Figura 3.

Frontis de la entrada al Ministerio de Educación

17

Organigrama y Funciones de las Principales Áreas

Figura 4.

Organigrama de la organización

Nota: Basada en la estructura organizacional típica de una entidad pública dedicada a la gestión educativa.

Director de la dirección

Coordinaciones Legales

Abogado

Asesores

Administración

Coordinador

Especialista

Especialistas Logísticos

Especialista de Control

Planeamiento y Presupuesto

Coordinador

Especialistas

PerúEduca

Coordinador

Especialistas

Mesa de ayuda

Proyectos Educativos

Coordinador

Especialistas

Innovación e Investigación Coordinador

Asistente

Recepción

18

Tabla 1.

Funciones de la organización

Área / Unidad Funciones principales

Dirección General Dirige, planifica y supervisa las acciones estratégicas orientadas a la gestión educativa digital y la innovación tecnológica. Representa a la organización pública PERÚ-EDUCACIÓN y vela por el cumplimiento de los objetivos institucionales.

Asistente de Dirección Apoya en la gestión administrativa y técnica de la dirección, coordina reuniones, redacta informes y controla el cumplimiento de actividades asignadas.

Recepción Atiende a usuarios internos y externos, canaliza la documentación, gestiona correspondencia y coordina la agenda institucional.

Coordinaciones Legales

Asesora en aspectos normativos, revisa convenios, contratos y garantiza el cumplimiento del marco legal vigente aplicable a los proyectos institucionales.

Administración Gestiona los recursos humanos, financieros y logísticos requeridos para el desarrollo de proyectos.

Supervisa el control presupuestal y la ejecución de gastos.

Planeamiento y Presupuesto

Elabora, evalúa y monitorea los planes operativos institucionales y presupuestos, alineándolos con los objetivos estratégicos de la entidad.

PerúEduca / Plataforma Educativa

Administra la principal plataforma digital educativa, asegurando la gestión de usuarios, la publicación de contenidos y el soporte técnico y pedagógico.

Mesa de Ayuda Brinda atención y soporte a los usuarios de las plataformas institucionales, registrando incidencias y garantizando la continuidad de los servicios digitales.

Proyectos Educativos Diseña, ejecuta y evalúa proyectos tecnológicos y pedagógicos que impulsen la mejora continua de la gestión educativa.

Innovación e Investigación

Promueve la creación de soluciones digitales, la investigación aplicada y la adopción de nuevas tecnologías en el ámbito educativo.

Especialistas Técnicos y Logísticos

Ejecutan tareas de desarrollo, mantenimiento, análisis y soporte de plataformas y sistemas institucionales, asegurando su operatividad y actualización.

Mapa de Procesos y Proceso(s) a Automatizar

Figura 5.

Mapa de procesos de la organización pública PERÚ-EDUCACIÓN

Fuente: Resolución de secretaria general N.º 075-2021-MINEDU

Figura 6.

Proceso actual (As-Is) de consolidación y generación de reportes educativos

Análisis FODA

Tabla 2.

Análisis FODA de la organización pública PERÚ-EDUCACIÓN

FORTALEZA DEBILIDAD

- Respaldo institucional del MINEDU para impulsar la transformación digital.
- Personal técnico con experiencia en gestión de plataformas educativas.
- Infraestructura tecnológica existente y bases de datos consolidadas.
- Capacidad de articulación con otras unidades y actores del sistema educativo.
- Dispersión de datos en múltiples sistemas sin integración efectiva.
- Procesos manuales de consolidación y análisis de información.
- Falta de estandarización en reportes e indicadores educativos.
- Limitado uso de herramientas avanzadas de analítica de datos.

OPORTUNIDAD AMENAZA

- Políticas de gobierno digital y educación basada en evidencias.
- Acceso a tecnologías de integración y visualización de bajo costo.
- Posibles alianzas con universidades y organismos internacionales.
- Mayor demanda de transparencia y generación de indicadores educativos.
- Cambios presupuestales o normativos que afecten la continuidad.

● Riesgos en la seguridad y protección de datos personales.

● Resistencia al cambio en la adopción de nuevas herramientas.

● Avances tecnológicos que exijan constante actualización.

22

Elicitación de Requisitos

Requisitos Funcionales:

La Tabla 3 presenta los requisitos funcionales identificados para la

plataforma de integración y visualización de datos educativos (PIVDE).

Tabla 3.

Requisitos Funcionales del Sistema

Código Nombre del Requisito

Descripción Actor

RF01 Integración de fuentes de datos

El sistema debe permitir la integración y consolidación de información proveniente de múltiples fuentes institucionales, tales como bases de datos relacionales, archivos CSV y Excel.

Analista de Datos

RF02 Carga de datos El sistema debe permitir la carga manual y automática de nuevos conjuntos de datos desde fuentes internas autorizadas.

Analista de Datos

RF03 Actualización de información

El sistema debe permitir la actualización periódica de los datos integrados, garantizando que la información esté disponible de forma oportuna.

Analista de Datos

RF04 Limpieza y depuración de datos

El sistema debe ejecutar procesos automáticos para eliminar registros duplicados, inconsistentes o incompletos antes del análisis.

Analista de Datos

RF05 Visualización de dashboards

El sistema debe mostrar dashboards interactivos con gráficos, indicadores y tendencias educativas.

Usuario Institucional

23

RF06 Filtros dinámicos

El sistema debe permitir aplicar filtros por año, curso, región u otros criterios definidos por el usuario.

Usuario Institucional

RF07 Generación de reportes

El sistema debe permitir la exportación de reportes y dashboards en formatos PDF, Excel y CSV.

Usuario Institucional

RF08 Gestión de usuarios

El sistema debe permitir gestionar perfiles de usuario con distintos niveles de acceso (administrativo y analítico).

Administrador

RF09 Registro de actividades

El sistema debe registrar las actividades realizadas por los usuarios, tales como cargas de datos y consultas.

Administrador

RF10 Panel principal de indicadores

El sistema debe mostrar un panel principal con indicadores clave resumidos para la toma de decisiones.

Gestor Educativo

Requisitos No Funcionales:

La Tabla 4 presenta los requisitos no funcionales, los cuales establecen las

condiciones de calidad, desempeño y operación del sistema.

Tabla 4.

Requisitos No Funcionales del Sistema

Código Categoría Descripción

RNF01 Usabilidad El sistema debe contar con una interfaz intuitiva y de fácil uso, accesible para usuarios con conocimientos básicos en herramientas digitales.

RNF02 Rendimiento El sistema debe procesar y mostrar la información integrada en tiempos de respuesta aceptables, incluso con grandes volúmenes de datos.

RNF03 Seguridad El sistema debe restringir el acceso a la información mediante autenticación y control de permisos por tipo de usuario.

RNF04 Disponibilidad El sistema debe estar disponible durante el horario laboral institucional, garantizando continuidad operativa.

RNF05 Escalabilidad El sistema debe permitir la incorporación de nuevas fuentes de datos y módulos sin afectar su funcionamiento actual.

RNF06 Mantenibilidad El sistema debe permitir la actualización y mejora de sus módulos de forma modular y controlada.

RNF07 Compatibilidad El sistema debe ser compatible con los principales navegadores web utilizados en la institución.

RNF08 Portabilidad El sistema debe poder desplegarse en servidores institucionales o entornos en la nube.

RNF09 Confiabilidad El sistema debe garantizar la integridad y consistencia de los datos procesados y visualizados.

RNF10 Cumplimiento normativo

El sistema debe respetar las políticas institucionales y normativas vigentes relacionadas con el manejo de datos.

Identificación de la Necesidad (Problema)

La organización pública PERÚ-EDUCACIÓN enfrenta la necesidad de integrar

datos educativos generados en múltiples plataformas digitales, con el fin de reducir los

tiempos de búsqueda, mejorar la precisión de los informes y contar con información

actualizada para la toma de decisiones.

25

Análisis del Problema

Para identificar la causa raíz de la demora en la generación de reportes institucionales y la

falta de integración de la información, se utilizó la técnica de los 5 Porqués, que permite analizar

de manera progresiva las razones que originan un problema específico.

Tabla 5.

Técnica de Análisis de Causa Raíz: Los 5 Porqué

¿Por qué? Respuesta

¿Por qué se demora la generación de reportes institucionales?

Porque la información se encuentra dispersa en diferentes plataformas y sistemas independientes.

¿Por qué la información está dispersa en múltiples sistemas?

Porque no existe un mecanismo centralizado que integre automáticamente los datos provenientes de las distintas áreas.

¿Por qué no existe un mecanismo centralizado de integración de datos?

Porque cada área realiza sus propias descargas de manera manual y consolida la información utilizando hojas de cálculo, sin un sistema común.

¿Por qué se consolidan los datos manualmente utilizando hojas de cálculo?

Porque no se han establecido procesos estandarizados ni herramientas compartidas para la administración, depuración e integración de datos institucionales.

¿Por qué no se han estandarizado los procesos ni implementado herramientas especializadas?

Porque en años anteriores no se priorizó el desarrollo de una solución tecnológica integrada para la gestión eficiente de la información institucional.

26

El análisis evidencia que la causa raíz del problema es la ausencia de una plataforma

centralizada y de procesos estandarizados para la gestión de datos, lo que obliga al personal a

realizar tareas manuales y lentas que afectan la oportunidad y precisión de los reportes institucionales.

Propuesta de Solución

Diseñar e implementar una plataforma de integración y visualización de datos educativos que unifique la información institucional, facilite su análisis y mejore la disponibilidad de datos en tiempo real.

Alternativas de Solución

Con el propósito de resolver la problemática relacionada con la dispersión de información, los procesos manuales de consolidación y la demora en la generación de reportes institucionales, se evaluaron diversas alternativas tecnológicas. Cada alternativa fue analizada considerando su costo estimado, tiempo de implementación, riesgos asociados y capacidad para resolver la causa raíz identificada.

27

Tabla 6.

Alternativas de solución

Alternativa Descripción Costo estimado

Tiempo de impleme ntación

Riesgo estimado

Conclusión

1. Continuar con el proceso manual mejorado (Excel + procedimie ntos internos)

Implementar mejoras menores en los procesos actuales utilizando hojas de cálculo, formatos estandarizado s y reglas de validación básicas para organizar la consolidación de datos.

S/. 12,000 (Comprende capacitacione s, supervisión y ajustes de procesos internos)

1 mes Alto • Alta probabilidad de errores humanos

• No soluciona la duplicidad de datos

• Depende del personal que consolida la información

• No mejora significativamente la rapidez ni la calidad de los reportes

2. Implement ar una herramient a de Business Intelligence (BI) comercial

Adquirir una solución empresarial como Power BI Pro, Tableau, Qlik u otra plataforma con conectores, dashboards automatizado s y soporte corporativo.

S/. 80,000 – S/. 120,000 (Licencias anuales, conectores, infraestructur a, capacitación y soporte)

2 – 3 meses

Medio • Costos recurrentes por licencias

• Dependencia de un proveedor externo

• Integración limitada con sistemas educativos públicos

• Requiere personal capacitado permanentemente

28

3. Desarrollo de una Plataforma Propia de Integración y Visualizaci ón de Datos

Construcción de una plataforma web desarrollada con Python, Streamlit y librerías de visualización, capaz de integrar diversas fuentes de datos institucionale s (PerúEduca, Mesa de Ayuda, encuestas, archivos administrativ os, etc.) y generar reportes consolidados en tiempo real.

S/. 31,320 (Mano de obra del equipo Scrum + costos tecnológicos)

4 meses Medio - Bajo

• Depende del equipo de desarrollo

• Permite total control y personalización

• No tiene costos de licenciamiento

• Se adapta a las políticas de datos del MINEDU

29

Factibilidad del Proyecto

La factibilidad del proyecto se evaluó considerando tres componentes principales:

factibilidad técnica, factibilidad operativa y factibilidad económica. Esta evaluación permite

determinar si el desarrollo e implementación de la plataforma de integración y visualización de

datos es viable en términos de recursos tecnológicos, capacidades organizacionales y presupuesto

institucional.

Factibilidad Técnica

La factibilidad técnica analiza si la organización cuenta con la infraestructura tecnológica

y las herramientas necesarias para la implementación del sistema. PERÚ-Educación dispone de

equipos informáticos, conectividad estable y personal con conocimientos en herramientas

digitales, lo que permite implementar la solución propuesta.

El proyecto se desarrollará utilizando tecnologías de código abierto como Python,

Streamlit, Pandas, NumPy, Plotly, Matplotlib y SQLAlchemy, que no requieren inversión en

licenciamiento. Asimismo, se utilizará un gestor de base de datos libre (PostgreSQL o MySQL),

lo cual reduce significativamente los costos.

30

Recursos Tecnológicos

Tabla 7.

Recursos tecnológicos disponibles y requeridos

Tipo de Recurso Descripción Cantidad Observaciones

Hardware Laptop o PC para desarrollo (Intel i5/i7, 16 GB RAM, SSD 512 GB)

2 Disponible en la institución

Hardware Servidor para despliegue (físico o en la nube) 8 GB RAM, 4 vCPU

1 Puede usarse servidor institucional

Software Python 3.x, Streamlit, Pandas, NumPy, Plotly, Matplotlib

- Software libre

Software PostgreSQL / MySQL 1 Software libre

Software Herramientas de control de versiones (GitHub/GitLab)

- Software libre

Internet Conectividad ≥ 50 Mbps 1 Contemplada en servicios institucionales

31

Factibilidad Operativa

La factibilidad operativa evalúa si el personal, los procesos y la cultura organizacional permiten implementar con éxito la solución. Actualmente, PERÚ-Educación presenta dificultades para integrar datos provenientes de distintas plataformas, lo que genera duplicidad, errores y tiempos de respuesta elevados.

Durante el levantamiento de información, las áreas involucradas (Plataforma PerúEduca, Mesa de Ayuda, Innovación Educativa, Proyectos Especiales y Analítica) mostraron una disposición favorable para adoptar una herramienta centralizada que facilite la gestión de datos institucionales.

La plataforma propuesta será usada por personal administrativo, analistas, responsables de mesa de ayuda y gestores de información educativa. La interfaz basada en Streamlit es amigable y de fácil adopción, reduciendo la curva de aprendizaje.

32

Recursos Humanos

Tabla 8.

Recursos humanos involucrados

Rol Función Participación

Product Owner Define necesidades institucionales y valida entregables

Representante DITE

Scrum Master Facilita la metodología ágil y elimina impedimentos

Equipo del proyecto

Desarrollador Full Stack Construcción del backend, frontend e integración

Equipo del proyecto

Analista de Datos Validación, limpieza y estandarización de datos

Personal de la institución

Usuario Clave Pruebas, retroalimentación y validación funcional

Personal administrativo

33

Factibilidad Económica

La factibilidad económica determina si la organización puede asumir los costos del proyecto y si la inversión es justificable frente a los beneficios obtenidos.

Se consideraron los gastos en recursos humanos (equipo Scrum), infraestructura tecnológica y servicios complementarios. Los valores estimados se basan en los TAP que me compartiste, ajustados a tu proyecto.

Tabla 9.

Costo de recursos humanos

Rol Horas/mes (A)

Costo hora (B)

Subtotal mensual (C=A×B)

Meses (D)

Total (E=C×D)

Scrum Master 20 45 S/. 900 4 S/. 3,600

Desarrollador Full Stack 105 55 S/. 5,775 4 S/. 23,100

Analista de Datos 40 45 S/. 1,800 4 S/. 7,200

Tester / QA 25 40 S/. 1,000 4 S/. 4,000

Total RRHH 190h/mes 4 S/. 37,900

34

Tabla 10.

Costo de recursos tecnológicos

Recurso tecnológico Cantidad Costo unitario Total

Laptop para desarrollo 1 S/. 3,000 S/. 3,000

Servidor / nube (4 meses) 1 S/. 500 S/. 500

Ofimática y herramientas 1 S/. 1,700 S/. 1,700

Total tecnologías S/. **5,200**

Tabla 11.

Costo total del proyecto

Concepto Monto (S/.)

Recursos Humanos 37,900

Recursos Tecnológicos 5,200

Subtotal **43,100**

Margen (50%) 21,550

Costo de venta 64,650

IGV (18%) 11,637

Total 76,287

35

Análisis Costo-Beneficio

El análisis costo–beneficio permite determinar si la implementación de la plataforma de integración y visualización de datos genera ventajas económicas frente a los métodos actuales utilizados por la institución. Actualmente, las áreas de PERÚ-Educación realizan procesos manuales de consolidación de datos utilizando hojas de cálculo, lo cual incrementa los tiempos, genera errores y demanda recursos humanos adicionales.

Con la implementación del sistema, los procesos se automatizan, reduciendo significativamente el tiempo de procesamiento, la duplicidad de tareas y los costos operativos.

Tabla 12.

Costo mensual del proceso sin sistema

Concepto Cantidad Costo unitario (S/.)

Subtotal (S/.)

Analistas dedicados a consolidación manual de datos

2 2,000 4,000

Personal de soporte para búsquedas de información

1 800 800

Materiales y otros (impresiones, almacenamiento, reprocesos)

- 500 500

Total mensual sin sistema 5,300

36

Tabla 13.

Costo mensual del proceso con sistema

Concepto Cantidad Costo unitario (S/.)

Subtotal (S/.)

Analista encargado de explotación del sistema 1 1,000 1,000

Personal de soporte 1 400 400

Materiales y otros - 200 200

Total mensual con sistema 1,600

Tabla 14.

Ahorro anual

Concepto Monto (S/.)

Gastos anuales sin sistema (5,300 × 12) 63,600

Gastos anuales con sistema (1,600 × 12) 19,200

Ahorro anual estimado 44,400

Relación Beneficio / Costo

El costo total del proyecto, según el análisis económico previo, asciende a:

S/. 76,287

37

Cálculo:

• Beneficio anual: **S/ 44,400**

• Costo del proyecto: **S/ 76,287**

Beneficio/Costo = 44,400 / 76,287 ≈ **0.58**

Conclusión del Análisis Costo–Beneficio

Del análisis realizado se concluye que, si bien el costo total del proyecto asciende a S/. 76,287, la implementación de la plataforma PIVDE genera un ahorro anual estimado de S/. 44,400 en costos operativos asociados a procesos manuales de consolidación y análisis de datos.

Aunque la relación Beneficio/Costo (0.58) indica que la inversión no se recupera íntegramente en el primer año, el proyecto resulta económicamente viable en un horizonte de mediano plazo, considerando que la plataforma no genera costos recurrentes de licenciamiento y que los beneficios operativos, organizacionales y estratégicos se incrementan en los años posteriores.

Asimismo, el valor del proyecto no solo se justifica en términos monetarios, sino también en beneficios intangibles como la mejora en la calidad de la información, la reducción de errores, la optimización del tiempo del personal y el fortalecimiento de la toma de decisiones institucionales basada en datos.

38

Acta de Constitución del Proyecto

Tabla 15.

Acta de Constitución del Proyecto

Elemento Descripción

Nombre del Proyecto Plataforma de integración y visualización de datos educativos provenientes de múltiples fuentes para la organización pública PERÚ-Educación (Periodo 2019–2025).

Fecha de Inicio 01 de agosto de 2025

Objetivo del Proyecto Diseñar, desarrollar e implementar una plataforma web que integre datos de diferentes sistemas institucionales, permitiendo generar reportes, visualizaciones y análisis oportunos para la toma de decisiones internas.

Alcance del Proyecto • Desarrollar una plataforma en Python y Streamlit. • Integrar información proveniente de PerúEduca, Mesa de Ayuda y otros sistemas institucionales. • Implementar dashboards dinámicos y consultas filtradas. • Desplegar la plataforma en servidor institucional. • Capacitar a usuarios clave.

Duración del Proyecto 4 meses / 15 semanas

Presupuesto Estimado S/. 44,349 (incluye recursos humanos, tecnológicos y margen operativo).

Equipo del Proyecto • Product Owner: Representante DITE. • Scrum Master: Responsable de metodología ágil. • Desarrollador Full Stack: Implementación del sistema. • Analista de Datos: Validación y modelado. • Tester QA: Pruebas y control de calidad.

Cliente Organización pública PERÚ-Educación (sector educación).

39

Sitio del Proyecto Lima – Perú

Entregables Principales • Plataforma web funcional. • Base de datos operativa e integrada. • Dashboards e informes dinámicos. • Actas de validación. • Documentación técnica. • Informe final del proyecto.

40

Capítulo II

Planificación del Proyecto

Enfoque de Desarrollo y del Ciclo de Vida

Enfoque a aplicar

El enfoque seleccionado para este proyecto es el marco de trabajo Scrum, ya que permite desarrollar productos tecnológicos de manera iterativa y con retroalimentación continua. Las necesidades de PERÚ-Educación demandan una solución flexible que pueda ajustarse durante el proceso de construcción, especialmente por la variedad de fuentes de datos y el dinamismo en los requerimientos institucionales.

Justificación del enfoque a aplicar

La elección de Scrum se justifica por las siguientes razones:

• Flexibilidad ante cambios: La institución puede modificar o priorizar requerimientos a medida que avanza el proyecto.

• Entregas tempranas: Cada sprint permite entregar versiones funcionales, facilitando la validación con usuarios.

• Colaboración constante: Se fortalece la comunicación entre el equipo Scrum y los usuarios clave de PERÚ-Educación.

41

• Reducción de riesgos: La retroalimentación continua permite identificar problemas antes de etapas avanzadas.

• Orientación a valor: Scrum prioriza las funcionalidades más importantes para la toma de decisiones institucional.

Arquitectura de Software a Utilizar

El proyecto emplea una arquitectura basada en:

1. Backend (Servidor de datos)

• Lenguaje: **Python 3.x**

• Framework: **Streamlit** para desarrollo de interfaces interactivas

• Librerías principales:

o **Pandas** (procesamiento de datos)

o **NumPy** (operaciones numéricas)

o **Plotly / Matplotlib** (visualizaciones dinámicas)

o **SQLAlchemy** (gestión de base de datos)

2. Base de datos

• **PostgreSQL** o **MySQL**

• Acceso centralizado con modelo unificado

3. Frontend (Interfaz de usuario)

• Renderizado dinámico desde Streamlit

• Grafos e indicadores interactivos

• Widgets de filtros y exportación

4. Integración

• Cargas automáticas

• Conectores hacia plataformas institucionales

• Lectura de archivos externos (CSV, Excel)

Modelos, métodos y artefactos a aplicar

Los modelos y artefactos utilizados en este proyecto incluyen:

• Product Backlog

• Sprint Backlog

• Planning Poker (estimación)

• Diagrama EDT (Estructura de Desglose de Trabajo)

• Historias de Usuario

• Burndown Chart

• Matriz de Riesgos

• Actas de Sprint Review y Retrospectivas

• Diagramas de proceso (AS-IS y TO-BE)

Estos artefactos aseguran una gestión adecuada del ciclo de vida del proyecto y permiten un seguimiento transparente del avance.

Planificación del Proyecto

La planificación se realizó tomando como base el enfoque Scrum, definiendo sprints de dos semanas y priorizando las funcionalidades críticas para la consolidación y visualización de datos institucionales.

Plan del Proyecto

Los principales elementos del plan del proyecto son:

• **Duración:** 4 meses

• **Framework:** Scrum

• **Número estimado de sprints:** 4

• **Entregables por sprint:** módulos incrementales de la plataforma PIVDE

• **Herramientas:** GitHub, Python, Streamlit, PostgreSQL, Jira, Sheets institucionales

Declaración de Alcance del Proyecto

El alcance del proyecto incluye:

Incluye:

• Desarrollo de una plataforma web en Python/Streamlit

• Integración de múltiples fuentes de datos institucionales

• Depuración, normalización y consolidación automática de datos

• Dashboards dinámicos para la toma de decisiones

• Reportes descargables en PDF, Excel y CSV

• Capacitación a usuarios clave

No incluye:

• Reemplazo completo de sistemas existentes

• Desarrollo de aplicaciones móviles

• Integración con plataformas externas no autorizadas por DITE

Objetivos del Proyecto (General y Específicos)

Objetivo General

Desarrollar e implementar una plataforma web que integre y visualice datos educativos provenientes de múltiples fuentes institucionales para mejorar la toma de decisiones en PERÚ-Educación.

Objetivos Específicos

1. Integrar información de diversas plataformas institucionales en un repositorio único.
2. Automatizar la depuración y normalización de datos mediante scripts programados.
3. Diseñar dashboards dinámicos que permitan análisis en tiempo real.
4. Implementar opciones de exportación de reportes institucionales.
5. Facilitar la consulta rápida de datos para usuarios autorizados.

Beneficios del Proyecto

• Reducción del tiempo de consolidación de datos.

• Eliminación de errores derivados de procesos manuales.

• Disponibilidad de información confiable y actualizada.

• Mayor eficiencia institucional.

• Facilita el seguimiento de indicadores educativos.

• Mejora la toma de decisiones de áreas estratégicas.

Cronograma

El proyecto se ejecutará en un periodo total de **4 meses,** distribuidos en **6 sprints** de dos semanas aproximadamente. Cada sprint incluye planificación, desarrollo, pruebas y revisión con el Product Owner.

Interesados

Tabla 16.

Interesados

Interesado	**Rol**	**Participación**
Director de Innovación Tecnológica	Sponsor	Supervisión y validación
Equipo DITE	Usuarios clave	Requerimientos y pruebas
Scrum Master	Facilitador	Gestión ágil
Desarrollador Full Stack	Desarrollo	Backend y frontend
Analista de Datos	Validación	Calidad de información

Supuestos

• El personal clave estará disponible para revisiones periódicas.

• La institución permitirá acceso seguro a las fuentes de datos.

• La infraestructura tecnológica se mantendrá estable.

• Se contará con retroalimentación constante del Product Owner.

47

Factores Críticos de Éxito (FCE)

• Participación del Product Owner.

• Disponibilidad de las fuentes de datos.

• Comunicación clara entre las áreas.

• Iteraciones constantes y validación continua.

Riesgos

Tabla 17.

Matriz de Riesgos del Proyecto

N° Riesgo identificado

Probabilidad (P)

Impacto (I)

Nivel de riesgo (P×I)

Clasificación Estrategia de respuesta

R1

Retraso en el acceso a datos institucionales

4 (Alta) 4 (Alta) 16 Alto

Mitigar: gestionar accesos anticipados y definir responsables

R2

Cambios en los requerimientos durante el desarrollo

3 (Media) 4 (Alta) 12 Alto

Controlar mediante backlog priorizado y gestión de cambios

R3

Sobrecarga operativa del personal institucional

3 (Media) 3 (Media) 9 Medio

Aceptar y coordinar cronograma de entregables

48

R4

Fallos en la infraestructura tecnológica

2 (Baja) 4 (Alta) 8 Medio

Mitigar mediante respaldos y uso de servicios en la nube

R5

Baja adopción del sistema por parte de los usuarios finales

3 (Media) 3 (Media) 9 Medio

Mitigar mediante capacitación y sensibilización

Comunicaciones

La estrategia de comunicaciones se define de acuerdo con la estructura Scrum:

• **Reuniones diarias (Daily Standup):** Seguimiento del avance del sprint.

• **Sprint Planning:** Definición de tareas del sprint.

• **Sprint Review:** Presentación de incrementos funcionales.

• **Sprint Retrospective:** Identificación de mejoras internas.

• **Informe semanal al Product Owner:** Estado del proyecto, riesgos y avances.

• **Canales principales:** Google Meet, Gmail institucional, Jira, GitHub y WhatsApp corporativo.

49

Capítulo III

Ejecución del Proyecto

Fase Inicio

La visión del proyecto PIVDE consiste en proporcionar a PERÚ-Educación una plataforma centralizada, automatizada y confiable para la integración, depuración, consolidación y visualización de datos educativos provenientes de diversas fuentes institucionales. Esta solución permitirá disponer de información actualizada en tiempo real, facilitando la toma de decisiones basada en evidencia y optimizando los procesos analíticos y operativos de la institución.

El proyecto busca eliminar los procesos manuales de consolidación con hojas de cálculo, reducir tiempos operativos, mejorar la calidad de los datos y entregar dashboards interactivos que brinden indicadores claros, filtros avanzados y reportes descargables para los usuarios

autorizados.

Presentación del Lanzamiento del Proyecto (Kick Off)

El proyecto inició formalmente con una reunión de lanzamiento (Kick-Off *Meeting)* con el Product Owner y los integrantes del equipo Scrum. En esta reunión se revisaron:

• Los objetivos generales del proyecto.

• El alcance funcional de la plataforma PIVDE.

• El marco de trabajo ágil a aplicar (Scrum).

• Los roles del equipo.

• Los riesgos iniciales.

• La planificación general de sprints.

Se estableció una comunicación continua entre el Product Owner (representante de la DITE) y el equipo de desarrollo, y se definieron los artefactos principales que se emplearán durante el proyecto, así como los criterios de aceptación que permitirán validar cada incremento.

Figura 7.

Presentación del Lanzamiento del Proyecto (Kick Off)

Requerimientos/Historias de Usuarios (HU)

Las historias de usuario permiten capturar las necesidades reales de los usuarios de PERÚ-Educación, enfocándose en el valor que aporta el sistema y en los objetivos que se desean alcanzar con su implementación. Este enfoque facilita la comprensión de los requerimientos desde la perspectiva del usuario final, promoviendo una comunicación clara entre los actores del proyecto.

Para la documentación de los requerimientos funcionales se ha utilizado el **modo de**

representación 1, el cual consiste en describir las historias de usuario mediante una estructura narrativa simple basada en el formato: *"Como [rol], deseo [funcionalidad], para [beneficio]".* Este modo de representación prioriza la claridad del requerimiento y el valor que aporta al usuario, permitiendo identificar de forma directa el propósito de cada funcionalidad y su contribución al sistema.

Según **Cohn (2004),** las historias de usuario son descripciones breves y comprensibles que expresan una necesidad del negocio desde el punto de vista del usuario, evitando especificaciones técnicas extensas y facilitando su validación y priorización. Asimismo, **Schwaber y Sutherland (2020)** señalan que este tipo de representación es ampliamente utilizado en marcos ágiles como Scrum, ya que favorece la adaptabilidad al cambio y el desarrollo incremental del producto.

En ese sentido, el uso del modo de representación 1 resulta adecuado para el presente Trabajo de Aplicación Profesional, ya que permite alinear los requerimientos del sistema PIVDE con las necesidades institucionales de PERÚ-Educación y con los objetivos del proyecto.

52

Tabla 18.

Historias de Usuario

Código Historia de Usuario
HU01 Como analista de datos, deseo integrar en un solo repositorio la información proveniente de PerúEduca y la Mesa de Ayuda para visualizarla de forma unificada.
HU02 Como analista, quiero que el sistema depure automáticamente registros duplicados o inconsistentes para asegurar la calidad de los datos.
HU03 Como gestor educativo, deseo visualizar dashboards con indicadores clave (KPIs) para monitorear tendencias institucionales.
HU04 Como usuario autorizado, deseo aplicar filtros dinámicos en los dashboards para analizar la información según mis necesidades.
HU05 Como usuario institucional, deseo descargar reportes en formatos PDF, Excel o CSV para análisis externo.
HU06 Como administrador, deseo configurar nuevas fuentes de datos o actualizar rutas existentes.
HU07 Como Product Owner, deseo visualizar métricas del uso de la plataforma para evaluar su impacto.

53

Product Backlog

El Product Backlog reúne todas las funcionalidades, ordenadas y alineadas a las

prioridades del Product Owner.

Tabla 19.

Product Backlog Inicial

ID Requerimiento Prioridad

PB01 Integración de fuentes de datos institucionales Alta

PB02 Módulo de depuración automática Alta

PB03 Consolidación en BD centralizada Alta

PB04 Dashboards institucionales Alta

PB05 Filtros dinámicos Media

PB06 Descarga de reportes Media

PB07 Configuración de nuevas fuentes Media

PB08 Métricas de uso Baja

54

Fase Planificación

Historias de Usuarios Priorizados

El Product Owner priorizó las historias según valor, urgencia y dependencia técnica.

Tabla 20.

Historias Priorizadas (MoSCoW)

Código Historia Prioridad

HU01 Integración de datos Must

HU02 Depuración automática Must

HU03 Dashboards Must

HU04 Filtros Should

HU05 Exportaciones Should

HU06 Configuración data sources Could

HU07 Métricas de uso Could

55

Estimación de Tareas de las HU

El equipo Scrum estimó cada historia en puntos de historia (Story Points), basado en complejidad, esfuerzo y dependencias.

Tabla 21.

Estimación de Historias

Historia Puntos (Story Points)

HU01 8

HU02 5

HU03 13

HU04 3

HU05 3

HU06 5

HU07 2

56

Product Backlog Priorizado y Estimado

Tabla 22.

Product Backlog Priorizado

ID Historia Prioridad SP

PB01 HU01 Alta 8

PB02 HU02 Alta 5

PB03 HU03 Alta 13

PB04 HU04 Media 3

PB05 HU05 Media 3

PB06 HU06 Media 5

PB07 HU07 Baja 2

57

Cronograma/Sprint Planning

El proyecto se divide en 4 **sprints,** cada uno con una duración aproximada de 3 a 4 semanas.

Sprint 1 – Integración y Fundamentos

• HU01 Integración de fuentes de datos

• HU02 Depuración automática

• Configuración de base de datos centralizada

• Carga inicial de datos institucionales

Sprint 2 – Visualización y Análisis

• HU03 Dashboards institucionales

• HU04 Filtros dinámicos

• Validación de indicadores con Product Owner

Sprint 3 – Reportes y Configuración

• HU05 Exportación de reportes (PDF, Excel, CSV)

• HU06 Configuración de nuevas fuentes de datos

• Ajustes de seguridad y permisos

58

Sprint 4 – Métricas, Optimización y Cierre

• HU07 Métricas de uso de la plataforma

• Optimización de rendimiento

• Preparación de entrega final

Descomposición de Tareas de las HU (Plantilla Scrum)

Aunque Scrum no utiliza una EDT tradicional, se presenta una descomposición de alto nivel de las actividades del proyecto con fines académicos y de trazabilidad, alineada a los incrementos por sprint y a las historias de usuario priorizadas.

Tabla 23.

Descomposición de actividades del proyecto (alineada a Scrum)

Código Entregable Actividad

1.0 Planificación 1.1 Sprint Planning 1.2 Backlog Refinement

2.0 Integración de Datos 2.1 Análisis 2.2 ETL 2.3 BD

3.0 Dashboards 3.1 Diseño 3.2 Implementación

4.0 Reportes 4.1 Exportación 4.2 Validación

5.0 Configuraciones 5.1 Fuentes nuevas 5.2 Seguridad

59

Fase de Implementación

Incrementos por Sprint

Sprint 1 – Incremento

• Integración de PerúEduca

• Integración de Mesa de Ayuda

• Depuración automática (primera versión)

• BD centralizada creada y cargada

Sprint 2 – Incremento

• Dashboards interactivos

• Indicadores clave (KPIs)

• Filtros dinámicos

Sprint 3 – Incremento

• Exportación avanzada (PDF, Excel, CSV)

• Métricas de uso

• Nuevas fuentes configurables

60

Daily Scrum Meeting

Las sesiones *Daily Scrum* se realizaron diariamente a las 10:00 a.m., con duración máxima de 15 minutos.

Se respondieron las preguntas clave:

• ¿Qué hice ayer?

• ¿Qué haré hoy?

• ¿Qué impedimentos tengo?

Desarrollo del Sprint Backlog

Cada historia de usuario fue dividida en tareas técnicas claras:

Ejemplo Sprint 2:

• Crear estructura del dashboard

• Conectar dashboard con BD

• Implementar librerías Plotly/Altair

• Programar filtros dinámicos

• Validar indicadores con Product Owner

61

Fase de Revisión y Retrospectiva

Scrum Board

Durante la fase de revisión y retrospectiva, el equipo Scrum utilizó un Scrum Board para

visualizar el estado de las tareas asociadas a cada historia de usuario del sprint. Este tablero permitió dar seguimiento al avance del trabajo, identificar bloqueos y verificar la finalización de las actividades antes de la revisión con el Product Owner.

El Scrum Board se estructuró en cuatro columnas principales: To Do, In Progress, Testing y Done, representando el ciclo de vida de cada tarea durante el sprint.

Tabla 24.

Scrum Board del Sprint

Tarea	To Do	In Progress	Testing	Done
Integración de datos desde MongoDB				✔
Carga de datos en MySQL (perueduca4)				✔
Depuración de registros duplicados				✔
Desarrollo de dashboards interactivos				✔
Implementación de filtros dinámicos				✔
Exportación de reportes (PDF, Excel, CSV)				✔
Validación de indicadores con PO				✔

Sprint Review

Al cierre del sprint, se realizó la **Sprint Review** con la participación del Product Owner y usuarios clave de PERÚ-Educación. Durante esta sesión se presentaron los incrementos desarrollados, los cuales incluyeron:

• Integración de fuentes de datos institucionales.

• Visualización de dashboards con indicadores clave.

• Aplicación de filtros dinámicos.

• Generación y exportación de reportes.

El Product Owner validó las funcionalidades presentadas y confirmó que los incrementos cumplían con los criterios de aceptación definidos, aprobando su incorporación al producto.

Burn Down Chart

El Burndown Chart fue utilizado para monitorear el avance del trabajo durante los sprints del proyecto, comparando el trabajo planificado con el trabajo real.

Para efectos de este Trabajo de Aplicación Profesional, se presenta un Burndown

Chart representativo correspondiente a uno de los sprints ejecutados, el cual evidencia la tendencia de reducción del trabajo pendiente, la velocidad del equipo y el cumplimiento de los objetivos del sprint.

63

Figura 8.

Burn Down Chart – Sprint 2 (Visualización y Análisis)

Impediment Log

Ejemplos típicos:

• Retrasos por acceso a fuentes de datos

• Inconsistencias estructurales en archivos originales

• Cambios en requisitos de visualización

• Demoras en validación de indicadores

64

Retrospectiva del Sprint

Las conclusiones comunes:

• Mejorar comunicación con el PO

• Ajustar carga de tareas por sprint

• Incrementar pruebas automáticas

• Refinar backlog con mayor detalle

Fase de Lanzamiento

Acta de Conformidad de Entregables

Se documentó la validación final de:

• Dashboards

• Módulos de reportes

• BD unificada

• Fuentes de datos

• Exportaciones

El Product Owner aprobó la solución final PIVDE.

65

Capítulo IV

Programación

Implementación de la Arquitectura de Software

La plataforma PIVDE fue desarrollada bajo una arquitectura modular de tipo cliente–servidor, orientada a la integración, procesamiento y visualización de datos educativos provenientes de múltiples fuentes institucionales. Este enfoque permitió desacoplar la capa de presentación de la lógica de negocio y de la gestión de datos, facilitando la mantenibilidad, escalabilidad y evolución del sistema.

La arquitectura se organiza en tres capas principales:

• **Capa de presentación (Frontend):**

Implementada mediante el framework Streamlit, responsable de la interacción con el usuario final. Esta capa permite la visualización de dashboards, aplicación de filtros dinámicos, consulta de indicadores clave y exportación de reportes en distintos formatos.

• **Capa de servicios y lógica de negocio (Backend):**

Implementada en Python, concentra los procesos de validación, transformación, consulta y consolidación de información. Esta capa gestiona la comunicación entre la interfaz de usuario y la base de datos central, así como la ejecución de procesos automáticos de actualización de datos.

66

• **Capa de datos:**

Compuesta por una base de datos relacional MySQL denominada *perueduca4,* la cual almacena información consolidada y normalizada proveniente de diversas fuentes institucionales, incluyendo bases de datos NoSQL (MongoDB) y sistemas relacionales

externos.

Este diseño arquitectónico permite que los procesos de integración y carga de datos (ETL) se ejecuten de manera independiente al consumo de información por parte de los usuarios, optimizando el rendimiento y garantizando la consistencia de los datos visualizados.

67

Figura 9.

Arquitectura general de la plataforma PIVDE

68

Nota: Diagrama de arquitectura cliente–servidor con frontend, backend y base de datos. Elaboración propia.

Configuración del Entorno de Desarrollo y Dependencias

El entorno de desarrollo de la plataforma PIVDE fue configurado utilizando el lenguaje de programación Python, debido a su flexibilidad, amplio ecosistema de librerías y facilidad de integración con distintos motores de bases de datos.

Para asegurar buenas prácticas de seguridad y despliegue, las credenciales de conexión, rutas y parámetros sensibles fueron gestionados mediante **variables de entorno,** definidas en un archivo .env, evitando su exposición directa en el código fuente.

La gestión de dependencias del proyecto se realizó mediante un archivo requirements.txt, el cual garantiza la compatibilidad entre librerías y la correcta ejecución del sistema en distintos entornos.

Tabla 25.

Principales librerías utilizadas en el proyecto

Librería Propósito

Streamlit Construcción de interfaces web interactivas

Pandas Manipulación y análisis de datos

NumPy Operaciones numéricas

Matplotlib / Altair Visualización de datos

PyMongo Conexión y extracción de datos desde MongoDB

mysql-connector-python Conexión a base de datos MySQL

Requests Consumo de servicios backend

Figura 10.

Archivo requirements del Backend

Figura 11.

Archivo requirements del Frontend

70

El uso de este entorno permitió una integración eficiente entre procesos ETL, base de datos y visualización de información.

Implementación de la Base de Datos

Como repositorio central de información se implementó una base de datos relacional MySQL denominada **perueduca4,** diseñada para almacenar datos consolidados y normalizados provenientes de múltiples sistemas educativos institucionales.

La estructura de la base de datos fue definida mediante el archivo structure_perueduca4.sql, el cual establece las tablas necesarias para soportar los procesos analíticos, consultas, generación de indicadores y reportes.

El diseño relacional permite:

• Garantizar la integridad y consistencia de los datos.

• Optimizar el tiempo de respuesta en consultas analíticas.

• Facilitar la correlación de información entre distintas plataformas educativas.

• Reducir la carga sobre los sistemas fuente originales.

La base de datos *perueduca4* actúa como un **repositorio analítico central,** separando los procesos operativos de las fuentes originales del consumo de información institucional.

71

Figura 12.

Estructura de la base de datos perueduca4

Nota: Captura del esquema de la base de datos MySQL. Elaboración propia.

72

Procesos de Integración y Carga de Datos (ETL)

Para garantizar la disponibilidad y calidad de la información, se implementó un proceso **ETL (Extract, Transform, Load)** encargado de extraer datos desde bases de datos NoSQL (MongoDB) y sistemas relacionales, transformarlos y cargarlos en la base de datos central MySQL *perueduca4.*

Este proceso fue desarrollado en Python mediante el script Exportar_MongoDB_ImportarMySQL_2.py, el cual ejecuta las siguientes etapas:

• **Extracción:**

Conexión a MongoDB y lectura de colecciones relevantes que contienen información académica y de usuarios.

• **Transformación:**

Normalización de estructuras, eliminación de registros duplicados, depuración de campos inconsistentes y adaptación de los datos a un esquema relacional.

• **Carga:**

Inserción controlada de los datos transformados en las tablas correspondientes de MySQL.

Este proceso permite desacoplar las fuentes operativas del sistema analítico, asegurando que los dashboards trabajen sobre información consistente y validada.

73

Figura 13.

Script ETL

74

Figura 14.

Flujo ETL

Gestión de Conexiones y Fuentes de Datos Institucionales

La plataforma PIVDE se integra con múltiples fuentes de datos institucionales que utilizan diferentes motores de base de datos, tales como MongoDB, MySQL y SQL Server. Estas conexiones son gestionadas mediante variables de entorno, lo que permite centralizar la configuración y proteger información sensible.

La integración de datos se basa en **campos clave comunes,** principalmente:

• Identificador de usuario (UserKeyCloak)

• Correo electrónico

75

• Número de documento

Esta estrategia permite una correcta correlación de registros entre sistemas heterogéneos, garantizando la unicidad de los usuarios y la coherencia de la información consolidada.

Tabla 26.

Fuentes de datos integradas en la plataforma PIVDE

Nombre de la fuente

Sigla Ubicac ión

Tipo de BD

Nombre de BD

Sistema origen

Informa ción utilizada

Campo(s) clave de integració n

Campus Nube

CN AWS MySQ L

moodledb_pr d

Moodle 3.9

Usuarios, cursos, accesos

UserKeyCl oak, Correo, Núm. Document o

PerúEduca ción

PE Externa Mongo DB

amdbuse1pep rd102

Platafor ma PerúEdu ca

Informaci ón académic a y de usuarios

UserKeyCl oak, Correo, Núm. Document o

PerúEduca ción Local

PE Local

Local MySQ L

perueduca4 Copia diaria PerúEdu ca

Datos consolida dos

UserKeyCl oak, Correo, Núm. Document o

Sistema de Capacitaci ón

SISC AP

Externa MySQ L

sisdigetecapv 5

Sistema de capacita ción

Registros de capacitac ión

Núm. Document o, Correo

76

Aula Virtual

AV Externa MySQ L

aulasvirtuales Moodle 2.9

Cursos y usuarios

Núm. Document o, Correo

Campus Virtual 1– 12

CV1– CV12

Externa MySQ L / SQL Server

db_aula_virtu al*

Moodle 3.6

Informaci ón académic a

UserKeyCl oak, Correo, Núm. Document o

El proceso de integración de datos se apoya en campos clave comunes entre los distintos sistemas, principalmente el identificador de usuario (UserKeyCloak), el correo electrónico y el número de documento. En el caso de plataformas Moodle, el UserKeyCloak se obtiene del campo username de la tabla mdl_user, mientras que el correo electrónico se extrae del campo email. El número de documento se obtiene a partir de campos personalizados definidos en mdl_user_info_field y sus valores correspondientes en mdl_user_info_data. Esta estrategia permite una correcta correlación de registros entre sistemas heterogéneos.

77

Implementación del Backend de la Plataforma

El backend del sistema fue desarrollado en Python y se encarga de la gestión de la lógica de negocio, validación de datos y comunicación con la base de datos. Entre sus principales funciones se encuentran:

• Consulta y agregación de información desde MySQL.

• Ejecución de procesos automáticos de actualización de datos.

• Preparación de datasets para visualización.

• Control de accesos y validación de parámetros.

Esta capa permite que la plataforma procese grandes volúmenes de información de

manera eficiente y segura.

78

Figura 15.

Estructura básica del backend

79

Figura 16.

Muestra del archivo main.py

80

Figura 17.

Muestra del archivo db.py

81

Implementación del Frontend – Dashboard

El frontend fue desarrollado utilizando Streamlit, permitiendo construir una interfaz web interactiva de rápida implementación. La interfaz incluye:

• Dashboards con gráficos dinámicos.

• Indicadores clave (KPIs).

• Filtros por periodo, curso y plataforma.

• Visualización clara y amigable de la información.

El uso de Streamlit facilitó la adopción del sistema por parte de los usuarios institucionales, reduciendo la curva de aprendizaje.

82

Figura 18.

Estructura básica del Frontend

83

Figura 19.

Muestra del archivo app.py

84

Figura 20.

Muestra del archivo cursos.py

85

Consultas, Reportes e Indicadores

El sistema implementa consultas analíticas optimizadas sobre la base de datos central, permitiendo generar reportes consolidados y exportarlos en formatos PDF, Excel y CSV. Estas funcionalidades apoyan el análisis externo y la elaboración de informes institucionales.

Figura 21.

Modelo de consulta de cursos

86

Figura 22.

Informe gráfico de conexiones por día

87

Figura 23.

Consulta de matrículas por usuarios en todos los campus

88

Figura 24.

Reporte exportado en formato Excel

89

Capítulo V

Pruebas de Calidad de Software

Objetivo de las Pruebas de Calidad

El objetivo principal de las pruebas de calidad de software en la plataforma PIVDE fue verificar el correcto funcionamiento del sistema, asegurar la integridad y confiabilidad de la información procesada, y validar que las funcionalidades implementadas cumplan con los requerimientos definidos por los usuarios institucionales y el Product Owner.

Asimismo, las pruebas permitieron identificar posibles errores en los procesos de integración, visualización y exportación de datos, garantizando que la solución entregue información consistente, oportuna y útil para la toma de decisiones.

Estrategia de Pruebas Aplicada

La estrategia de pruebas aplicada se basó en un enfoque progresivo alineado con el marco de trabajo ágil Scrum, realizando validaciones iterativas al cierre de cada sprint.

Se aplicaron pruebas unitarias y pruebas integrales. La ejecución de las pruebas fue documentada mediante capturas de pantalla del sistema, las cuales evidencian el correcto funcionamiento de los módulos evaluados y se presentan en los apartados siguientes.

90

Pruebas Unitarias

Las pruebas unitarias se realizaron con el objetivo de validar el correcto funcionamiento de los componentes individuales del sistema, principalmente aquellos relacionados con la lógica de negocio, procesos ETL y consultas a la base de datos.

Componentes evaluados:

• Scripts de integración y carga de datos (ETL).

• Conexión a bases de datos MongoDB y MySQL.

• Funciones de transformación y depuración de datos.

• Consultas SQL para obtención de indicadores.

• Funciones de generación de gráficos.

A continuación, se presentan las evidencias de la ejecución de las pruebas unitarias.

91

Figura 25.

Ejecución del proceso ETL Cargar conexiones diarias

Descripción: Captura de la ejecución del script ETL mostrando carga exitosa de datos.

92

Figura 26.

Consulta SQL validada

Descripción: Captura de resultados correctos de consulta SQL.

93

Figura 27.

Generación de gráfico unitario

Descripción: Captura de gráfico generado correctamente.

94

Pruebas Integrales

Las pruebas integrales permitieron validar el funcionamiento del sistema como un todo, verificando la correcta interacción entre los distintos módulos de la plataforma PIVDE.

Escenarios evaluados:

• Integración completa de datos.

• Visualización de dashboards.

• Aplicación de filtros dinámicos.

• Exportación de reportes.

A continuación, se presentan las evidencias de las pruebas integrales.

95

Figura 28.

Dashboard principal

Descripción: Visualización general del estado de conexiones.

96

Figura 29.

Aplicación de filtros

Descripción: Filtros dinámicos aplicados correctamente.

97

Figura 30.

Muestra de catálogo de cursos

Descripción: Cursos listados con los campos necesarios.

98

Figura 31.

Exportación de reporte

Descripción: Reporte exportado en formato Excel.

99

Registro de Incidencias y Correcciones

Durante la ejecución de las pruebas se identificaron incidencias menores, principalmente relacionadas con:

• Inconsistencias en formatos de datos provenientes de algunas fuentes.

• Ajustes en la visualización de gráficos para mejorar la legibilidad.

• Optimización de consultas para mejorar tiempos de respuesta.

Estas incidencias fueron corregidas oportunamente dentro de los sprints correspondientes, permitiendo mejorar la estabilidad, usabilidad y rendimiento del sistema antes de su entrega final.

Validación por el Usuario

La validación final del sistema fue realizada por usuarios clave de PERÚ-Educación y por el Product Owner, quienes revisaron las funcionalidades implementadas y confirmaron que la plataforma cumple con los objetivos planteados en el proyecto.

Los usuarios destacaron principalmente:

• La facilidad de uso de la interfaz.

• La rapidez en la generación de reportes.

• La disponibilidad de información consolidada en tiempo real.

• La reducción significativa de tareas manuales.

100

Como resultado, el sistema fue **aprobado para su uso institucional,** quedando listo para su despliegue y operación.

101

Conclusiones

1. La implementación de la plataforma PIVDE permitió **integrar y centralizar información educativa proveniente de múltiples fuentes heterogéneas,** eliminando la dispersión de datos y los procesos manuales de consolidación que afectaban la oportunidad y confiabilidad de los reportes institucionales.

2. El uso de una **arquitectura modular cliente–servidor,** junto con tecnologías de código abierto como Python, Streamlit y MySQL, facilitó el desarrollo de una solución flexible, escalable y alineada con las necesidades operativas de la organización pública PERÚ-Educación.

3. La incorporación de procesos ETL automatizados permitió **mejorar la calidad de los datos,** reduciendo duplicidades e inconsistencias provenientes de los sistemas fuente, y garantizando que los dashboards y reportes se basen en información validada y consistente.

4. La aplicación del marco de trabajo ágil Scrum favoreció la **entrega incremental de funcionalidades,** la validación continua con el Product Owner y la rápida corrección de incidencias, contribuyendo al cumplimiento del cronograma del proyecto.

5. Las pruebas de calidad de software evidenciaron que la plataforma PIVDE cumple con los **requerimientos funcionales y no funcionales definidos,** demostrando estabilidad, facilidad de uso y tiempos de respuesta adecuados para apoyar la toma de decisiones institucionales.

102

Recomendaciones

1. Se recomienda **implementar la ejecución programada del proceso ETL,** mediante

tareas automáticas (por ejemplo, cron jobs), con el fin de asegurar la actualización periódica de la información sin intervención manual.

2. Es recomendable ampliar progresivamente la plataforma PIVDE incorporando **nuevas fuentes de datos institucionales,** así como nuevos indicadores y métricas que fortalezcan el análisis y la gestión educativa basada en evidencia.

3. Se sugiere establecer **políticas de gobernanza y calidad de datos,** que definan responsables, estándares y controles, garantizando la sostenibilidad y confiabilidad de la información a largo plazo.

4. Se recomienda brindar **capacitaciones periódicas a los usuarios finales,** a fin de maximizar el aprovechamiento de las funcionalidades de la plataforma y fomentar una cultura organizacional orientada al uso de datos para la toma de decisiones.

5. Finalmente, se sugiere evaluar el despliegue de la plataforma en **entornos de nube institucional,** con el objetivo de mejorar la disponibilidad, escalabilidad y continuidad operativa del sistema.

Referencias Bibliográficas

Ministerio de Educación del Perú. (2024). Dirección de Innovación Tecnológica en Educación – DITE. MINEDU. Recuperado de https://www.gob.pe/minedu

SCRUM, metodologías ágiles y ciclo de vida

3,7 **Schwaber, K., & Sutherland, J. (2020).** *The Scrum Guide: The definitive guide to Scrum.* Scrum.org. https://scrumguides.org/

Cohn, M. (2004). *User stories applied: For agile software development.* Addison-Wesley.

Pressman, R. S., & Maxim, B. R. (2019). *Software engineering: A practitioner's approach* (9th ed.). McGraw-Hill Education.

Sommerville, I. (2016). *Software engineering* (10th ed.). Pearson.

Gestión de proyectos, planificación y procesos

4 **Project Management Institute. (2021).** *A guide to the project management body of knowledge (PMBOK Guide)* (7th ed.). Project Management Institute.

Kerzner, H. (2017). *Project management: A systems approach to planning, scheduling, and controlling* (12th ed.). Wiley.

104

Harmon, P. (2019). *Business process change: A business process management guide for managers and process professionals* (4th ed.). Morgan Kaufmann.

Arquitectura de software, ingeniería y patrones

Bass, L., Clements, P., & Kazman, R. (2012). *Software architecture in practice* (3rd ed.). Addison-Wesley.

Freeman, E., & Robson, E. (2021). *Head First Design Patterns* (2nd ed.). O'Reilly Media.

Analítica de datos, dashboards y visualización

Few, S. (2013). *Information dashboard design: Displaying data for at-a-glance monitoring* (2nd ed.). Analytics Press.

Knaflic, C. N. (2015). *Storytelling with data: A data visualization guide for business professionals.* Wiley.

Turban, E., Sharda, R., & Delen, D. (2014). *Business intelligence and analytics: Systems for decision support* (10th ed.). Pearson.

Gobernanza, integración y calidad de datos

Imai, M. (2012). *Gemba Kaizen: A commonsense low-cost approach to management* (2nd ed.). McGraw-Hill.

Ishikawa, K. (1985). *What is total quality control? The Japanese way.* Prentice Hall.

Loshin, D. (2019). *Master data management and data governance.* Morgan Kaufmann.

105

Tecnologías usadas en tu proyecto (Python, Streamlit, Base de Datos)

Van Rossum, G., & Drake, F. L. (2009). *Python 3 reference manual.* CreateSpace.

Streamlit Inc. (2023). *Streamlit documentation.* https://docs.streamlit.io/

Hernández, M. A., & García-Molina, H. (2019). *Database systems: The complete book* (2nd ed.). Pearson.

Cohn, M. (2004). User Stories Applied: For Agile Software Development. Addison-Wesley.

Schwaber, K., & Sutherland, J. (2020). The Scrum Guide. Scrum.org.

Anexos

Anexo 01. Arquitectura general de la plataforma PIVDE

Descripción

El presente anexo muestra la arquitectura general de la plataforma PIVDE, la cual fue diseñada bajo un enfoque modular de tipo cliente–servidor. En el diagrama se evidencian las capas de presentación, lógica de negocio y datos, así como los procesos de integración y carga de información provenientes de múltiples fuentes institucionales.

Contenido del anexo

• Diagrama de arquitectura del sistema.

• Relación entre frontend, backend, procesos ETL y base de datos.

• Fuentes de datos integradas y repositorio central.

Figura 32.

Arquitectura general de la plataforma PIVDE

108

Anexo 02. Modelo de la base de datos perueduca4

Descripción

En este anexo se presenta el modelo de la base de datos *perueduca4,* utilizada como repositorio central de información consolidada. El diseño de la base de datos permite almacenar, relacionar y explotar información educativa proveniente de diferentes sistemas institucionales.

Contenido del anexo

• Esquema de la base de datos MySQL.

• Tablas principales y relaciones.

• Estructura utilizada para consultas analíticas y reportes.

109

Figura 33.

Esquema de la base de datos perueduca4

110

Anexo 03. Script de integración y carga de datos (ETL)

Descripción

El presente anexo muestra extractos representativos del script desarrollado en Python para la extracción, transformación y carga de datos desde bases de datos NoSQL y relacionales hacia la base de datos central *perueduca4.* Este proceso garantiza la calidad, consistencia y disponibilidad de la información utilizada por la plataforma PIVDE.

Figura 34.

Ejecución del proceso ETL MongoDB – MySQL

111

Anexo 04. Configuración del entorno de desarrollo

Descripción

Este anexo presenta la configuración del entorno de desarrollo utilizada en la plataforma PIVDE, incluyendo la definición de variables de entorno y las dependencias principales del proyecto. Esta configuración permite asegurar buenas prácticas de seguridad y facilitar el despliegue del sistema en distintos entornos.

Figura 35.

Archivo .env (sin credenciales reales)

112

Figura 36.

Archivo requirements.txt del Backend

Figura 37.

Archivo requirements.txt del Frontend

113

Anexo 05. Capturas de funcionamiento de la plataforma PIVDE

Descripción

En este anexo se incluyen capturas de pantalla que evidencian el funcionamiento de la plataforma PIVDE. Las imágenes muestran los dashboards interactivos, filtros dinámicos, generación de reportes y resultados de los procesos de integración de datos.

Figura 38.

Pantalla de acceso

114

Figura 39.

Ejecución del Backend – API de Integración de Datos (FastAPI / Uvicorn)

Descripción: Consola de ejecución del backend de la plataforma PIVDE, mostrando el inicio correcto del servicio API, la carga de motores de integración y la validación del endpoint

de salud (/health).

115

Figura 40.

Ejecución del Frontend – Dashboard en Streamlit

Descripción: Consola de ejecución del módulo frontend de la plataforma PIVDE, donde se evidencia el levantamiento del dashboard interactivo mediante Streamlit y su disponibilidad a través de un navegador web.