



UNIVERSIDAD TÉCNICA DE AMBATO

**FACULTAD DE INGENIERÍA EN SISTEMAS, ELECTRÓNICA E
INDUSTRIAL**

CARRERA DE SOFTWARE

Tema:

**IMPLEMENTACIÓN DE UN SISTEMA DE REPORTERÍA Y VISUALIZACIÓN DE DATOS
CON FUNCIONES DE RECOMENDACIÓN Y PREDICCIÓN, UTILIZANDO TÉCNICAS
DE BUSINESS INTELLIGENCE (BI) E INTELIGENCIA ARTIFICIAL (IA), EN LA
DIRECCIÓN DE INVESTIGACIÓN Y DESARROLLO (DIDE) DE LA UNIVERSIDAD
TÉCNICA DE AMBATO (UTA)**

Trabajo de titulación modalidad Proyecto de Investigación, presentado previo a la
obtención del título de Ingeniero de Software

ÁREA: Desarrollo de software

LÍNEA DE INVESTIGACIÓN: Software, Tecnologías de la información y Ciencia
de datos

AUTOR: Christopher Paul Santamaría Márquez

TUTOR: Ing. Rubén Eduardo Nogales Portero, PhD.

Ambato - Ecuador

julio– 2026

APROBACIÓN DEL TUTOR

En calidad de tutor del trabajo de titulación con el tema: IMPLEMENTACIÓN DE UN SISTEMA DE REPORTERÍA Y VISUALIZACIÓN DE DATOS CON FUNCIONES DE RECOMENDACIÓN Y PREDICCIÓN, UTILIZANDO TÉCNICAS DE BUSINESS INTELLIGENCE (BI) E INTELIGENCIA ARTIFICIAL (IA), EN LA DIRECCIÓN DE INVESTIGACIÓN Y DESARROLLO (DIDE) DE LA UNIVERSIDAD TÉCNICA DE AMBATO (UTA), desarrollado bajo la modalidad Proyecto de Investigación por el señor Christopher Paul Santamaría Márquez, estudiante de la Carrera de Software, de la Facultad de Ingeniería en Sistemas, Electrónica e Industrial, de la Universidad Técnica de Ambato, me permito indicar que el estudiante ha sido tutorado durante todo el desarrollo del trabajo hasta su conclusión, de acuerdo a lo dispuesto en el Artículo 17 del Reglamento para la Titulación de Grado en la Universidad Técnica de Ambato y el numeral 6.3 del instructivo del reglamento referido.

Ambato, julio 2026.



Validar únicamente en FIrmaEC.
Firmado electrónicamente por:
**RUBEN EDUARDO
NOGALES PORTERO**

Ing. Rubén Eduardo Nogales Portero, PhD.

TUTOR

AUTORÍA

El presente trabajo de titulación con el tema: IMPLEMENTACIÓN DE UN SISTEMA DE REPORTERÍA Y VISUALIZACIÓN DE DATOS CON FUNCIONES DE RECOMENDACIÓN Y PREDICCIÓN, UTILIZANDO TÉCNICAS DE BUSINESS INTELLIGENCE (BI) E INTELIGENCIA ARTIFICIAL (IA), EN LA DIRECCIÓN DE INVESTIGACIÓN Y DESARROLLO (DIDE) DE LA UNIVERSIDAD TÉCNICA DE AMBATO (UTA), es absolutamente original, auténtico y personal y ha observado los preceptos establecidos en la Disposición General Quinta del Reglamento para la Titulación de Grado en la Universidad Técnica de Ambato. En tal virtud, el contenido, efectos legales y académicos que se desprenden del mismo son de exclusiva responsabilidad del autor.

Ambato, julio 2026.



Christopher Paul Santamaría Márquez

C.C. 1805358643

AUTOR

DERECHOS DE AUTOR

Autorizo a la Universidad Técnica de Ambato para que reproduzca total o parcialmente este trabajo de titulación dentro de las regulaciones legales e institucionales correspondientes. Además, cedo todos mis derechos de autor a favor de la institución con el propósito de su difusión pública, por lo tanto, autorizo su publicación en el repositorio virtual institucional como un documento disponible para la lectura y uso con fines académicos e investigativos de acuerdo con la Disposición General Cuarta del Reglamento para la Titulación de Grado en la Universidad Técnica de Ambato.

Ambato, julio 2026.



Christopher Paul Santamaría Márquez

C.C. 1805358643

AUTOR

APROBACIÓN DEL TRIBUNAL DE GRADO

En calidad de par calificador del informe final del trabajo de titulación presentado por el señor Christopher Paul Santamaría Márquez, estudiante de la Carrera de Software, de la Facultad de Ingeniería en Sistemas, Electrónica e Industrial, bajo la Modalidad Proyecto de Investigación, titulado IMPLEMENTACIÓN DE UN SISTEMA DE REPORTERÍA Y VISUALIZACIÓN DE DATOS CON FUNCIONES DE RECOMENDACIÓN Y PREDICCIÓN, UTILIZANDO TÉCNICAS DE BUSINESS INTELLIGENCE (BI) E INTELIGENCIA ARTIFICIAL (IA), EN LA DIRECCIÓN DE INVESTIGACIÓN Y DESARROLLO (DIDE) DE LA UNIVERSIDAD TÉCNICA DE AMBATO (UTA), nos permitimos informar que el trabajo ha sido revisado y calificado de acuerdo al Artículo 19 del Reglamento para la Titulación de Grado en la Universidad Técnica de Ambato y el numeral 6.4 del instructivo del reglamento referido. Para cuya constancia suscribimos, juntamente con el señor Presidente del Tribunal.

Ambato, julio 2026.



Validar únicamente en FirmaEC.
Firmado electrónicamente por:
**FRANKLIN OSWALDO
MAYORGA MAYORGA**

Ing. Franklin Oswaldo Mayorga Mayorga, Mg.
PRESIDENTE DEL TRIBUNAL



Validar únicamente en FirmaEC.
Firmado electrónicamente por:
**HERNAN FABRICIO
NARANJO AVALOS**

Ing. Hernán Fabricio Naranjo Avalos, Mg.
PROFESOR CALIFICADOR



Validar únicamente en FirmaEC.
Firmado electrónicamente por:
**MARCOS RAPHAEL
BENITEZ ALDAS**

Ing. Marcos Raphael Benitez Aldas, MSc.
PROFESOR CALIFICADOR

DEDICATORIA

Dedico este trabajo de titulación, en primer lugar, a Dios, por ser mi guía, mi fortaleza y mi mayor sustento durante todo este proceso. A Él le atribuyo el mérito principal de haberme permitido avanzar, superar las dificultades y culminar esta etapa tan importante de mi vida académica y profesional.

A mis padres, por su amor, sacrificio, confianza y apoyo incondicional. Gracias por acompañarme en cada decisión, por motivarme a seguir adelante y por enseñarme con su ejemplo el valor del esfuerzo, la responsabilidad y la perseverancia.

*A mis hermanos, **Alex Santamaría, Andrés Galora y Sebastián de Jesús**, por estar presentes en este camino y por ser parte fundamental de mi vida. De manera especial, dedico este logro a **Sebastián de Jesús**, quien ha sido una gran inspiración, motivo de orgullo y una razón especial para continuar con firmeza hasta culminar mi carrera.*

Este trabajo representa no solo el cumplimiento de una meta académica, sino también el resultado de la fe, el apoyo familiar y la motivación que me impulsaron a no rendirme.

AGRADECIMIENTO

Expreso mi más sincero agradecimiento a Dios, por brindarme sabiduría, fortaleza y perseverancia para culminar este trabajo de titulación.

Agradezco de manera especial a mi tutor, por su orientación, acompañamiento y apoyo durante el desarrollo de esta investigación. Su guía fue fundamental no solo en la toma de decisiones académicas, sino también en decisiones estudiantiles y personales que impulsaron la continuidad y culminación de este proyecto.

De igual manera, agradezco a la Universidad Técnica de Ambato y a la Dirección de Investigación y Desarrollo, por el apoyo institucional brindado, la apertura para el levantamiento de información y la colaboración necesaria para orientar esta propuesta hacia una necesidad real.

A mis padres y hermanos, por confiar en mí, por acompañarme en cada etapa y por darme la motivación necesaria para seguir adelante. Su apoyo fue esencial para alcanzar esta meta.

Finalmente, agradezco a mis amigos y personas cercanas que, con sus consejos, apoyo y palabras oportunas, me ayudaron a tomar buenas decisiones durante este proceso. Cada aporte fue valioso para mantenerme enfocado y culminar esta etapa con responsabilidad y gratitud.

ÍNDICE GENERAL DE CONTENIDOS

PORTADA.....	ii
APROBACIÓN DEL TUTOR.....	ii
AUTORÍA.....	iii
DERECHOS DE AUTOR	iv
APROBACIÓN DEL TRIBUNAL DE GRADO.....	v
DEDICATORIA.....	vi
AGRADECIMIENTO	vii
ÍNDICE GENERAL DE CONTENIDOS	viii
ÍNDICE DE TABLAS.....	xiii
ÍNDICE DE FIGURAS.....	xvi
ÍNDICE DE ANEXOS	xix
RESUMEN EJECUTIVO	xx
ABSTRACT	xxi
CAPÍTULO I. MARCO TEÓRICO.....	1
1.1 Tema de investigación.....	1
1.1.1 Planteamiento del problema.....	1
1.2 Antecedentes investigativos	4
1.2.1 Resumen de antecedentes.....	18
1.3 Fundamentación teórica	23

1.3.1 Business Intelligence.....	23
1.3.2 IA en entornos BI.....	24
1.3.3 Ciencia de datos y toma de decisiones.....	24
1.3.4 Reportería dinámica y dashboards	24
1.3.5 Sistemas de recomendación y predicción	24
1.4 Objetivos	25
1.4.1 Objetivo general	25
1.4.2 Objetivos específicos	25
CAPÍTULO II. METODOLOGÍA.....	26
2.1 Materiales.....	26
2.2 Métodos.....	27
2.2.1 Modalidad de la investigación	27
2.2.2 Población y muestra	27
2.2.3 Recolección de información.....	28
2.2.4 Procesamiento y análisis de datos	31
CAPÍTULO III. RESULTADOS Y DISCUSIÓN.....	35
3.1 Análisis del contexto y requerimientos	35

3.1.1 Procesamiento de la información	35
3.1.2 Análisis de requerimientos	39
3.1.3 Requerimientos funcionales y no funcionales	40
3.1.4 Diagramas UML del sistema.....	46
3.2 Diseño de la propuesta	77
3.2.1 Arquitectura general del sistema.....	77
3.2.2 Diseño del modelo transaccional OLTP	81
3.2.3 Diseño del modelo analítico OLAP	92
3.2.4 Diseño visual del sistema	106
3.3 Desarrollo de la propuesta.....	107
3.3.1 Selección de la metodología de desarrollo	107
3.3.2 Aplicación de la metodología.....	110
3.3.3 Selección de herramientas tecnológicas.....	114
3.3.4 Desarrollo modular del sistema.....	115
3.3.5 Implementación de funciones de recomendación y predicción	118
3.4 Pruebas del sistema	128

3.4.1 Estrategia de pruebas.....	129
3.4.2 Pruebas de carga masiva e integridad de datos	130
3.4.3 Pruebas de integración del registro individual	131
3.4.4 Pruebas smoke y end-to-end parcial	132
3.4.5 Pruebas de reportería, ETL y exportación	133
3.4.6 Pruebas de carga funcional y rendimiento operativo	134
3.4.7 Consideraciones sobre pruebas unitarias	135
3.4.8 Resultados generales de las pruebas	135
3.5 Resultados obtenidos.....	136
3.5.1 Resultados relacionados con el análisis del estado actual de la gestión informativa	137
3.5.2 Resultados relacionados con el desarrollo del sistema BI e IA	138
3.5.3 Resultados relacionados con la evaluación del desempeño del sistema	140
3.5.4 Métricas del sistema.....	145
3.6 Implantación del sistema.....	148
3.6.1 Estrategia de implantación	148
3.6.2 Verificación del despliegue.....	151
3.6.3 Resultado de la implantación	152
3.6.4 Escalabilidad y mantenimiento	153
CAPÍTULO IV. CONCLUSIONES Y RECOMENDACIONES	155
4.1 Conclusiones	155

4.2 Recomendaciones.....	159
REFERENCIAS BIBLIOGRÁFICAS	161
ANEXOS	167

ÍNDICE DE TABLAS

Tabla 1. Resultados obtenidos por base de datos y cadenas de búsqueda.	4
Tabla 2. Criterios de inclusión y exclusión.	5
Tabla 3. Resultados obtenidos tras aplicar criterios de inclusión y exclusión.	5
Tabla 4. Criterios de pertinencia aplicados.	6
Tabla 5. Resultados obtenidos tras aplicar los criterios de pertinencia.	6
Tabla 6. Problemáticas identificadas durante el proceso de levantamiento de información.	37
Tabla 7. Necesidades institucionales identificadas para el desarrollo del sistema. ...	38
Tabla 8. Requerimientos funcionales del sistema.	42
Tabla 9. Requerimientos no funcionales del sistema.	44
Tabla 10. Tecnologías utilizadas en la propuesta tecnológica.	79
Tabla 11. Agrupación funcional de entidades del modelo transaccional OLTP.	82
Tabla 12. Entidades principales y propósito funcional dentro del modelo OLTP.	89
Tabla 13. Relaciones funcionales principales del modelo transaccional OLTP.	90
Tabla 14. Tablas de hechos y dimensiones de la constelación dimensional OLAP institucional.	98
Tabla 15. Procedimientos y componentes utilizados dentro del proceso ETL institucional.	102
Tabla 16. Procesos de transformación y validación aplicados durante el ETL.	102
Tabla 17. Vistas analíticas implementadas para reportería institucional.	104
Tabla 18. Indicadores institucionales soportados por el entorno BI.	105

Tabla 19. Criterios de evaluación para la selección de la metodología de desarrollo.	108
Tabla 20. Resultados de la matriz ponderada de metodologías ágiles.	109
Tabla 21. Roles adaptados para la aplicación de Scrum + XP.	111
Tabla 22. Plan resumido de sprints aplicados al desarrollo del sistema.	111
Tabla 23. Prácticas XP aplicadas durante el desarrollo.	113
Tabla 24. Criterios resumidos de Definition of Done aplicados al proyecto.	113
Tabla 25. Herramientas tecnológicas utilizadas durante el desarrollo del sistema. .	115
Tabla 26. Módulos desarrollados dentro del sistema.	116
Tabla 27. Funciones principales del módulo de Inteligencia Artificial.	119
Tabla 28. Algoritmos utilizados en el módulo IA.	121
Tabla 29. Métricas de evaluación del modelo predictivo.	122
Tabla 30. Indicadores utilizados para calcular la preparación del módulo IA.	123
Tabla 31. Fórmulas principales utilizadas en el módulo IA.	124
Tabla 32. Tipos de pruebas aplicadas al sistema.	129
Tabla 33. Casos principales de pruebas de carga masiva e integridad.	130
Tabla 34. Pruebas de integración del registro individual.	131
Tabla 35. Pruebas smoke aplicadas al sistema.	132
Tabla 36. Pruebas de reportería y exportación.	133
Tabla 37. Datasets utilizados en pruebas de carga funcional.	134
Tabla 38. Resumen general de resultados de pruebas.	136

Tabla 39. Resultados asociados al análisis del contexto institucional.	137
Tabla 40. Componentes desarrollados como resultado del sistema propuesto.	138
Tabla 41. Evidencia funcional del módulo de Inteligencia Artificial.	139
Tabla 42. Resultados generales de pruebas aplicadas al sistema.	141
Tabla 43. Base de medición para la reducción estimada de tiempos de reportería..	143
Tabla 44. Tiempos de reportería por tipo de operación.	143
Tabla 45. Resultados de la entrevista estructurada post-implantación basada en dimensiones CSUQ.	144
Tabla 46. Métricas generales del sistema desarrollado.	145
Tabla 47. Trazabilidad entre objetivos específicos, evidencias y resultados alcanzados.	147
Tabla 48. Información operativa no sensible del entorno de implantación	150
Tabla 49. Verificaciones realizadas durante la implantación del sistema.....	152

ÍNDICE DE FIGURAS

Figura 1. Frecuencia de metodologías utilizadas en estudios sobre BI e IA.....	19
Figura 2. Frecuencia de modelos implementados en los estudios analizados.....	20
Figura 3. Principales algoritmos de IA y Aprendizaje Automático utilizados.	21
Figura 4. Frecuencia de herramientas tecnológicas y plataformas de apoyo empleadas en los estudios.	22
Figura 5. Tipos y porcentajes de uso de fuentes de datos utilizados.....	23
Figura 6. Promedio por ítem de la entrevista estructurada.....	31
Figura 7. Porcentaje de aceptación por ítem de la entrevista estructurada.	32
Figura 8. Distribución general de respuestas por escala Likert.	33
Figura 9. Promedio por dimensión evaluada en la entrevista estructurada.....	33
Figura 10. Flujo del procesamiento y análisis de la información para la definición de requerimientos del sistema.	36
Figura 11. Mapa funcional general del sistema propuesto.....	41
Figura 12. Diagrama de casos de uso general del sistema propuesto.	47
Figura 13. Diagrama de casos de uso del módulo de registro de artículos científicos.	49
Figura 14. Diagrama de casos de uso del módulo de carga masiva de información..	51
Figura 15. Diagrama de casos de uso del módulo de workflow y revisión institucional.	54
Figura 16. Diagrama de casos de uso del módulo de integración con APIs externas.	56
Figura 17. Diagrama de casos de uso del módulo de reportería y Business Intelligence.	58

Figura 18. Diagrama de casos de uso del módulo de administración y seguridad.....	60
Figura 19. Diagrama de casos de uso del módulo de Inteligencia Artificial para recomendación y predicción.	62
Figura 20. Diagrama de actividades del proceso de registro de artículos científicos.	64
Figura 21. Diagrama de actividades del proceso de carga masiva de información. ..	65
Figura 22. Diagrama de actividades del workflow de revisión institucional.	66
Figura 23. Diagrama de actividades del proceso de consulta mediante APIs externas.	67
Figura 24. Diagrama de actividades del proceso de generación de reportes y ETL. .	68
Figura 25. Diagrama de actividades del proceso de recomendación y predicción mediante Inteligencia Artificial.....	69
Figura 26. Diagrama de secuencia del proceso de registro de artículos científicos...	71
Figura 27. Diagrama de secuencia del proceso de carga masiva de información.....	72
Figura 28. Diagrama de secuencia del proceso de consulta mediante APIs externas.	73
Figura 29. Diagrama de secuencia del workflow institucional.	74
Figura 30. Diagrama de secuencia del proceso de generación de dashboards y reportes.	75
Figura 31. Diagrama de secuencia del proceso de recomendación y predicción mediante Inteligencia Artificial.....	76
Figura 32. Arquitectura general del sistema basada en Clean Architecture [48].....	79
Figura 33. Flujo general de la arquitectura del sistema.	80
Figura 34. Esquema general del modelo transaccional OLTP.	82
Figura 35. Arquitectura funcional del modelo transaccional OLTP extensible.....	85

Figura 36. Estructura funcional del modelo basado en metadatos y componentes dinámicos.	86
Figura 37. Arquitectura funcional de staging, workflow y trazabilidad institucional.	87
Figura 38. Integración del modelo OLTP con procesos ETL y Business Intelligence.	88
Figura 39. Relación funcional de entidades principales del modelo transaccional OLTP.....	89
Figura 40. Estrategias estructurales aplicadas en el modelo transaccional OLTP.....	92
Figura 41. Arquitectura general del modelo analítico OLAP institucional.	95
Figura 42. Modelo conceptual tipo constelación del Data Warehouse institucional.	96
Figura 43. Arquitectura general del proceso ETL institucional.	99
Figura 44. Arquitectura operativa del proceso ETL institucional.....	101
Figura 45. Arquitectura funcional del módulo de Inteligencia Artificial.....	126

ÍNDICE DE ANEXOS

Anexo 1. Entrevista estructurada inicial con escala tipo Likert.....	167
Anexo 2. Entrevista semiestructurada con preguntas abiertas.	169
Anexo 3. Entrevista estructurada post-implantación basada en dimensiones CSUQ.	171
Anexo 4. Product Backlog.	174
Anexo 5. Plan de sprints y cronograma de desarrollo.....	177
Anexo 6. Definition of Done (DoD).	178
Anexo 7. Plan de aseguramiento de calidad del software.	179
Anexo 8. Casos de prueba de carga masiva e integridad de datos.	183
Anexo 9. Matriz de tiempos de reportería.....	186
Anexo 10. Selección de metodología Scrum + XP.	189
Anexo 11. Diseño de sistema.	195
Anexo 12. Informe técnico del módulo de Inteligencia Artificial.	210
Anexo 13. Informe del módulo de Inteligencia Artificial.	215
Anexo 14. Modelo dimensional tipo constelación implementado en el Data Warehouse institucional.	220

RESUMEN EJECUTIVO

El presente trabajo de titulación tuvo como tema la implementación de un sistema de reportería y visualización de datos con funciones de recomendación y predicción, utilizando técnicas de Business Intelligence e Inteligencia Artificial, en la Dirección de Investigación y Desarrollo de la Universidad Técnica de Ambato. La problemática identificada se relacionó con la gestión manual de información científica, trazabilidad y centralización de los datos. La investigación se desarrolló bajo una modalidad aplicada, de campo y no experimental, considerando como población a cinco funcionarios vinculados con los procesos de gestión informativa. Para el diagnóstico se empleó una entrevista estructurada con escala tipo Likert y una entrevista semiestructurada complementaria, permitiendo identificar ocho problemáticas y diez necesidades institucionales.

Como resultado, se desarrolló una plataforma web basada en Scrum + XP, integrada por gestión de artículos científicos, carga masiva, workflow institucional, integración externa, procesos ETL, modelo OLTP extensible, Data Warehouse tipo constelación, dashboards y reportes exportables. Además, se implementó un módulo de IA con predicción mensual de producción científica, predicción por facultad y línea de investigación, recomendaciones de calidad de datos, recomendaciones editoriales y análisis de preparación IA mediante indicadores ponderados. La evaluación evidenció 80 operaciones de reportería exitosas, tasa de éxito del 100 %, tiempo promedio de 10,53 segundos y reducción estimada del 99,66 %. La entrevista estructurada post-implantación basada en CSUQ obtuvo 4,59/5, aceptación del 100 % y Alfa de Cronbach de 0,973. Se concluye que el sistema fortalece la gestión informativa y apoya la toma de decisiones basada en datos.

Palabras claves: Business Intelligence, inteligencia artificial, reportería, Data Warehouse, toma de decisiones, producción científica.

ABSTRACT

This thesis focused on the implementation of a data reporting and visualization system with recommendation and prediction functions, using Business Intelligence and Artificial Intelligence techniques, at the Research and Development Directorate of the Technical University of Ambato. The identified problem was related to the manual management of scientific information, data traceability, and centralization. The research was conducted using an applied, field-based, and non-experimental approach, considering five staff members involved in information management processes as the population. A structured interview with a Likert scale and a complementary semi-structured interview were used for the diagnostic process, allowing the identification of eight problems and ten institutional needs.

As a result, a web platform based on Scrum + XP was developed, integrating scientific article management, bulk upload, institutional workflow, external integration, ETL processes, an extensible OLTP model, a constellation-type data warehouse, dashboards, and exportable reports. Furthermore, an AI module was implemented with monthly scientific output forecasting, forecasting by faculty and research area, data quality recommendations, editorial recommendations, and AI readiness analysis using weighted indicators. The evaluation showed 80 successful reporting operations, a 100% success rate, an average time of 10.53 seconds, and an estimated reduction of 99.66%. The post-implementation structured interview based on CSUQ received a score of 4.59/5, 100% acceptance, and a Cronbach's alpha of 0.973. It is concluded that the system strengthens information management and supports data-driven decision-making.

Keywords: Business Intelligence, artificial intelligence, reporting, Data Warehouse, decision-making, scientific production.

CAPÍTULO I. MARCO TEÓRICO

1.1 Tema de investigación

IMPLEMENTACIÓN DE UN SISTEMA DE REPORTERÍA Y VISUALIZACIÓN DE DATOS CON FUNCIONES DE RECOMENDACIÓN Y PREDICCIÓN, UTILIZANDO TÉCNICAS DE BUSINESS INTELLIGENCE (BI) E INTELIGENCIA ARTIFICIAL (IA), EN LA DIRECCIÓN DE INVESTIGACIÓN Y DESARROLLO (DIDE) DE LA UNIVERSIDAD TÉCNICA DE AMBATO (UTA),

1.1.1. Planteamiento del problema

En la era digital, las organizaciones buscan ventajas competitivas mediante el uso estratégico de la tecnología [1]. Herramientas como BI, junto con tecnologías emergentes como la IA, big data e Internet of Things (IoT), se han consolidado como pilares clave para optimizar procesos, gestionar recursos y mejorar la toma de decisiones [1]. En [2], estas tecnologías permiten centralizar y analizar grandes volúmenes de información, facilitando la generación de modelos predictivos eficaces. Sin embargo, su adopción exitosa depende de factores como la capacidad de almacenamiento, procesamiento, analítica y la calidad de los sistemas implementados [3, 4]. En países en vías de desarrollo, muchas organizaciones enfrentan barreras tecnológicas como paneles poco intuitivos, falta de automatización en la reportería y escasa integración de soluciones avanzadas, lo que limita su competitividad [5].

En [6] se señala que la integración de sistemas como Customer Relationship Management (CRM), Enterprise Resource Planning (ERP), Business Intelligence (BI) y técnicas de Inteligencia Artificial (IA) en entornos de comercio electrónico mejora significativamente la eficiencia empresarial. Asimismo, la incorporación de sistemas de recomendación y predicción, junto con una adecuada aplicación de BI, permite optimizar la asignación de recursos y realizar una evaluación más precisa de la segmentación del mercado. El estudio se desarrolló con una muestra de 30 consumidores mensuales, cuyos aportes contribuyeron positivamente al desarrollo de las pequeñas y medianas empresas (PYMEs). Además, este enfoque resulta

extrapolable al ámbito académico, donde la toma de decisiones basada en datos se ha vuelto un componente esencial.

Así, un análisis bibliométrico sobre el uso de BI e IA en Instituciones de Educación Superior (IES) reveló un crecimiento sostenido del interés académico, con 755 publicaciones en Scopus entre 2019 y 2023, evidenciando un crecimiento anual del 9,9%, alcanzando 178 artículos en 2023, con una edad promedio de 2,86 años y 11,57 citas por documento, lo que indica una rápida evolución y consolidación científica. Asimismo, las aplicaciones más relevantes identificadas fueron la minería de datos educativos, la analítica del aprendizaje y los modelos predictivos, empleando técnicas como árboles de decisión, redes neuronales y regresión logística [7]. En [8], la predicción del rendimiento académico alcanzó una precisión del 96% mediante un algoritmo basado en el análisis del comportamiento de entrega de tareas. En universidades peruanas, la implementación de un sistema BI mejoró significativamente los tiempos de respuesta en la toma de decisiones académicas mediante el uso de datamarts y procesos ETL (Extract, Transform, Load) [9]. Esta evidencia sugiere que las tecnologías de BI e IA no solo optimizan procesos y recursos, sino que también contribuyen a mejorar la retención estudiantil, la personalización del aprendizaje y la calidad institucional. Sin embargo, se observa una baja representación geográfica en Sudamérica, lo que subraya la necesidad de impulsar investigaciones contextualizadas que favorezcan una adopción efectiva en la región.

Pese al reconocimiento de la inteligencia empresarial como herramienta estratégica en el ámbito corporativo, su adopción en IES del país continúa siendo limitada y desigual. En estas organizaciones, donde la gestión de datos es esencial para el monitoreo de proyectos, la evaluación de indicadores académicos y la toma de decisiones en áreas de investigación y desarrollo, la ausencia de soluciones integradas de análisis y visualización representa un obstáculo significativo [5, 7]. A esto se suma la presencia de factores como la Ansiedad Tecnológica (AT), definida como el temor o resistencia frente a nuevas tecnologías, la cual ha demostrado tener un impacto negativo en la adopción de sistemas BI [10]. Esta problemática es especialmente relevante en contextos de países en desarrollo, donde se ha identificado que los profesionales tienden a evitar la incertidumbre y presentan niveles elevados de AT, afectando la implementación efectiva de sistemas tecnológicos [11]. Como consecuencia, muchas

unidades académicas optan por seguir utilizando herramientas tradicionales como hojas de cálculo, formularios físicos o reportes manuales, en lugar de adoptar plataformas inteligentes que integren inteligencia artificial y analítica avanzada. Este fenómeno no solo limita la eficiencia operativa, sino que también retrasa la transformación digital de las instituciones, comprometiendo su capacidad para generar conocimiento, tomar decisiones oportunas y competir en un entorno educativo cada vez más dinámico y orientado por los datos.

En el caso particular de la DIDE - UTA, se identifica la necesidad de fortalecer la gestión de la información relacionada con la producción científica institucional. En la actualidad, la inexistencia de un sistema integrado limita la centralización, organización, consulta y análisis oportuno de los datos requeridos para el seguimiento de artículos científicos y la generación de reportes estratégicos. Esta situación provoca una alta dependencia de procesos manuales, registros dispersos y mecanismos poco automatizados, lo que dificulta la obtención de información confiable, actualizada y útil para la toma de decisiones. De igual manera, la ausencia de herramientas de visualización interactiva, reportería automatizada, cargas masivas de información y estructuras configurables restringe la capacidad institucional para responder de forma ágil a las necesidades de control, evaluación y planificación. A ello se suma la falta de integración con fuentes externas de información científica, lo que reduce las posibilidades de enriquecimiento, validación y aprovechamiento analítico de los datos disponibles. Como consecuencia, los procesos de gestión investigativa presentan limitaciones que afectan la eficiencia operativa y la calidad del análisis realizado dentro de la DIDE. Frente a este escenario, resulta pertinente plantear el desarrollo de una solución tecnológica basada en técnicas de Business Intelligence e Inteligencia Artificial que permita optimizar la gestión informativa y fortalecer el soporte a las decisiones estratégicas. En este contexto, surge la siguiente pregunta de investigación: ¿de qué manera la implementación de un sistema de reportería y visualización de datos con funciones de recomendación y predicción puede optimizar la toma de decisiones estratégicas en la Dirección de Investigación y Desarrollo de la Universidad Técnica de Ambato?

1.2 Antecedentes investigativos

La recopilación de antecedentes investigativos se realizó a través de diversas bases de datos científicas reconocidas, tales como ScienceDirect, ACM Digital Library, Heliyon, IEEE Xplore y Wiley. Estas fuentes permitieron obtener información actual, relevante y validada sobre la implementación de soluciones de inteligencia de negocios (BI) e inteligencia artificial (IA) aplicadas a la gestión institucional. Cada una de estas bases de datos se destaca por su rigor en la publicación de estudios revisados por pares en áreas de ciencias aplicadas y tecnología, lo que garantiza la confiabilidad de la información recopilada. La revisión de estos antecedentes facilitó la construcción de un marco teórico sólido, la identificación de enfoques vigentes y el reconocimiento de tendencias recientes en el uso de herramientas BI e IA para la visualización de datos, reportería dinámica, sistemas de recomendación y modelos predictivos aplicados al proceso de toma de decisiones estratégicas en entornos académicos y organizacionales. Para la búsqueda se utilizaron las siguientes cadenas de consulta, y se muestran los resultados obtenidos en la Tabla 1.

- ("Business Intelligence") AND ("data mining") AND (("institutions") OR ("organizations") OR ("university"))
- ("Data Science") AND ("dashboards") AND ("decision support") AND (("artificial intelligence") OR ("machine learning")) AND (("recommendation systems") OR ("predictive models"))

Tabla 1. Resultados obtenidos por base de datos y cadenas de búsqueda.

Base de datos	Cantidad	Cadena de búsqueda
ACM Digital Library	1079	Cadena 1
ACM Digital Library	165	Cadena 2
Heliyon	46	Cadena 1
Heliyon	1	Cadena 2
IEEE Xplore	1333	Cadena 1
IEEE Xplore	0	Cadena 2
ScienceDirect	3538	Cadena 1
ScienceDirect	185	Cadena 2

Wiley	1367	Cadena 1
Wiley	52	Cadena 2
Total	7766	

Para realizar un primer filtro de la información recopilada, se definieron criterios de inclusión y exclusión con el fin de asegurar la pertinencia y calidad de los estudios seleccionados. En la Tabla 2 se detallan los criterios establecidos, los cuales permitieron delimitar el alcance de la revisión hacia investigaciones centradas en el uso de inteligencia de negocios, ciencia de datos, y tecnologías de predicción y recomendación en entornos institucionales. Se priorizó la selección de estudios recientes y empíricos, garantizando la actualidad y el sustento metodológico de los hallazgos. Asimismo, se enfocó la búsqueda en investigaciones que abordaran la implementación de dashboards, reportería automatizada y apoyo a la toma de decisiones mediante IA. Este proceso permitió reducir significativamente el volumen de resultados iniciales, seleccionando un conjunto más preciso y relevante para la investigación. En la Tabla 3 se presentan los estudios previos resultantes de esta depuración.

Tabla 2. Criterios de inclusión y exclusión.

Criterios
Inclusión
1. El estudio se publicó entre 2020 y 2025.
2. El estudio está escrito en inglés.
3. El estudio utiliza herramientas y metodologías para BI.
4. El texto completo del estudio se encuentra disponible.
5. El estudio realiza o aplica BI, IA, modelos, etc.
Exclusión
1. El estudio es un estudio secundario, como una revisión bibliográfica o un artículo teórico.
2. El estudio no presenta un estudio empírico (solo metaanálisis).
3. El estudio no está disponible en su totalidad.

Tabla 3. Resultados obtenidos tras aplicar criterios de inclusión y exclusión.

Base de datos	Cantidad	Cadena de búsqueda
ACM Digital Library	6	Cadena 1
ACM Digital Library	7	Cadena 2
Heliyon	1	Cadena 1

Heliyon	0	Cadena 2
IEEE Xplore	1	Cadena 1
IEEE Xplore	0	Cadena 2
ScienceDirect	78	Cadena 1
ScienceDirect	46	Cadena 2
Wiley	1	Cadena 1
Wiley	4	Cadena 2
Total	144	

Los criterios de pertinencia se aplicaron como un segundo filtro con el objetivo de verificar el aporte real de los estudios seleccionados a la presente investigación. En la Tabla 4 se presentan los criterios definidos, asegurando que los estudios no solo sean relevantes, sino que también contribuyan de forma significativa al análisis de herramientas BI, dashboards, sistemas de recomendación, y modelos predictivos aplicados a contextos institucionales. Al aplicar estos criterios, se dio prioridad a investigaciones que ofrecieran resultados concretos, metodologías replicables y aplicaciones prácticas donde se demuestra los criterios aplicados mostrados en la Tabla 5.

Tabla 4. Criterios de pertinencia aplicados.

Pertinencia
1. Propone una metodología, modelo o enfoque aplicado.
2. Expone las técnicas principales o algoritmos utilizados.
3. Presenta herramientas tecnológicas, plataformas o tecnologías de BI o IA.
4. Evidencia la inclusión o exclusión de módulos de reportería dinámica o dashboards.
5. Muestra el uso de visualizaciones para facilitar el análisis de información.
6. Presenta el sector al que se realiza el estudio.

Tabla 5. Resultados obtenidos tras aplicar los criterios de pertinencia.

Base de datos	Cantidad	Cadena
Science Direct	10	Cadena 1
Science Direct	7	Cadena 2
IEEE Xplore	1	Cadena 1
IEEE Xplore	0	Cadena 2
Heliyon	2	Cadena 1
Heliyon	0	Cadena 2

ACM Digital Library	0	Cadena 2
ACM Digital Library	0	Cadena 1
Wiley	0	Cadena 1
Wiley	0	Cadena 2
Total	20	

El resultado de la revisión permitió recopilar información relevante y actualizada, alineada con los criterios de inclusión, exclusión y pertinencia definidos en un total de 20 resultados. Esta evidencia respalda la necesidad de implementar soluciones basadas en BI e IA para optimizar la toma de decisiones en contextos institucionales, especialmente en el ámbito educativo. Los antecedentes investigativos seleccionados constituyen un soporte fundamental para el presente estudio, al aportar un marco teórico robusto y evidenciar la efectividad de herramientas como dashboards interactivos, reportería dinámica, sistemas de recomendación y modelos predictivos.

El artículo [6] muestra que las PYMEs orientadas a proyectos carecen de herramientas de BI y Project Management Office (PMO) accesibles y adaptadas a su tamaño. La investigación desarrolla un marco conceptual que emplea Microsoft Excel como base de datos interna (data.xlsx) y Microsoft Power BI para dashboards dinámicos de KPIs, complementado con una red neuronal entrenada en PyTorch que sugiere automáticamente el mejor modelo de datos predefinido según las entidades disponibles (clientes, proyectos, tareas, hitos, riesgos). Asimismo, se muestra un chatbot que asiste en la recolección guiada de datos donde facilitan el uso de las herramientas aún sin conocimientos técnicos en BI. Al sistema realizado, muestra una solución low-cost de fácil implementación permitiendo a los mandos medios y alta dirección manejar información estandarizada sobre tiempo, costos, calidad y riesgos sin necesidad de software especializado. Además, en cuanto a la metodología y Kpis se utilizó el “triángulo de hierro” (tiempo, presupuesto, alcance/calidad): entrega a tiempo, horas planificadas vs reales, valor ganado, índice de desempeño de cronograma, rentabilidad, márgenes, utilización de recursos, matriz de riesgo, entre otros. Como resultado, se muestra que el marco reduce costos, agiliza la toma de decisiones y mejora la gestión estratégica, mostrando un alto potencial para hacer viable la gestión de cartera de proyectos en PYMEs con escasos recursos.

El artículo [19] muestra un estudio realizado en 15 parques industriales de la provincia de Henan (China) donde se investigó cómo las pequeñas y medianas empresas manufactureras pueden mitigar los efectos de la incertidumbre del entorno competitivo y mejorar su rendimiento mediante el uso de BI. El estudio muestra la aplicación de una encuesta cuantitativa estructurada aplicada a 429 participantes, estas analizaron las relaciones entre la intensidad competitiva, la eficiencia de los procesos internos, las capacidades de inteligencia empresarial y el rendimiento organizacional. La investigación se desarrolló bajo el modelo Tecnología-Organización-Entorno (TOE) y en la Teoría de las Capacidades Dinámicas, mostrando así factores tecnológicos, organizacionales y ambientales para explicar la adaptación estratégica de las PYMES. En cuestión de análisis de datos utilizaron herramientas estadísticas y de ciencia de datos como STATA 15.0 y MPLUS 7.4, aplicando técnicas como el Análisis Factorial Confirmatorio (AFC), el uso del estimador CLC, y la implementación de modelos de ecuaciones estructurales. Asimismo, los principales KPIs que se evaluaron incluyen la eficiencia interna, la capacidad de detección de información de mercado y la capacidad de respuesta estratégica. Como resultado, se confirman que la capacidad de inteligencia empresarial impacta positivamente en el rendimiento de las empresas, mediada por la eficiencia en los procesos internos, lo que sugiere que la analítica y el conocimiento del entorno competitivo son claves para la resiliencia y el éxito de las PYMES manufactureras ante contextos inciertos.

El artículo [20] expone el modelo de éxito de sistemas de información de DeLone & McLean para analizar el uso BI por contadores gerenciales, diferenciando entre uso rutinario (acceso a informes predefinidos, consultas básicas) y uso avanzado (analítica predictiva, visualización interactiva, creación de dashboards). Mediante una encuesta a 362 contadores en Sudáfrica (usuarios de SAP BusinessObjects, Power BI y Oracle), se identificó que la calidad del sistema (usabilidad, estabilidad), calidad de datos (integridad, completitud) y calidad del servicio (soporte técnico) son determinantes clave para ambos tipos de uso, aunque el impacto es mayor en el uso avanzado ($\beta = 0.59$ para calidad de información vs. 0.20 en uso rutinario). La autoeficacia del usuario y la complejidad de tareas (ej.: análisis de varianzas, pronósticos) mostraron influencia significativa solo en el uso avanzado, mientras que la satisfacción del usuario correlacionó principalmente con este último. Los resultados revelaron que el uso

avanzado mejora significativamente el desempeño individual ($\beta = 0.68$ vs. 0.26 del uso rutinario), especialmente en entornos de alta complejidad (efecto moderador: $*p* < 0.001$). El estudio destaca la necesidad de capacitación en funciones avanzadas de BI y la optimización de la calidad de datos para maximizar el ROI en sistemas de inteligencia de negocios.

El artículo [21] demuestra cómo la implementación de un sistema de soporte a decisiones, diseñado bajo el enfoque metodológico Design Science Research (DSR), permite optimizar la gestión operativa de terminales portuarias mediante el uso de herramientas de Business Intelligence. Se puede evidenciar componentes esenciales como un Data Warehouse, procesos de ETL (Extract, Transform, Load), dashboards interactivos, reportes dinámicos y motores analíticos, todo integrado mediante tecnologías como Microsoft SQL Server, Power BI, ASP.NET, y servicios web para el consumo de datos en tiempo real. Entre los KPIs principales considerados se encuentran: tiempo medio de espera de vehículos, productividad de equipos de carga y descarga, eficiencia de los procesos logísticos, y cumplimiento de tiempos de operación. Para los datos históricos se utilizaron en tiempo real, lo cual facilita el análisis descriptivo, predictivo y prescriptivo en las operaciones portuarias. Como resultado de la investigación se confirma que el uso integrado de tecnologías BI, herramientas de visualización, y analítica de datos permite fortalecer la toma de decisiones tácticas y estratégicas, siendo altamente replicable en contextos como el comercio electrónico, mantenimiento de infraestructura tecnológica o soporte técnico organizacional.

El artículo [22] muestra cómo los sistemas BI actúan como catalizadores de innovación organizacional, particularmente en las PYMEs, aplicando un enfoque metodológico riguroso apoyado en herramientas tecnológicas y estadísticas avanzadas. Se utilizaron herramientas como SmartPLS 4.1.2 para el modelado de ecuaciones estructurales mediante la técnica PLS-SEM (Partial Least Squares – Structural Equation Modeling), ideal para tratar datos no normales y modelos complejos con múltiples indicadores, IBM SPSS 23 para análisis descriptivos y gráficos, MS Excel para el manejo preliminar de datos, y Typeform como plataforma de recolección de datos mediante encuestas web dirigidas a 150 altos directivos y gerentes. A nivel estadístico, se aplicaron análisis factoriales confirmatorios (AFC), pruebas de confiabilidad y validez

convergente/discriminante, así como algoritmos como el PLS para estimación de pesos y cargas, bootstrapping con 5000 muestras para determinar la significancia de relaciones, y los criterios Fornell-Larcker y HTMT para evaluar la validez discriminante. Además, los coeficientes R^2 y Q^2 permitieron medir el poder explicativo y la relevancia predictiva del modelo estructural mencionado.

Además, el estudio indica que herramientas como data warehousing, OLAP, minería de datos, Big Data y sistemas de información ejecutiva son utilizadas para transformar datos históricos y actuales en conocimiento aplicable a decisiones estratégicas, tácticas y operativas. Se muestra que estas tecnologías impulsan mejoras en eficiencia operativa, diferenciación competitiva, segmentación de clientes, velocidad de respuesta, y generación de valor agregado. A partir de los resultados, el estudio plantea un modelo conceptual conformado por cuatro dimensiones clave: (1) liderazgo estratégico digital, (2) infraestructura tecnológica de BI, (3) repositorios de datos como base del conocimiento organizacional, y (4) cultura organizacional orientada a la innovación. El estudio concluye que el éxito en la adopción de BI no radica solo en la tecnología, sino en su integración estratégica, el compromiso de la alta dirección, y el desarrollo de capacidades analíticas que permitan traducir datos en decisiones efectivas y sostenibles para la innovación.

El artículo [23] muestra cómo las IES enfrentan el desafío de gestionar grandes volúmenes de datos para asegurar la calidad educativa, proponiendo el uso de tableros BI como solución estratégica. Para ello, las herramientas que se incluyeron fueron las de autoservicio (SSBI) como Power BI, Tableau, QlikView y Birst, así como tecnologías orientadas a servicios (SoBI) como Cube.js, complementadas con tecnologías web como PHP y HTML, permitiendo la integración de dashboards interactivos en entornos accesibles. La técnica que aplica y menciona fue la de arquitecturas adaptativas que incluyen modelos tradicionales basados en almacenes de datos (DW), orientados a servicios (SoBI) y autoservicio (SSBI), seleccionados en función de los recursos técnicos y presupuestarios de cada institución. A través de estas arquitecturas, se visualizan y monitorean indicadores clave de rendimiento (KPIs) relacionados con la calidad académica, el cumplimiento de estándares de acreditación y el rendimiento institucional. Se muestra además la capacidad de estas herramientas para ofrecer análisis personalizados en tiempo real, sin depender exclusivamente de

personal técnico especializado. Para el desarrollo, realizaron prototipos y combinándolos con visualizadores como Power BI y motores de extracción como Cube.js, evidencian un enfoque flexible y escalable. Como resultado, les permitieron validar que la implementación de BI en el ámbito educativo mejora la toma de decisiones estratégicas y facilita el cumplimiento de estándares de calidad.

El artículo [24] muestra un enfoque integral para abordar la pérdida de clientes en el sector bancario, mediante el uso de técnicas de Machine Learning (ML) y el desarrollo de una aplicación interactiva para la visualización de resultados y toma de decisiones. La investigación incluye modelos de clasificación como XGBoost, Random Forest, regresión logística y Support Vector Machine (SVM), aplicados sobre un dataset de 10,000 clientes, además, se utilizó la técnica Synthetic Minority Oversampling Technique (SMOTE) para balancear las clases desbalanceadas durante el preprocesamiento. Como soporte tecnológico, desarrollaron una aplicación en RShiny que permite al usuario interactuar con pestañas de análisis de datos, predicción y evaluación de modelos, incorporando además gráficos dinámicos con Plotly y procesamiento con Python. El artículo muestra como la plataforma facilita la exploración de variables relevantes para la pérdida de clientes, como género, saldo, número de productos contratados o nacionalidad, y permite exportar los reportes generados para la gestión operativa. Entre los KPIs utilizados para evaluar los modelos destacan: precisión, sensibilidad, especificidad, AUC y F1 Score, donde el modelo XGBoost obtuvo el mejor rendimiento. Como resultado, el desarrollo demuestra la eficacia de integrar herramientas de ML con entornos visuales interactivos para apoyar decisiones estratégicas en entornos corporativos.

El artículo [25] un estudio PERISCOPE que desarrolló y validó modelos de inteligencia artificial para la detección temprana de infecciones postoperatorias utilizando datos retrospectivos de historias clínicas electrónicas, abarcando más de 252.000 procedimientos quirúrgicos entre 2014 y 2023. Se usó un modelo principal, el XGBoost, un algoritmo de aprendizaje automático basado en potenciación de gradiente, seleccionado por su capacidad de manejar grandes volúmenes de datos y relaciones no lineales complejas. Se implementaron procesos de validación cruzada estratificada (10 veces por hospital), así como particiones temporales en tres hospitales (A, B y C) para validar la estabilidad del rendimiento a lo largo del tiempo,

evidenciando una evaluación rigurosa en contextos multicéntricos. Para la calibración y ajuste de hiperparámetros los realizaron utilizando Python 3.8. El enfoque que llevaron a cabo se contrastó con herramientas tradicionales como ACS NSQIP (regresión logística clásica) y sistemas más avanzados como MySurgeryRisk, que emplea modelos de machine learning como random forests y redes neuronales, resaltando así las ventajas del enfoque propuesto: aprendizaje continuo, adaptabilidad a nuevos datos y mayor precisión al modelar patrones no lineales. Como resultado, el estudio demuestra cómo la integración de IA con historias clínicas electrónicas puede ofrecer una solución escalable y precisa para la predicción clínica postoperatoria.

En artículo [26] empleó técnicas de BI, como ETL para integrar datos de ERP, repositorios internos y fuentes gubernamentales, y On-Line Analytical Processing (OLAP) para análisis multidimensional, apoyado en dashboards interactivos desarrollados en Microsoft Power BI (Power Query para ETL, Power Pivot para modelado y Power View para visualización). En cuanto a la arquitectura se siguió un modelo con un data warehouse de 32 tablas interrelacionadas que monitoreaban KPIs clave: niveles de presas, eficiencia en requisiciones de materiales (RMS), cumplimiento de seguridad radiológica y métricas financieras (presupuesto vs. real, rentabilidad). La metodología fue investigación-acción, con ciclos iterativos entre marzo y julio de 2022, donde ingenieros y gerentes validaron los dashboards. Además, el proyecto no incluyó IA avanzada, sentó las bases para su integración futura, destacando el potencial de herramientas como PyTorch para predecir riesgos u optimizar recursos. Como resultado se demostraron mejoras en la transparencia operativa y el cumplimiento regulatorio, el estudio sugiere ampliar la automatización con IoT y escalar el Business Intelligence Systems (BIS) a otras áreas.

El artículo [27] implementa un sistema de BI mediante técnicas de ETL (Extraer, Transformar, Cargar) para integrar datos estructurados y semiestructurados (CSV, JSON, bases de datos SQL) en un Data Warehouse diseñado bajo el modelo estrella, utilizando herramientas como Microsoft SQL Server (SSIS para ETL, SSAS para cubos OLAP) y Power BI para visualización interactiva. La metodología Design Science Research (DSR) lideró el desarrollo iterativo, validando el artefacto con métricas como KPIs de eficiencia operativa (tiempos de respuesta, rentabilidad) y calidad de datos. Los dashboards generados permiten análisis multidimensionales

(drill-down, slice-and-dice) sobre indicadores clave como costos, satisfacción de clientes y riesgos regulatorios, optimizando la toma de decisiones. Adicional, el estudio propone futura integración con PyTorch para modelos predictivos. Asimismo, la arquitectura, escalable y low-cost, destaca la importancia de la capacitación de usuarios y la alineación con estándares de Industry 4.0. Los resultados evidenciaron mejoras del 30% en accesibilidad de datos y un 25% en reducción de errores manuales. También, el enfoque combina tecnología accesible (Power BI), metodologías ágiles (DSR) y gobernanza de datos.

El artículo [28] demuestra el desarrollo de un dashboard de BI para laboratorios clínicos utilizando el paquete flexdashboard de R, una alternativa gratuita y personalizable frente a herramientas comerciales como Power BI o Tableau. Se presenta un tutorial detallado la creación de visualizaciones interactivas para monitorear métricas clave como el reporte de resultados críticos, incluyendo tiempos de notificación y cumplimiento de protocolos, utilizando datos simulados de 2167 casos. La arquitectura del dashboard se basa en componentes flexibles de R Markdown, permitiendo la integración de gráficos estáticos (ggplot2) e interactivos (plotly, dygraphs), tablas dinámicas (DT) y cuadros de valores con formato condicional. Los KPIs evaluados incluyen el porcentaje de resultados reportados a tiempo, volumen diario de casos y eficiencia por área clínica, con capacidades de filtrado por fecha, tipo de prueba y personal involucrado. La metodología enfatiza la reproducibilidad (código abierto en GitHub) y la automatización, aprovechando paquetes como tidyverse para el procesamiento de datos y lubridate para el análisis temporal. Como resultado, el enfoque está dirigido a profesionales con habilidades básicas en R, ofrece una solución escalable para laboratorios que requieren dashboards adaptables a regulaciones como CLIA o Joint Commission, combinando la potencia analítica de R con la accesibilidad de interfaces web responsivas.

El artículo [29] en el contexto de manufactura inteligente y sostenible, muestra como se ha desarrollado una metodología basada en inteligencia artificial (IA) que se estructura en fases: captura de datos, análisis y aprendizaje, toma de decisiones y evaluación. Para ello, emplearon modelos de aprendizaje automático supervisados y no supervisados, redes neuronales, clustering, y sistemas autónomos como los Digital Twins. Además, muestra el apoyo de la metodología utilizada en tecnologías como

sensores IoT, dispositivos OPC-UA y sistemas ciberfísicos (CPS), todos interconectados mediante arquitecturas Edge y Cloud Computing. Las herramientas incluyeron un toolkit de IA desarrollado por Tekniker, que permite desplegar módulos inteligentes para control, predicción y mantenimiento. Se utilizaron lenguajes como Python y C++, junto con servicios web y conectividad estándar para integrar con sistemas MES y tableros BI. Como resultado, las métricas evaluadas incluyen indicadores clave como tiempo de ciclo, OEE, eficiencia energética y disponibilidad de máquinas. Asimismo, la propuesta demuestra cómo integrar IA en entornos industriales para lograr sostenibilidad, autonomía y mejora continua de los procesos.

El artículo [30] presenta una metodología basada en inteligencia IA para impulsar la manufactura inteligente y sostenible, la estructuraron en cuatro etapas: adquisición de datos, análisis y aprendizaje, toma de decisiones autónoma, y evaluación continua. En cuanto a los algoritmos, muestran técnicas de aprendizaje automático supervisado y no supervisado, redes neuronales, modelos de clustering y regresión, así como análisis de series temporales. De acuerdo con lo expuesto, la arquitectura planteada en el estudio se basó en tecnologías de sensores IoT y dispositivos OPC-UA para la captura de datos en tiempo real, integrándose con sistemas ciberfísicos mediante arquitecturas de computación en la nube y en el borde (edge computing). Además, se implementaron Digital Twins para simular procesos industriales, facilitando la predicción de fallos, la optimización del rendimiento y el soporte a la toma de decisiones. En lo referente a herramientas, se utilizaron lenguajes como Python y C++, así como servicios web RESTful y conectividad estándar con software de planta tipo Manufacturing Execution Systems (MES). Asimismo, en el marco de la evaluación del sistema propuesto, se utilizaron diversos KPI, como el tiempo de ciclo, la eficiencia energética, la disponibilidad de los equipos, el Overall Equipment Effectiveness (OEE) y el número de paradas no planificadas. Finalmente, el estudio muestra la implementación de módulos de IA integrables, desarrollados con un toolkit propio que permite personalizar soluciones para diferentes procesos industriales.

El artículo [31] muestra el empleo algoritmos de aprendizaje automático como Random Forest, SVM (con kernels RBF, lineal y polinomial), KNN, regresión logística, AdaBoost, XGBoost y un clasificador de apilamiento, implementados en Python con librerías como Scikit-learn y Optuna. Para el preprocesamiento, el presente

estudio aplicó una codificación one-hot-encoding a variables categóricas, transformaciones logarítmicas a datos continuos y balanceo de clases con RandomOverSampler, descartaron SMOTE por bajo rendimiento. La selección de características la optimizaron mediante regularización LASSO, mientras que los hiperparámetros los ajustaron con un enfoque híbrido: búsqueda bayesiana (Optuna) y exhaustiva (GridSearchCV). En cuestión de validación, el estudio incluyó división 80%-20% de los datos, validación cruzada estratificada de 5 pliegues y una prueba temporal con datos de 2023. Además, las métricas evaluadas fueron precisión, recall, F1, AUC-ROC y matrices de confusión, respaldadas por análisis SHAP para interpretabilidad. Como resultado, el modelo Stacking destacó con un 99.1% de precisión, seguido por Random Forest (98.3%) y SVM-RBF (97.7%), mientras que la regresión logística tuvo (53.9%). También, se mostró la optimización con Optuna mejoró la precisión de Random Forest del 96.94% al 98.3%, y el balanceo aumentó el recall de la clase minoritaria del 92% al 98%.

El artículo [32] muestra un análisis del uso de IA en educación donde presenta el uso de algoritmos como Random Forest, SVM y redes neuronales, aplicados en sistemas de tutoría inteligente y dashboards predictivos de rendimiento estudiantil. El estudio manifiesta que emplearon metodologías rigurosas, incluyendo encuestas estandarizadas, análisis de datos históricos y talleres de co-diseño con estudiantes para evaluar percepciones éticas. En cuestión de métricas, exponen la precisión predictiva (80-90% en estudiantes tradicionales vs. 60-70% en no tradicionales), tasas de falsos positivos y niveles de transparencia algorítmica. Además, utilizaron herramientas como SHAP y Fairlearn ayudaron a identificar sesgos, revelando que los modelos eran 20-30% menos precisos para grupos subrepresentados. Como resultado, mostraron que el 60% de los estudiantes exigían mayor control sobre sus datos, mientras que solo el 15% de las herramientas ofrecían explicaciones interpretables. Además, aplicaron validación cruzada y auditorías éticas donde mencionan la necesidad de equilibrar precisión con equidad, especialmente en datos históricos cargados de desigualdades. Finalmente, los hallazgos del estudio subrayan la urgencia de desarrollar marcos éticos estandarizados para IA educativa.

El artículo [33] muestra el desarrollo de sistemas de alerta temprana para enfermedades infecciosas sensibles al clima mediante el uso integrado de herramientas de BI e IA,

donde emplearon plataformas como HealthMap, EpiWATCH y modelos estadísticos (ARIMA, SARIMA, GAM) y de machine learning (Random Forest, LSTM, XGBoost) con la finalidad de analizar datos climáticos, epidemiológicos y socioambientales. En cuestión de algoritmos, el estudio utiliza el aprendizaje supervisado y no supervisado para la predicción de brotes y la identificación de patrones ocultos, también utilizan sistemas de recomendación basados en umbrales de riesgo. Los KPIs evaluados incluyen precisión, sensibilidad, RMSE y AUC-ROC para validar la eficacia de los modelos que se utilizaron, respaldados por metodologías robustas como cross-validation temporal y feature engineering. Asimismo, para el procesamiento de datos desde la recolección y limpieza el estudio expone el uso de técnicas como KNN hasta el modelado espaciotemporal con enfoques Bayesianos (BYM, CAR) y el despliegue de dashboards interactivos para la visualización de riesgos. Como resultado, se muestra que la combinación de deep learning (redes neuronales, autoencoders) con modelos interpretables busca equilibrar precisión y transparencia, mientras que futuras mejoras contemplan el uso de redes neuronales de grafos y modelos de lenguaje avanzados para optimizar la vigilancia epidemiológica.

El artículo [34] muestra como empresas manufactureras escalan IA mediante pipelines de datos estructurados (sensores IoT, almacenamiento en la nube y catálogos etiquetados), combinados con algoritmos de predicción (aprendizaje supervisado para mantenimiento predictivo y por refuerzo para sistemas autónomos), integrando datos externos (clima, tráfico) para mejorar precisión. Asimismo, el estudio muestra sistemas de recomendación que se implementan en dashboards accesibles para usuarios no técnicos, apoyados por metodologías ágiles de co-creación con clientes y equipos multidisciplinarios. Como resultado, el estudio incluye KPIs de eficiencia (reducción del 15% en paradas no planificadas, 25% en optimización operacional) e ingresos por servicios basados en IA.

El artículo [35] muestra un sistema de inteligencia estratégica para la gestión educativa utilizando la suite Pentaho (v7.1), incluyendo Pentaho Data Integration para procesos ETL, Business Analytics para cubos OLAP, y Schema Workbench para modelado dimensional, junto con LinceBI 9.0 para dashboards interactivos. En cuestión de bases de datos, muestran BD relacionales (PostgreSQL/MySQL) para almacenar data marts y cubos OLAP con consultas MDX, mientras que R (con el paquete Caret) se usa para

análisis predictivo mediante modelos de regresión. La metodología que el estudio sigue es la de CRISP-DM para minería de datos y un enfoque ágil de ETL para transformar datos desde fuentes heterogéneas (Excel, sistemas legacy) hacia esquemas estandarizados. Además, las herramientas de visualización (Pentaho Analyzer, LinceBI) utilizadas para crear tablas pivotantes, informes estáticos y paneles configurables, validados mediante pruebas de consistencia y contraste con datos históricos. Como resultado, el estudio muestra que existe una mayor precisión (25%) en reportes frente a métodos tradicionales, con capacidad predictiva demostrada en la estimación de necesidades escolares (52 centros ajustados de 4,436). Finalmente, la plataforma desarrollada la evaluaron con 40 técnicos donde destacaron la interoperabilidad (integración de Pentaho, R y LinceBI) y escalabilidad en entornos educativos complejos.

El artículo [36] muestra el impacto de BI en PYMEs indonesias del sector creativo mediante PLS-SEM, se puede evidenciar que BI no afecta directamente su rendimiento (coef. 0.002, $*p* = 0.977$), pero sí indirectamente a través de mediadores clave. En cuestión de hallazgos, se manifiesta que el aprendizaje organizacional (coef. 0.731, $*p* = 0.000$) y la innovación (coef. 0.285, $*p* = 0.000$) actúan como mediadores en serie, mientras que la anticipación del valor del cliente (CVA) (coef. 0.615, $*p* = 0.000$) impulsa la innovación y el rendimiento. Contrariamente, el aprendizaje en red no fomenta significativamente la innovación (coef. -0.031, $*p* = 0.704$), aunque sí impacta el rendimiento (coef. 0.207, $*p* = 0.004$). Como resultados, el estudio utiliza la teoría RBV-KBV para destacar la necesidad de que las PYMEs integren BI con herramientas accesibles y fortalezcan el aprendizaje interno y la colaboración con proveedores. Además, no se presenta herramientas o manejo de datos concretos, este enfoque está destinado a mostrar la relevancia de tener entornos de BI e IA para el desarrollo de las PYMEs. Finalmente, el estudio muestra que las limitaciones incluyen el enfoque regional (Java Oriental) y la sugerencia de futuros estudios en sectores específicos o con variables demográficas.

El artículo [37] muestra un sistema de recomendación de salud basado en deep learning, combinando una Restricted Boltzmann Machine (RBM) con una Red Neuronal Convolutiva (CNN) para mejorar la precisión en la sugerencia de tratamientos y hospitales. Utiliza TensorFlow y Python como herramientas para su

implementación, evaluando su desempeño mediante métricas clave como RMSE (2.56), MAE (0.057), precisión (94.71%), recall (89%) y F-measure (92%), superando métodos tradicionales como Matrix Factorization (MF) y Singular Value Decomposition (SVD). La metodología incluye filtrado colaborativo híbrido, procesamiento de big data bajo las 3V's (volumen, variedad, velocidad) y técnicas de privacidad diferencial para proteger datos sensibles. Durante el análisis, el estudio del sistema demuestra avances en personalización y reducción de falsos positivos (4.45%), enfrenta desafíos como el problema de "cold start" y la falta de integración con estándares clínicos como HL7 o FHIR. Finalmente, el enfoque que utiliza da un aporte significativo a BI e IA en salud, al transformar datos médicos en recomendaciones accionables, aunque requeriría mayor énfasis en explicabilidad (XAI) y escalabilidad para su adopción clínica real.

1.2.1 Resumen de antecedentes

Los artículos revisados muestran una fuerte preferencia por metodologías mixtas, combinando enfoques cuantitativos y cualitativos como se puede evidenciar en Figura 1. Predomina el uso de Data Mining, Machine Learning, y BI como enfoques estratégicos para analizar datos organizacionales. Se destacan marcos como CRISP-DM, que guía el ciclo completo de minería de datos, y Design Science Research como marco para diseñar soluciones en contextos reales. Algunas investigaciones integran Scrum, PMI o Lean Six Sigma en procesos analíticos o de desarrollo. Se observa también el enfoque agile analytics, el uso del Ciclo de Vida del Desarrollo de Software (SDLC), y marcos conceptuales adaptados a sectores específicos como manufactura o logística. Varios artículos proponen modelos conceptuales extendidos para mejorar la agilidad o madurez digital de las empresas. Estas metodologías permiten una integración coherente entre los objetivos del negocio y las soluciones tecnológicas.

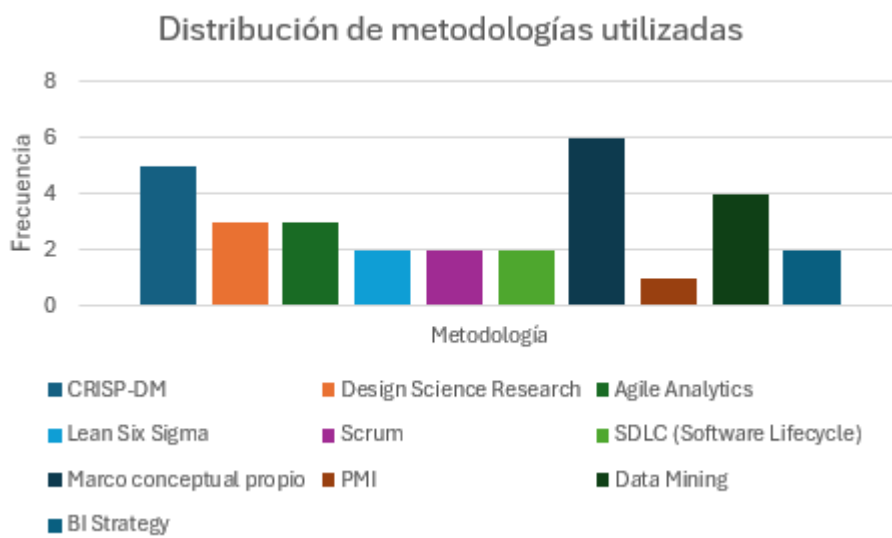


Figura 1. Frecuencia de metodologías utilizadas en estudios sobre BI e IA.

Respecto a los modelos propuestos, se identifican principalmente modelos predictivos, de recomendación, y de madurez digital. Algunos trabajos proponen modelos de decisión multicriterio (MCDM) y modelos conceptuales híbridos basados en inteligencia artificial. Se observa la implementación de modelos de optimización de portafolios de proyectos con criterios financieros, técnicos y de desempeño. Otros modelos están orientados a la detección de anomalías en procesos industriales o a la mejora de la toma de decisiones empresariales en tiempo real. También se emplean modelos de simulación y modelos de redes neuronales para análisis en contextos logísticos y manufactura inteligente. Varios modelos son extensiones de marcos existentes, integrando aprendizaje automático y temporalidad de los datos. En resumen, hay una evolución desde modelos clásicos hacia modelos híbridos y adaptativos vistos en Figura 2.

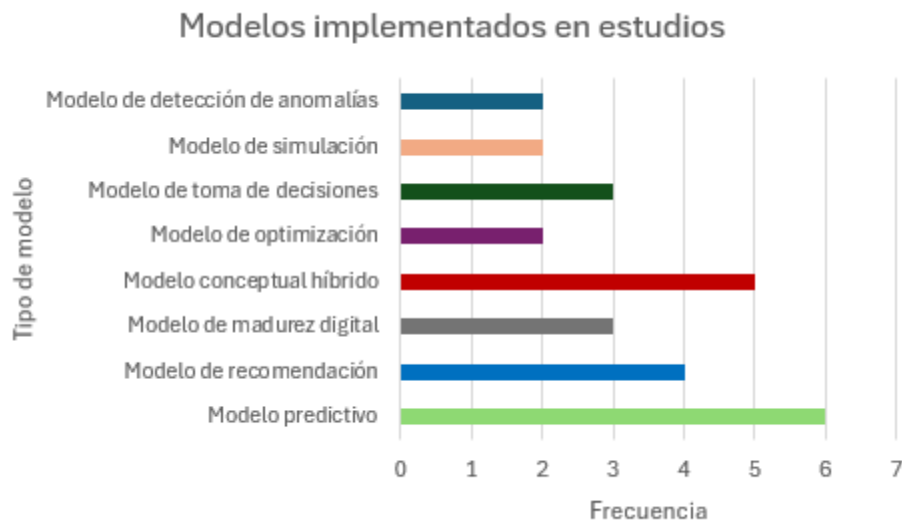


Figura 2. Frecuencia de modelos implementados en los estudios analizados.

En cuanto a los algoritmos más utilizados en los estudios revisados, estos pertenecen al campo del aprendizaje automático supervisado y no supervisado. Predomina el uso de Redes Neuronales Artificiales (ANN), árboles de decisión, Random Forest, y K-means como técnicas de agrupamiento. Se identifican también algoritmos como SVM (Máquinas de Vectores de Soporte), XGBoost y algoritmos genéticos para optimización. Algunos estudios utilizan Apriori y algoritmos de reglas de asociación para minería de patrones. Se emplean algoritmos híbridos, combinando clasificación con clustering o reducción de dimensionalidad (como PCA). Estos algoritmos permiten extraer patrones de comportamiento, hacer predicciones sobre variables clave y segmentar clientes o procesos. La diversidad de algoritmos empleados responde a la variedad de sectores y tipos de datos que se puede evidenciar en Figura 3.

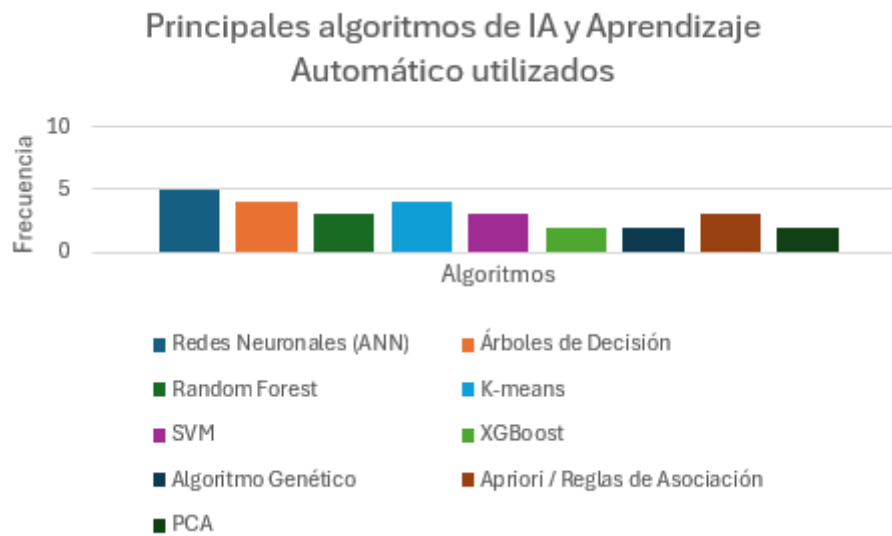


Figura 3. Principales algoritmos de IA y Aprendizaje Automático utilizados.

Respecto a software y herramientas utilizadas, los estudios evidencian un uso extendido de plataformas como Power BI, Python, R, y entornos como RapidMiner y WEKA visto en Figura 4. También se emplean Excel avanzado y complementos como Solver o Power Query. Herramientas específicas para minería de datos como Orange y KNIME se mencionan en aplicaciones prácticas. A nivel empresarial, algunos artículos integran soluciones ERP como SAP o Oracle BI. Además, en sectores industriales se utilizan SCADA, sistemas embebidos o entornos simulados como AnyLogic y Simulink. Para visualización y dashboards, Power BI y Tableau lideran por su integración con sistemas organizacionales. Algunas propuestas desarrollan software propio o módulos web basados en frameworks como Django, Flask o React. El uso del software es determinante para la implementación real de los modelos propuestos.

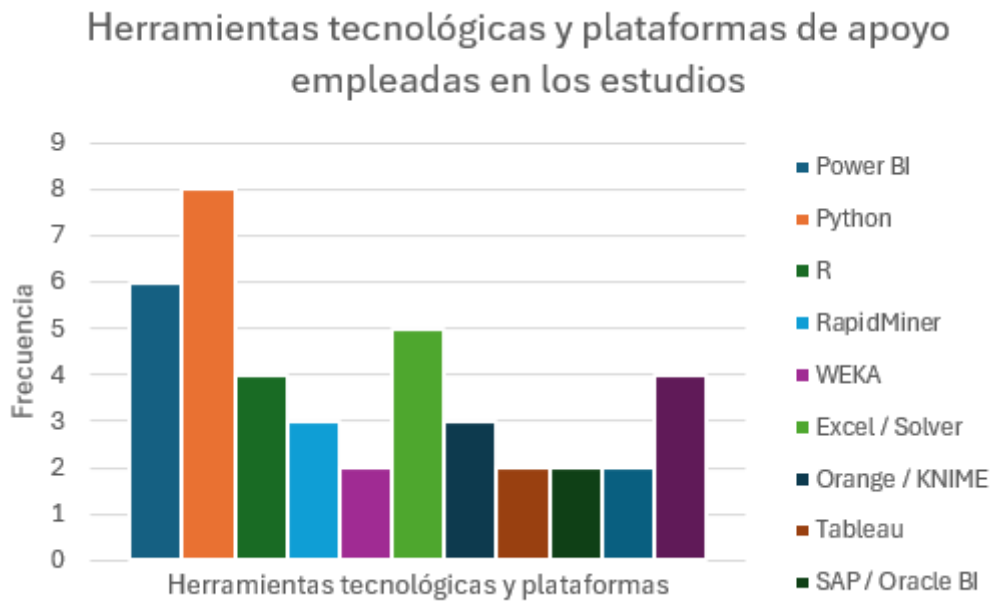


Figura 4. Frecuencia de herramientas tecnológicas y plataformas de apoyo empleadas en los estudios.

En cuanto a bases de datos y fuentes de información, las investigaciones utilizan fuentes de datos tanto internas (empresariales reales) como externas (datasets públicos y simulados). Se destacan conjuntos de datos estructurados provenientes de sistemas ERP, CRM o sensores industriales (IoT). Algunos estudios usan bases como UCI Machine Learning Repository, Kaggle, o datos de series temporales específicas del sector manufactura o salud. Otros desarrollan bases de datos propias mediante encuestas, recolección de logs o datos históricos de desempeño empresarial. Se nota un interés por datos de alta frecuencia y granularidad para análisis predictivo. El uso de series temporales y datos en tiempo real es clave en aplicaciones de mantenimiento predictivo y monitoreo. Algunas investigaciones también incluyen análisis de texto no estructurado y minería de opiniones que se puede evidenciar en Figura 5.



Figura 5. Tipos y porcentajes de uso de fuentes de datos utilizados.

1.3 Fundamentación teórica

El desarrollo de esta investigación realizó una revisión de estudios previos para comprender el estado actual de las investigaciones en el área. Este análisis permite identificar aspectos cualitativos aplicados en las investigaciones previas. A partir de esta base, se identifican los recursos y se profundiza en los conceptos clave necesarios para llevar a cabo el estudio. Con estos elementos claros, se establece un punto de partida sólido para el desarrollo de la investigación. Finalmente, esta fase inicial garantizará que la investigación esté alineada con los avances y necesidades actuales del campo.

1.3.1 Business Intelligence

BI hace referencia al conjunto de metodologías, procesos, arquitecturas y tecnologías que permiten transformar grandes volúmenes de datos en información útil y conocimiento accionable, con el objetivo de respaldar la toma de decisiones estratégicas [28]. A través de la integración, análisis y visualización de datos, BI facilita la identificación de oportunidades, mejora la eficiencia operativa y permite un monitoreo continuo del desempeño organizacional [29]. Su implementación en

instituciones de educación superior permite optimizar la planificación académica, mejorar la asignación de recursos y fortalecer la evaluación de resultados de investigación.

1.3.2 IA en entornos BI

La IA complementa los procesos de BI mediante algoritmos capaces de generar predicciones y recomendaciones a partir de datos históricos y en tiempo real. Técnicas como redes neuronales, árboles de decisión y aprendizaje automático o machine learning se utilizan para anticipar resultados, detectar patrones y proponer acciones automatizadas [30]. La fusión de BI e IA ha dado lugar a un enfoque más dinámico conocido como Augmented Analytics, que automatiza procesos analíticos complejos y mejora la experiencia de los usuarios en la toma de decisiones basada en datos [31].

1.3.3 Ciencia de datos y toma de decisiones

La ciencia de datos integra estadística, informática y análisis de datos con el fin de extraer valor de conjuntos de datos complejos. En entornos organizacionales, su aplicación permite descubrir patrones ocultos, predecir comportamientos y generar conocimiento estratégico [32]. En el contexto universitario, la ciencia de datos contribuye a tomar decisiones informadas en la gestión de proyectos, evaluación de indicadores de producción científica y seguimiento de resultados investigativos.

1.3.4 Reportería dinámica y dashboards

La reportería dinámica se refiere a la generación automática de informes interactivos que permiten una visualización clara y flexible de los datos en tiempo real. Los tableros de control o **dashboards** son herramientas gráficas que sintetizan indicadores clave de desempeño (KPIs), facilitando una comprensión rápida y accionable de la información [33]. Esta visualización mejora la eficiencia en la toma de decisiones y permite monitorear el progreso hacia los objetivos estratégicos institucionales.

1.3.5 Sistemas de recomendación y predicción

Los sistemas de recomendación utilizan técnicas de IA para sugerir acciones, contenidos o decisiones basadas en patrones previos de comportamiento o preferencia.

En el contexto de BI, los sistemas de predicción permiten anticipar resultados como niveles de productividad o riesgos de baja eficiencia mediante modelos estadísticos y de aprendizaje automático [34]. Estas herramientas permiten personalizar la experiencia del usuario institucional, mejorar la planificación y orientar decisiones con base en evidencia empírica.

1.4 Objetivos

1.4.1 Objetivo general

Implementar un sistema de reportería y visualización de datos con funciones de recomendación y predicción, para optimizar la toma de decisiones estratégicas en la DIDE de la UTA, mediante el uso de técnicas de BI e IA, durante el periodo de Agosto 2025 - Enero 2026.

1.4.2 Objetivos específicos

- Analizar el estado actual de la gestión informativa en la DIDE, con el propósito de identificar las principales limitaciones en la generación, análisis y presentación de los datos.
- Desarrollar un sistema de reportería y visualización interactiva con funciones de recomendación y predicción, basado en técnicas de BI e IA, orientado a fortalecer los procesos de planificación y evaluación estratégica en la DIDE.
- Evaluar el desempeño del sistema desarrollado mediante métricas como reducción de tiempos de reportería, nivel de adopción institucional y utilidad percibida por los usuarios, en un entorno operativo que represente las condiciones reales de gestión en la DIDE.

CAPÍTULO II. METODOLOGÍA.

2.1 Materiales

Para la presente investigación se emplearon instrumentos de recolección de información dirigidos a los cinco funcionarios de la Dirección de Investigación y Desarrollo de la Universidad Técnica de Ambato, quienes conforman la población de estudio y participan directamente en los procesos relacionados con la gestión de información científica institucional. Estos instrumentos permitieron recopilar información sobre la situación actual del proceso de registro, reportería, organización de datos, uso de herramientas digitales y necesidades asociadas a la implementación de una solución basada en técnicas de Business Intelligence e Inteligencia Artificial.

En primera instancia, se diseñó y aplicó una entrevista estructurada con escala tipo Likert, orientada a obtener datos cuantificables sobre la percepción de los participantes con respecto a los procesos actuales de gestión informativa. Este instrumento permitió valorar aspectos relacionados con la agilidad del registro de artículos, el tiempo de generación de reportes, la organización de los datos, la disposición al uso de herramientas digitales, la utilidad esperada del sistema BI-IA, la calidad de los informes, las funciones de predicción y recomendación, la facilidad de uso y la participación en la validación de la propuesta. La entrevista estructurada estuvo conformada por ítems cerrados valorados en una escala de 1 a 5, donde 1 representa el menor grado de acuerdo y 5 el mayor. La versión final del instrumento estructurado se presenta en el Anexo 1.

Adicionalmente, se consideró una entrevista semiestructurada compuesta por preguntas abiertas, orientada a profundizar en aspectos relacionados con la gestión de información científica, reportería institucional, validación de registros, trazabilidad, integración de información y necesidades tecnológicas. Este instrumento permitió recopilar información cualitativa complementaria sobre el contexto institucional y los procesos analizados. El guion de la entrevista semiestructurada se presenta en el Anexo 2.

2.2 Métodos

2.2.1 Modalidad de la investigación

La presente investigación se desarrolló bajo una modalidad aplicada, debido a que estuvo orientada a la atención de una necesidad institucional concreta en la Dirección de Investigación y Desarrollo de la Universidad Técnica de Ambato. Esta modalidad fue pertinente porque el estudio se dirigió al planteamiento de una solución tecnológica frente a una problemática real del entorno organizacional. Su carácter aplicado se justifica en la utilización de conocimientos de ingeniería de software, gestión de datos y análisis de información con una finalidad práctica dentro del contexto universitario. Asimismo, la investigación se complementó con un componente de campo, ya que la información necesaria para comprender la situación objeto de estudio fue obtenida directamente en la unidad institucional correspondiente. Esto permitió mantener coherencia entre la realidad observada y el propósito de la investigación dentro del escenario donde se presenta la problemática. De igual manera, el estudio se mantuvo en un contexto no experimental, puesto que no se manipularon deliberadamente las condiciones naturales en las que se desarrollan los procesos institucionales analizados. En conjunto, la modalidad adoptada resultó adecuada para sustentar metodológicamente una investigación orientada a la aplicación práctica y contextualizada del conocimiento.

2.2.2 Población y muestra

La población objeto de estudio estuvo conformada por los cinco funcionarios de la DIDE - UTA, quienes participan directamente en los procesos de gestión y seguimiento de la información institucional asociada al área de investigación. Esta población fue considerada pertinente debido a su relación directa con el contexto administrativo y operativo en el que se desarrolla la presente investigación. Dado que el número de integrantes fue reducido y plenamente identificable, no resultó necesario aplicar una fórmula de muestreo estadístico para seleccionar participantes. En consecuencia, se trabajó con la totalidad de la población disponible, por lo que la muestra coincidió con los cinco funcionarios de la DIDE-UTA. Esta decisión

metodológica permitió disponer de información completa y contextualizada de todos los actores involucrados en la realidad analizada. No obstante, debido al tamaño reducido de la población, los resultados derivados del estudio deben interpretarse como evidencia descriptiva propia del contexto institucional examinado. Por ello, su alcance se circunscribe a la DIDE-UTA y no pretende generalizarse de manera directa a otras dependencias o entornos distintos.

2.2.3 Recolección de información

La recolección de información se desarrolló de manera planificada, considerando como instrumento principal la entrevista estructurada con escala tipo Likert. Esta fue aplicada a los cinco funcionarios que conformaron la población de estudio, con el propósito de obtener datos cuantificables sobre la percepción de los participantes frente al estado actual de la gestión informativa y la pertinencia de implementar una solución tecnológica basada en BI e IA.

La entrevista estructurada permitió valorar la percepción de los participantes sobre la eficiencia del registro de artículos, el tiempo requerido para generar reportes, la organización y actualización de los datos, la preparación para utilizar herramientas digitales, la utilidad esperada del sistema propuesto, la mejora de la calidad de los informes, el aporte de las funciones de predicción y recomendación, la confianza en la utilidad práctica del sistema, la importancia de una interfaz simple y la disposición para participar en la validación de la propuesta. Estos aspectos fueron seleccionados debido a su relación directa con el problema de investigación y con los objetivos planteados.

La aplicación del instrumento se realizó de forma presencial, utilizando ítems cerrados con escala de valoración de 1 a 5. Esta estructura permitió recopilar información homogénea entre los participantes y facilitar el procesamiento descriptivo de los resultados. La información obtenida permitió establecer una línea base sobre las condiciones iniciales del proceso de gestión informativa, especialmente en relación con tiempos de reportería, eficiencia operativa, disposición tecnológica y aceptación inicial de una solución BI-IA.

De forma complementaria, se aplicó una entrevista semiestructurada con preguntas abiertas, con el propósito de ampliar la comprensión del contexto institucional y

recoger información cualitativa relacionada con los procesos de gestión científica, reportería, validación, trazabilidad e integración de información. La aplicación de ambos instrumentos permitió disponer de una visión cuantitativa y cualitativa del problema, manteniendo coherencia con la modalidad aplicada y de campo de la investigación.

Análisis de Resultados de entrevista estructurada

Ítem 1. Los procesos actuales de registro de artículos son ágiles y eficientes. El promedio obtenido fue de 2,60, con una desviación estándar de 0,55 y 0 % de respuestas favorables en los niveles 4 y 5. Este resultado evidencia una percepción baja respecto a la agilidad del proceso actual de registro. Por tanto, se determina que existe una limitación operativa relacionada con el ingreso y gestión de información científica, lo que justifica la necesidad de automatizar y centralizar el proceso.

Ítem 2. El tiempo de generación de reportes es adecuado para las necesidades institucionales.

El promedio fue de 2,80, con una desviación estándar de 0,45 y 0 % de aceptación en los niveles altos. Este resultado muestra que los participantes no consideran suficiente el tiempo actual de generación de reportes. En consecuencia, se evidencia la necesidad de implementar mecanismos de reportería automatizada, dashboards y exportación de información que reduzcan la dependencia de procesos manuales.

Ítem 3. Los datos registrados se mantienen organizados y actualizados. El promedio alcanzado fue de 3,60, con una desviación estándar de 0,55 y 60 % de respuestas favorables. Este resultado refleja una percepción moderada: existen ciertos niveles de organización, pero no suficientes para garantizar actualización, centralización y disponibilidad analítica de la información. Por ello, se justifica la implementación de una base institucional integrada y mecanismos de validación de datos.

Ítem 4. Me siento preparado/a para utilizar herramientas digitales más avanzadas. El promedio fue de 4,40, con una desviación estándar de 0,55 y 100 % de respuestas favorables. Este resultado evidencia una disposición positiva hacia el uso de herramientas tecnológicas más avanzadas. Por tanto, se determina que existe un escenario favorable para la adopción de una solución basada en BI e IA.

Ítem 5. La implementación de un sistema BI-IA sería beneficiosa para la gestión institucional.

El promedio obtenido fue de 4,80, con una desviación estándar de 0,45 y 100 % de aceptación. Este resultado muestra una valoración muy favorable hacia la propuesta tecnológica. En consecuencia, se evidencia que los participantes perciben el sistema como una herramienta útil para mejorar la gestión informativa y apoyar procesos de decisión.

Ítem 6. Un sistema de reportería inteligente mejoraría la calidad de los informes institucionales.

El promedio fue de 4,60, con una desviación estándar de 0,55 y 100 % de respuestas favorables. Este resultado permite determinar que existe una percepción positiva sobre el aporte de la reportería inteligente en la calidad de los informes. Por ello, se justifica la incorporación de dashboards, KPIs, reportes PDF/Excel y visualizaciones analíticas.

Ítem 7. Las funciones de predicción y recomendación ayudarían a tomar mejores decisiones.

El promedio alcanzado fue de 4,60, con una desviación estándar de 0,55 y 100 % de aceptación. Este resultado respalda la incorporación de funciones de Inteligencia Artificial en el sistema. Se concluye que los participantes perciben valor en contar con predicciones y recomendaciones como apoyo para anticipar tendencias y orientar decisiones institucionales.

Ítem 8. Confío en la utilidad práctica del sistema propuesto. El promedio fue de 4,60, con una desviación estándar de 0,55 y 100 % de respuestas favorables. Este resultado evidencia confianza en la aplicación práctica del sistema. Por tanto, la propuesta no solo se percibe como técnicamente pertinente, sino también como útil para los procesos institucionales relacionados con gestión y reportería.

Ítem 9. Considero que la interfaz debe ser simple y de fácil uso. El promedio obtenido fue de 4,80, con una desviación estándar de 0,45 y 100 % de aceptación. Este resultado evidencia que la facilidad de uso es un criterio clave para la adopción del sistema. Por ello, el diseño visual debe priorizar navegación clara, organización modular, dashboards comprensibles e interfaces intuitivas.

Ítem 10. Me gustaría participar en la validación o pruebas del nuevo sistema. El promedio fue de 4,60, con una desviación estándar de 0,55 y 100 % de respuestas favorables. Este resultado muestra disposición de los participantes para colaborar en la validación del sistema. En consecuencia, se evidencia una actitud favorable hacia la evaluación de la propuesta y condiciones favorables para adopción progresiva.

2.2.4 Procesamiento y análisis de datos

El procesamiento de los datos se realizó a partir de la entrevista estructurada con escala tipo Likert aplicada a los cinco funcionarios de la DIDE-UTA. Para cada ítem se calcularon medidas descriptivas como promedio, desviación estándar, frecuencia de respuestas y porcentaje de aceptación. La escala utilizada fue de 1 a 5, donde 1 representa el menor nivel de acuerdo y 5 el mayor. Para interpretar los resultados, se consideraron como respuestas favorables aquellas ubicadas en los valores 4 y 5, debido a que reflejan acuerdo o total acuerdo con la afirmación evaluada.

El análisis permitió identificar dos comportamientos principales. Por una parte, los ítems relacionados con el proceso actual de gestión informativa presentaron valores bajos o moderados, especialmente en la agilidad del registro de artículos y en el tiempo de generación de reportes. Por otra parte, los ítems relacionados con la utilidad esperada del sistema BI-IA, la mejora de informes, las funciones de predicción y recomendación, la facilidad de uso y la disposición para participar en la validación obtuvieron valores altos. Esto evidencia que, aunque existen limitaciones en los

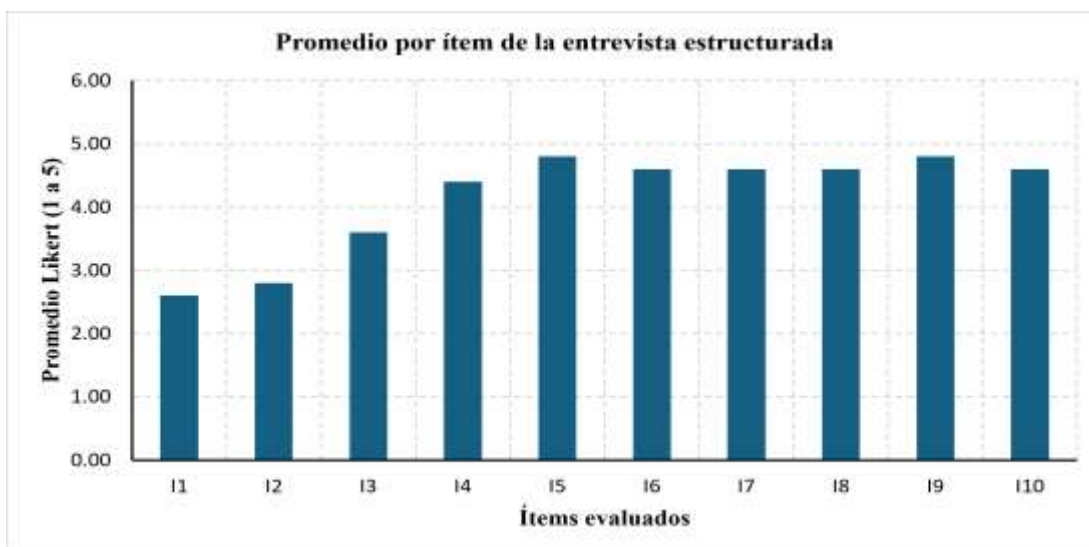


Figura 6. Promedio por ítem de la entrevista estructurada.

procesos actuales, también existe una percepción favorable hacia la implementación de una solución tecnológica basada en Business Intelligence e Inteligencia Artificial.

La Figura 6 muestra el promedio obtenido en cada ítem evaluado. Los valores más bajos corresponden al Ítem 1, relacionado con la agilidad del registro actual, con un promedio de 2,60, y al Ítem 2, relacionado con el tiempo de generación de reportes, con un promedio de 2,80. Estos resultados evidencian que los participantes perciben limitaciones en los procesos actuales de registro y reportería. En contraste, los ítems relacionados con la utilidad esperada del sistema, la facilidad de uso y la mejora de informes presentan promedios entre 4,40 y 4,80, lo que refleja una valoración positiva hacia la propuesta tecnológica.



Figura 7. Porcentaje de aceptación por ítem de la entrevista estructurada.

La Figura 7 presenta el porcentaje de aceptación por ítem, considerando las respuestas ubicadas en los valores 4 y 5 de la escala Likert. Los ítems I1 e I2 obtuvieron 0 % de aceptación, lo que confirma que los procesos actuales de registro y reportería no son considerados adecuados por los participantes. En cambio, desde el ítem I4 hasta el I10 se obtuvo una aceptación del 100 %, evidenciando una percepción favorable hacia el uso de herramientas digitales, la implementación del sistema BI-IA, la mejora de reportes, las funciones de predicción y recomendación, y la participación en la validación de la propuesta.

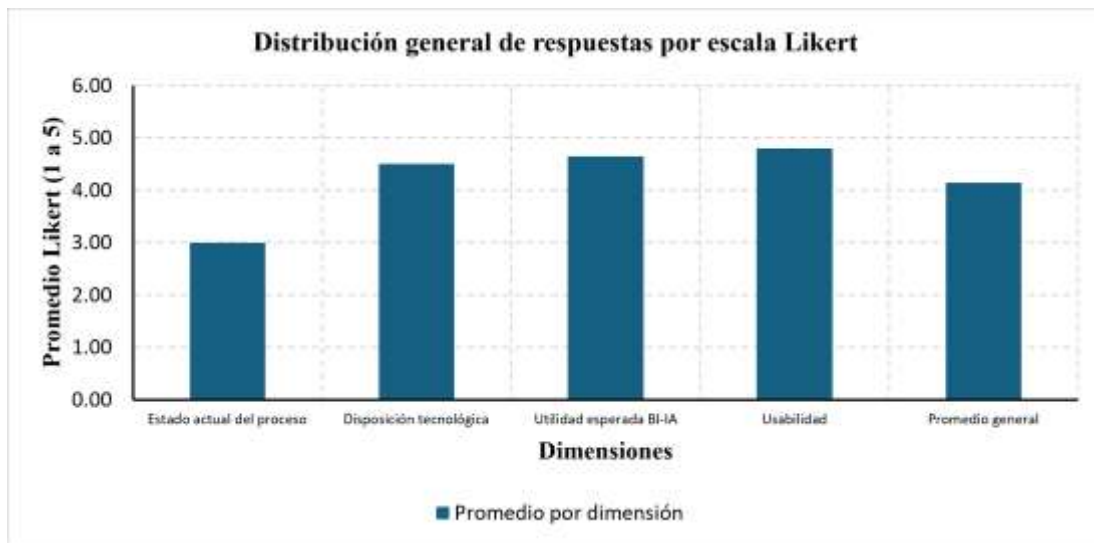


Figura 9. Promedio por dimensión evaluada en la entrevista estructurada.

La Figura 8 agrupa los resultados por dimensiones de análisis. La dimensión relacionada con el estado del proceso actual obtuvo el promedio más bajo, con 3,00, lo que confirma la existencia de oportunidades de mejora en la gestión informativa. En cambio, las dimensiones relacionadas con utilidad esperada del sistema, funciones de IA, usabilidad y participación en la validación presentaron promedios superiores a 4,40. Estos resultados respaldan la pertinencia de implementar una plataforma que centralice la información, automatice la reportería y permita incorporar funciones de recomendación y predicción.

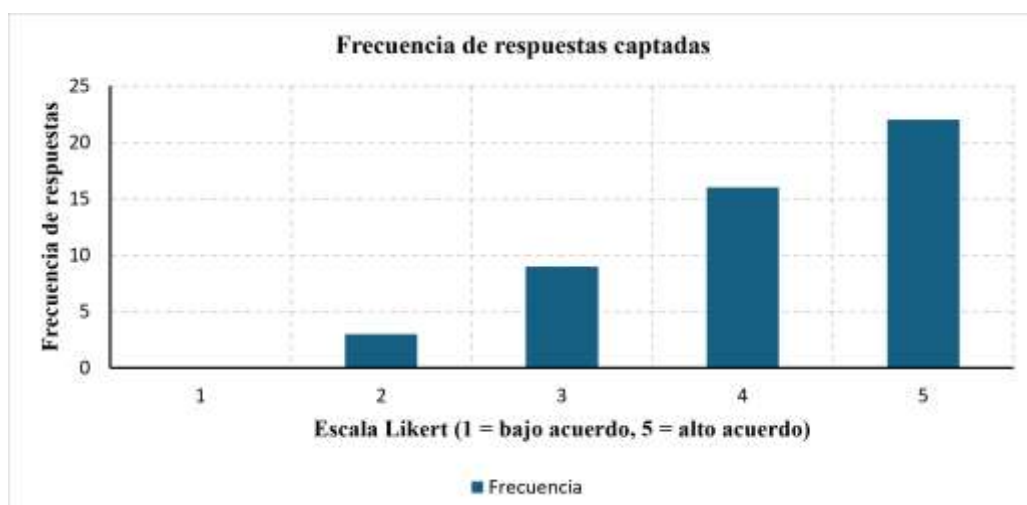


Figura 8. Distribución general de respuestas por escala Likert.

La Figura 9 evidencia que la mayoría de las respuestas se concentran en los valores altos de la escala, principalmente 4 y 5. Esta tendencia refleja una valoración favorable hacia la implementación del sistema propuesto. No obstante, la presencia de respuestas en niveles bajos para los ítems relacionados con el proceso actual confirma que existen limitaciones en la agilidad del registro y en la generación de reportes. De esta manera, la distribución general de respuestas permite sustentar tanto la problemática identificada como la necesidad de desarrollar una solución tecnológica orientada a reportería, visualización, recomendación y predicción.

Adicionalmente, se calculó el Alfa de Cronbach como referencia de consistencia interna del instrumento estructurado. Este indicador permitió valorar la coherencia de los ítems aplicados dentro de la entrevista estructurada. Debido a que la población estuvo conformada por cinco participantes, el resultado debe interpretarse como un indicador descriptivo de apoyo y no como una validación estadística generalizable.

En síntesis, el procesamiento de los datos permitió establecer una línea base sobre la situación inicial de la gestión informativa en la DIDE-UTA. Los resultados evidencian limitaciones en los procesos actuales de registro y reportería, pero también una alta aceptación hacia la implementación de un sistema BI-IA. Por ello, la información obtenida mediante la entrevista estructurada respalda la necesidad de desarrollar una plataforma que automatice procesos, mejore la visualización de información y apoye la toma de decisiones mediante funciones de recomendación y predicción.

CAPÍTULO III. RESULTADOS Y DISCUSIÓN

3.1 Análisis del contexto y requerimientos

La presente sección describe el proceso de análisis realizado para identificar las necesidades institucionales, operativas y funcionales relacionadas con la gestión de información científica dentro de la DIDE - UTA. Para ello, se empleó información obtenida mediante dos entrevistas orientadas tanto al análisis del contexto institucional como al levantamiento de requerimientos del sistema propuesto. La segunda entrevista correspondió a una entrevista semiestructurada, la cual permitió establecer un primer acercamiento con los entrevistados y profundizar en las necesidades reales de los procesos involucrados dentro de la dirección. Dichas entrevistas fueron desarrolladas en horarios previamente establecidos mediante un cronograma coordinado con el personal responsable de la institución, permitiendo una adecuada organización del proceso de recopilación de información. A partir de este análisis se identificaron limitaciones asociadas a la gestión manual de información, la ausencia de mecanismos centralizados de reportería, la limitada trazabilidad de los procesos y la necesidad de incorporar herramientas de BI e IA para fortalecer la toma de decisiones estratégicas. Asimismo, la información recopilada fue procesada y organizada mediante técnicas de análisis cualitativo apoyadas en el uso de una herramienta informática, permitiendo estructurar categorías y necesidades institucionales que posteriormente sirvieron como base para el diseño y desarrollo de la propuesta tecnológica.

3.1.1 Procesamiento de la información

El procesamiento de la información se desarrolló a partir de los datos obtenidos durante las entrevistas aplicadas al personal involucrado en los procesos de gestión de información científica dentro de la DIDE-UTA, constituyéndose estas como una fuente principal para el levantamiento inicial de requerimientos y el análisis del contexto institucional. La información recopilada permitió identificar aspectos relacionados con el registro, validación, almacenamiento y generación de reportes asociados a la producción científica institucional. Posteriormente, los datos obtenidos fueron organizados y clasificados de acuerdo con criterios funcionales, operativos y

tecnológicos, permitiendo establecer relaciones entre las necesidades institucionales y las funcionalidades requeridas para el sistema propuesto.

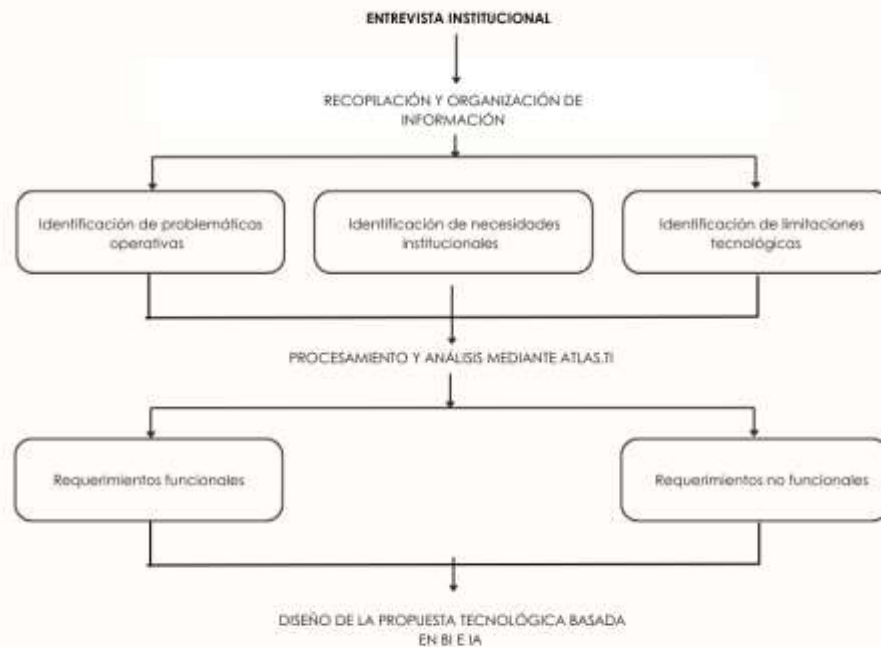


Figura 10. Flujo del procesamiento y análisis de la información para la definición de requerimientos del sistema.

Como se observa en la Figura 10, el análisis realizado permitió identificar problemáticas operativas, necesidades institucionales y limitaciones tecnológicas que posteriormente fueron procesadas mediante una herramienta informática para la estructuración de los requerimientos del sistema. Asimismo, durante las etapas posteriores de análisis y retroalimentación institucional, se identificaron nuevas necesidades relacionadas con la integración de información científica externa, la automatización de procesos de validación, el fortalecimiento de la reportería analítica y la incorporación de funcionalidades orientadas a Business Intelligence e Inteligencia Artificial para apoyar la toma de decisiones estratégicas. Finalmente, la información obtenida permitió establecer la base conceptual y funcional utilizada posteriormente durante las etapas de diseño y desarrollo de la propuesta tecnológica.

A partir del análisis realizado fue posible identificar limitaciones relacionadas con la gestión de información científica, los procesos de reportería institucional y el manejo de datos dentro de la dirección. Asimismo, el estudio permitió clasificar las

necesidades institucionales desde una perspectiva funcional, operativa y tecnológica, facilitando la estructuración de la propuesta tecnológica de acuerdo con los requerimientos identificados. El uso de una herramienta informática de análisis cualitativo permitió establecer relaciones entre las problemáticas detectadas y las funcionalidades requeridas por la institución, contribuyendo a una mejor comprensión de los procesos involucrados. De igual manera, se evidenció la necesidad de incorporar mecanismos de trazabilidad, automatización, integración de información externa y reportería para optimizar la gestión institucional y fortalecer la toma de decisiones. Finalmente, los resultados obtenidos permitieron consolidar la base conceptual y funcional utilizada posteriormente durante las etapas de diseño, desarrollo e implementación del sistema propuesto.

Tabla 6. Problemáticas identificadas durante el proceso de levantamiento de información.

Código	Limitación identificada	Impacto institucional	Proceso
P01	Gestión manual de información científica	Retrasos en el procesamiento y validación de registros	Registro, validación y generación de reportes
P02	Información almacenada en múltiples fuentes y formatos	Dificultad para consolidar información institucional	Consolidación y consulta de información
P03	Ausencia de mecanismos centralizados de reportería	Limitaciones en la generación de indicadores estratégicos	Reportería institucional
P04	Escasa trazabilidad en los procesos de validación	Dificultad para realizar seguimiento de registros	Validación y seguimiento
P05	Procesos manuales de revisión y aprobación	Incremento de errores e inconsistencias en la información	Revisión y aprobación de registros
P06	Limitada integración con fuentes externas de información científica	Restricciones para obtener información académica actualizada	Integración y actualización de datos
P07	Ausencia de dashboards y visualización analítica	Baja capacidad de análisis para la toma de decisiones	Análisis y visualización de información

P08	Generación manual de reportes institucionales	Incremento en tiempos de elaboración de informes	Elaboración de reportes institucionales
-----	---	--	---

La Tabla 6 presenta las principales problemáticas identificadas durante el proceso de levantamiento de información realizado en la DIDE-UTA. Entre las limitaciones más relevantes se evidenciaron la gestión manual de registros científicos, la dispersión de información institucional y la ausencia de mecanismos centralizados de reportería y trazabilidad. Asimismo, se identificaron dificultades relacionadas con los procesos manuales de validación y aprobación, ocasionando inconsistencias, duplicidad de información y retrasos en el procesamiento de registros científicos. De igual manera, la limitada integración con fuentes externas de información académica restringía la obtención automatizada de información actualizada para la consolidación institucional. Además, la ausencia de dashboards y herramientas de visualización analítica dificultaba el análisis estratégico y la generación oportuna de indicadores institucionales.

Durante el levantamiento de información también se identificó el interés institucional por incorporar funcionalidades orientadas a Inteligencia Artificial como mecanismo de apoyo para la toma de decisiones, análisis predictivo y fortalecimiento de futuras mejoras tecnológicas dentro de la dirección. Estas funcionalidades fueron consideradas como un componente estratégico para potenciar el análisis institucional y modernizar los procesos relacionados con la gestión de información científica. Finalmente, las problemáticas identificadas permitieron justificar la necesidad de implementar una solución tecnológica orientada a la automatización de procesos, integración de información científica, reportería analítica y fortalecimiento de la toma de decisiones institucionales.

Tabla 7. Necesidades institucionales identificadas para el desarrollo del sistema.

Código	Necesidad institucional	Prioridad
N01	Centralizar la gestión de información científica institucional	Alta
N02	Automatizar los procesos de validación y aprobación de registros	Alta
N03	Incorporar mecanismos de trazabilidad y seguimiento de información	Alta
N04	Implementar reportería analítica mediante dashboards y KPIs	Alta

N05	Integrar información científica proveniente de fuentes externas como Scopus	Alta
N06	Reducir tiempos en la generación de reportes institucionales	Media-Alta
N07	Gestionar procesos de carga masiva de información	Media-Alta
N08	Permitir la configuración dinámica de catálogos y campos del sistema	Media
N09	Fortalecer el análisis institucional mediante herramientas BI e IA	Alta
N10	Implementar mecanismos de seguridad mediante roles y permisos	Alta

La Tabla 7 presenta las principales necesidades institucionales identificadas durante el proceso de levantamiento y análisis de información realizado en la DIDE-UTA. Entre las necesidades con mayor prioridad se identificaron la centralización de la información científica institucional, la automatización de procesos de validación y aprobación, y la incorporación de mecanismos de reportería analítica orientados al fortalecimiento de la toma de decisiones estratégicas. Asimismo, se evidenció la necesidad de integrar información proveniente de fuentes externas como Scopus, permitiendo consolidar registros científicos institucionales dentro de una plataforma centralizada y actualizada. De igual manera, se identificó la importancia de incorporar mecanismos de trazabilidad, seguridad y control de acceso para optimizar la gestión de procesos relacionados con la producción científica institucional. Finalmente, las necesidades identificadas permitieron establecer la base funcional y tecnológica utilizada posteriormente para la definición de requerimientos funcionales y no funcionales del sistema propuesto.

3.1.2 Análisis de requerimientos

El análisis de requerimientos se desarrolló mediante el uso de una herramienta informática, la cual permitió procesar y estructurar la información obtenida durante las entrevistas realizadas al personal involucrado en los procesos de gestión de información científica dentro de la DIDE-UTA. El uso de esta herramienta facilitó la organización de la información cualitativa mediante procesos de codificación, categorización y análisis de relaciones entre las necesidades identificadas durante el levantamiento de información. A partir de este procedimiento fue posible reconocer patrones asociados a la gestión de artículos científicos, los procesos de validación, la

reportería institucional, la trazabilidad de registros y la integración con fuentes externas de información académica.

Asimismo, el análisis permitió establecer relaciones entre las problemáticas detectadas y las funcionalidades requeridas para la propuesta tecnológica, contribuyendo a la definición inicial de requerimientos alineados con los procesos operativos y administrativos de la institución. Del mismo modo, la categorización de la información permitió identificar necesidades relacionadas con la automatización de procesos, el fortalecimiento de mecanismos de control, la centralización de registros científicos y la incorporación de herramientas de BI e IA orientadas al análisis institucional y apoyo a la toma de decisiones estratégicas. Finalmente, los resultados obtenidos sirvieron como base para la estructuración de categorías funcionales y tecnológicas consideradas posteriormente durante las etapas de diseño y desarrollo del sistema propuesto.

3.1.3 Requerimientos funcionales y no funcionales

Los requerimientos funcionales y no funcionales del sistema fueron definidos a partir del análisis de la información obtenida durante las entrevistas, las problemáticas institucionales identificadas y las necesidades operativas de la DIDE-UTA. Este proceso permitió establecer una relación directa entre los procedimientos actuales de gestión de producción científica y las funcionalidades necesarias para automatizar, controlar y analizar la información institucional. A partir de este análisis se consideraron módulos relacionados con seguridad, gestión de artículos científicos, carga masiva de registros, workflow institucional, integración con fuentes externas, reportería analítica, Business Intelligence, procesos ETL y funcionalidades orientadas al análisis predictivo y recomendación de información científica mediante Inteligencia Artificial.

Los requerimientos funcionales se orientaron a describir las acciones y operaciones que el sistema debe ejecutar para responder a las necesidades operativas y administrativas de la institución. Por su parte, los requerimientos no funcionales permitieron definir características técnicas relacionadas con seguridad, rendimiento, escalabilidad, mantenibilidad, usabilidad e interoperabilidad. De esta manera, la definición de requerimientos constituyó la base para orientar el diseño arquitectónico, el desarrollo modular y las pruebas posteriores del sistema propuesto.

Con base en las necesidades institucionales y los procesos identificados durante el levantamiento de información, se definieron los principales módulos funcionales que conforman la propuesta tecnológica. Como se observa en la Figura 11, el sistema fue estructurado mediante módulos orientados a la gestión de información científica, validación institucional, reportería analítica, integración de información externa y funcionalidades de apoyo predictivo basadas en Inteligencia Artificial.

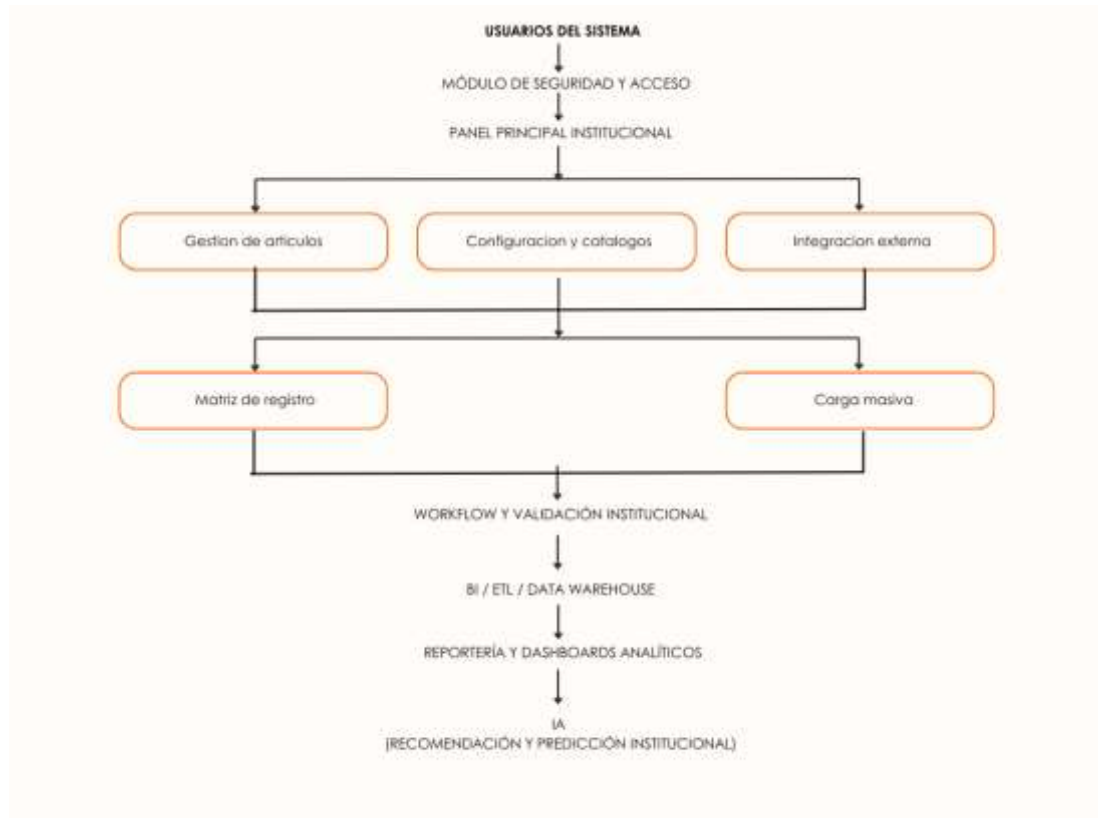


Figura 11. Mapa funcional general del sistema propuesto.

Los requerimientos funcionales fueron definidos a partir de las necesidades identificadas durante el levantamiento de información y del análisis de los procesos institucionales relacionados con la gestión de producción científica. Estos requerimientos describen las funcionalidades y operaciones que el sistema debe ejecutar para permitir el registro, validación, procesamiento, análisis y generación de reportes relacionados con la información científica institucional. Asimismo, los requerimientos funcionales contemplan procesos asociados a seguridad, workflow institucional, integración externa, carga masiva de información, Business Intelligence y funcionalidades orientadas al análisis predictivo y recomendación como apoyo a la

toma de decisiones estratégicas. La Tabla 8 presenta los principales requerimientos funcionales definidos para la propuesta tecnológica.

Tabla 8. Requerimientos funcionales del sistema.

Código	Requerimiento funcional	Descripción
RF01	Gestionar autenticación de usuarios	El sistema deberá permitir el inicio de sesión y validación de usuarios mediante credenciales seguras.
RF02	Gestionar roles y permisos	El sistema deberá administrar roles y controlar el acceso a módulos según políticas de autorización.
RF03	Registrar artículos científicos	El sistema deberá permitir el registro individual de artículos científicos y sus metadatos asociados.
RF04	Gestionar autores y coautores	El sistema deberá permitir asociar participantes y roles de autoría a cada artículo registrado.
RF05	Gestionar revistas y métricas editoriales	El sistema deberá administrar revistas, cuartiles y métricas editoriales utilizadas en reportería institucional.
RF06	Administrar formularios dinámicos	El sistema deberá permitir configurar formularios y campos dinámicos para procesos de registro.
RF07	Gestionar matrices de registro	El sistema deberá permitir registrar información mediante estructuras matriciales dinámicas.
RF08	Ejecutar procesos de carga masiva	El sistema deberá importar información científica mediante archivos estructurados y validación por lotes.
RF09	Validar registros importados	El sistema deberá detectar errores e inconsistencias durante el procesamiento de cargas masivas.
RF10	Gestionar workflow institucional	El sistema deberá administrar procesos de revisión, aprobación y devolución de registros.
RF11	Integrar información externa	El sistema deberá consultar información científica desde APIs externas como Scopus y Crossref.
RF12	Ejecutar procesos ETL	El sistema deberá transformar y cargar información operativa hacia el modelo analítico.
RF13	Generar dashboards analíticos	El sistema deberá mostrar KPIs, gráficos y métricas institucionales mediante paneles interactivos.
RF14	Exportar reportes institucionales	El sistema deberá generar reportes en formatos PDF y Excel.
RF15	Gestionar trazabilidad de procesos	El sistema deberá registrar acciones y estados asociados al flujo de validación institucional.
RF16	Gestionar catálogos institucionales	El sistema deberá administrar catálogos utilizados en formularios, reportería y validaciones.

RF17	Analizar producción científica institucional	El sistema deberá permitir analizar información científica mediante indicadores y métricas BI.
RF18	Implementar funciones de recomendación y predicción mediante IA	El sistema deberá permitir generar recomendaciones y predicciones institucionales a partir de información histórica consolidada en el modelo analítico.

Los requerimientos funcionales definidos permitieron establecer la estructura operativa del sistema propuesto para la DIDE-UTA. Entre las funcionalidades con mayor relevancia se identifican los procesos de gestión de artículos científicos, workflow institucional, carga masiva de información, integración con fuentes externas y reportería analítica basada en Business Intelligence. Asimismo, se incorporaron funcionalidades relacionadas con trazabilidad, administración dinámica de formularios y control de acceso mediante roles y permisos, fortaleciendo la gestión institucional de información científica y el control de los procesos internos. Del mismo modo, la definición de requerimientos funcionales permitió establecer una base estructurada para el diseño modular y escalable del sistema, incorporando capacidades orientadas al análisis predictivo y recomendación institucional mediante técnicas de Inteligencia Artificial.

Los requerimientos no funcionales fueron definidos con el propósito de establecer las condiciones técnicas y operativas necesarias para garantizar el correcto funcionamiento de la plataforma dentro del entorno institucional. Estos requerimientos contemplan aspectos relacionados con seguridad, rendimiento, disponibilidad, interoperabilidad, mantenibilidad, escalabilidad y capacidad de crecimiento del sistema. Asimismo, se consideraron características orientadas a la modularidad y extensibilidad de la arquitectura, facilitando la incorporación de nuevas funcionalidades, procesos analíticos y mecanismos de apoyo predictivo sin afectar el funcionamiento general de la solución tecnológica. La Tabla 9 presenta los principales requerimientos no funcionales considerados para el desarrollo de la propuesta tecnológica.

Tabla 9. Requerimientos no funcionales del sistema.

Código	Requerimiento no funcional	Descripción
RNF01	Seguridad de acceso	El sistema deberá implementar mecanismos de autenticación y autorización mediante roles y políticas de acceso seguras.
RNF02	Escalabilidad del sistema	El sistema deberá permitir la incorporación de nuevos módulos, funcionalidades y componentes sin afectar la arquitectura principal.
RNF03	Extensibilidad funcional	El sistema deberá soportar la creación de formularios dinámicos, campos configurables y catálogos institucionales adaptables.
RNF04	Rendimiento de procesamiento	El sistema deberá procesar consultas, validaciones y operaciones de carga masiva en tiempos adecuados para el entorno institucional.
RNF05	Disponibilidad de la información	El sistema deberá garantizar acceso continuo a la información institucional y a los servicios principales de reportería y gestión.
RNF06	Integridad de datos	El sistema deberá mantener consistencia y validación de información durante procesos de registro, revisión y carga masiva.
RNF07	Trazabilidad operativa	El sistema deberá registrar las acciones realizadas por los usuarios y los cambios efectuados dentro de los procesos institucionales.
RNF08	Interoperabilidad	El sistema deberá integrarse con servicios externos y APIs académicas mediante estándares de comunicación REST y JSON.
RNF09	Usabilidad	El sistema deberá presentar interfaces intuitivas y organizadas para facilitar la interacción de los usuarios institucionales.
RNF10	Compatibilidad web	El sistema deberá funcionar correctamente en navegadores web modernos y diferentes resoluciones de pantalla.
RNF11	Mantenibilidad	El sistema deberá estar estructurado mediante una arquitectura modular que facilite futuras actualizaciones y mantenimiento.
RNF12	Soporte para Business Intelligence	El sistema deberá permitir el procesamiento y análisis de información mediante modelos analíticos y reportería institucional.
RNF13	Soporte para procesamiento ETL	El sistema deberá permitir la extracción, transformación y carga de información hacia el modelo analítico institucional.

RNF14	Capacidad de crecimiento de datos	El sistema deberá soportar el incremento progresivo de registros científicos, procesos y datos analíticos institucionales.
RNF15	Generación de reportes	El sistema deberá permitir exportar información institucional en formatos PDF y Excel de manera estructurada.
RNF16	Seguridad de información institucional	El sistema deberá proteger la información almacenada mediante mecanismos de control y validación de acceso.
RNF17	Adaptabilidad tecnológica	El sistema deberá permitir integraciones con módulos de Inteligencia Artificial orientados a recomendación y predicción institucional.
RNF18	Portabilidad	El sistema deberá poder desplegarse en diferentes entornos compatibles con tecnologías web y SQL Server.

Los requerimientos no funcionales definidos permitieron garantizar que el sistema propuesto contara con características técnicas adecuadas para soportar los procesos institucionales relacionados con la gestión y análisis de producción científica. Entre los aspectos más relevantes se destacaron la escalabilidad, interoperabilidad, seguridad y mantenibilidad de la plataforma, asegurando la continuidad operativa y el crecimiento progresivo del sistema dentro del entorno institucional. Asimismo, la incorporación de una arquitectura modular facilitó la integración de nuevos módulos, procesos ETL y funcionalidades orientadas a Business Intelligence y análisis predictivo mediante Inteligencia Artificial. Del mismo modo, los requerimientos relacionados con trazabilidad, integridad y disponibilidad de la información fortalecieron la confiabilidad de los procesos gestionados por la plataforma. Finalmente, estos requerimientos constituyeron la base técnica que posteriormente orientó el diseño arquitectónico, el desarrollo modular y las pruebas del sistema propuesto.

La definición de requerimientos funcionales y no funcionales permitió establecer una base estructurada para el desarrollo de la propuesta tecnológica orientada a la gestión y análisis de información científica institucional. A partir de estos requerimientos fue posible identificar las funcionalidades principales del sistema y las condiciones técnicas necesarias para garantizar su funcionamiento, crecimiento y sostenibilidad dentro del entorno institucional. Asimismo, los requerimientos definidos sirvieron como referencia para el diseño arquitectónico, la estructuración modular y la implementación de procesos relacionados con workflow institucional, integración de

información científica, ETL, reportería analítica y Business Intelligence. Finalmente, esta información permitió orientar las etapas posteriores de diseño, desarrollo y pruebas de la solución tecnológica propuesta.

3.1.4 Diagramas UML del sistema

Los diagramas UML fueron elaborados para representar la estructura funcional, la interacción de los usuarios y el comportamiento de los principales procesos del sistema propuesto para la DIDE-UTA. Estos modelos permitieron describir funcionalidades relacionadas con gestión de artículos científicos, carga masiva, workflow institucional, integración con APIs externas, reportería analítica, administración de usuarios e Inteligencia Artificial. La utilización de diagramas de casos de uso, actividades y secuencia facilitó la comprensión de los módulos que conforman la plataforma y sirvió como apoyo para las etapas de diseño arquitectónico, desarrollo e implementación del sistema.

a. Diagrama de casos de uso general

El diagrama de casos de uso general representa la interacción entre los actores principales y las funcionalidades globales del sistema propuesto para la DIDE-UTA. Como se observa en la Figura 12, la plataforma integra procesos de gestión de artículos científicos, carga masiva, workflow institucional, reportería analítica, administración de usuarios, procesos ETL e Inteligencia Artificial. Este diagrama permite visualizar de forma general la relación entre los usuarios institucionales y los módulos funcionales que conforman la solución tecnológica.

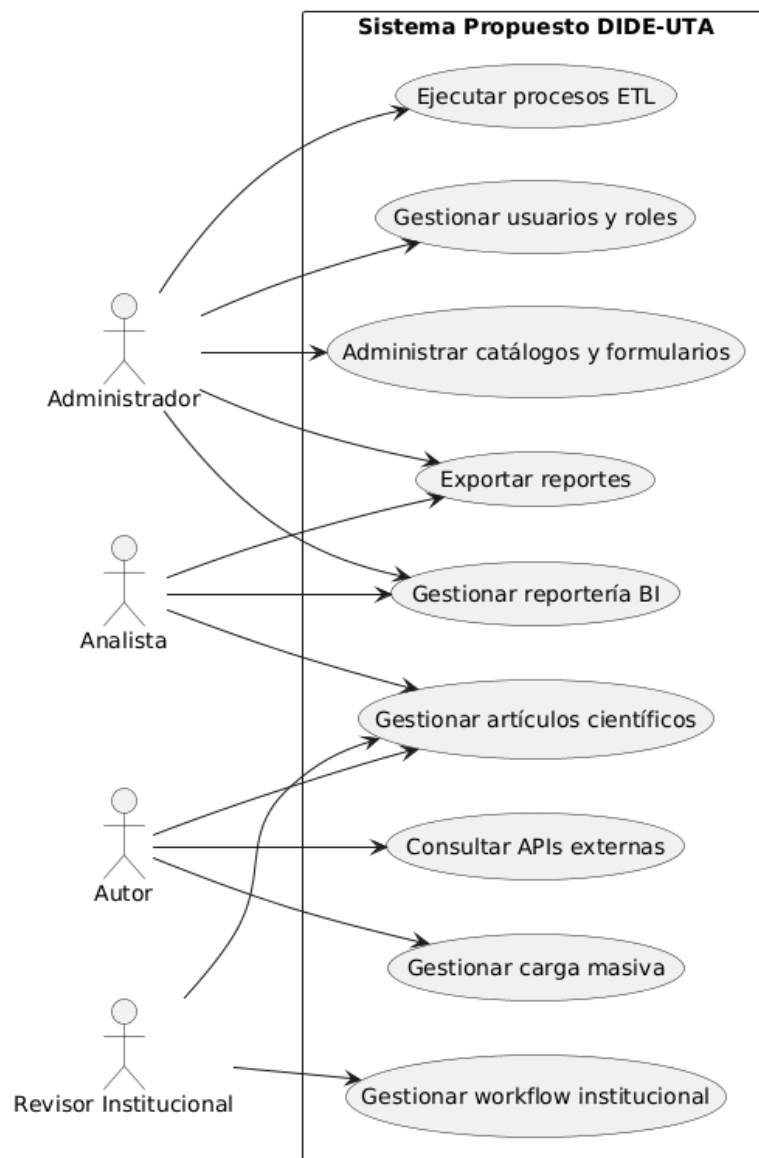


Figura 12. Diagrama de casos de uso general del sistema propuesto.

b. Diagrama de casos de uso por módulo

Los diagramas de casos de uso por módulo permiten detallar las funcionalidades principales de cada componente del sistema. A través de estos modelos se describen las acciones realizadas por los actores institucionales, las condiciones necesarias para ejecutar cada proceso y los resultados esperados dentro de la plataforma. Esta representación facilita comprender el alcance funcional de los módulos de registro,

carga masiva, workflow, integración externa, reportería, seguridad e Inteligencia Artificial.

Módulo de registro de artículos

El módulo de registro de artículos permite almacenar información bibliográfica, académica y editorial relacionada con la producción científica institucional. Como se observa en la Figura 13, este módulo contempla el registro de artículos, autores, revistas, indexaciones, clasificación OCDE y validación de información antes de su envío al flujo de revisión institucional.

CU-03 Registrar artículo científico

Actor principal: Analista DIDE / Investigador / Autor

Descripción:

Permite registrar información relacionada con artículos científicos dentro del sistema institucional, incluyendo datos bibliográficos, autores, métricas editoriales, clasificación OCDE y metadatos asociados para su posterior validación y procesamiento institucional.

Precondiciones:

- Usuario autenticado.
- Catálogos institucionales cargados.
- Permisos de acceso habilitados.

Flujo principal:

- El usuario accede al módulo de registro de artículos.
- El sistema carga catálogos y formularios dinámicos.
- El usuario ingresa información bibliográfica del artículo.
- El usuario registra autores y coautores.
- El usuario selecciona revista, indexación y clasificación OCDE.
- El sistema valida la información ingresada.
- El usuario confirma el registro.
- El sistema almacena la información.
- El sistema permite enviar el artículo al flujo de revisión institucional.

Flujos alternos:

- Si el DOI ya existe, el sistema muestra un mensaje de advertencia.
- Si existen campos obligatorios incompletos, el sistema bloquea el registro.
- Si la revista no existe, el sistema permite registrarla o actualizarla.

Postcondiciones:

- Artículo registrado correctamente.
- Información disponible para validación y reportería institucional.

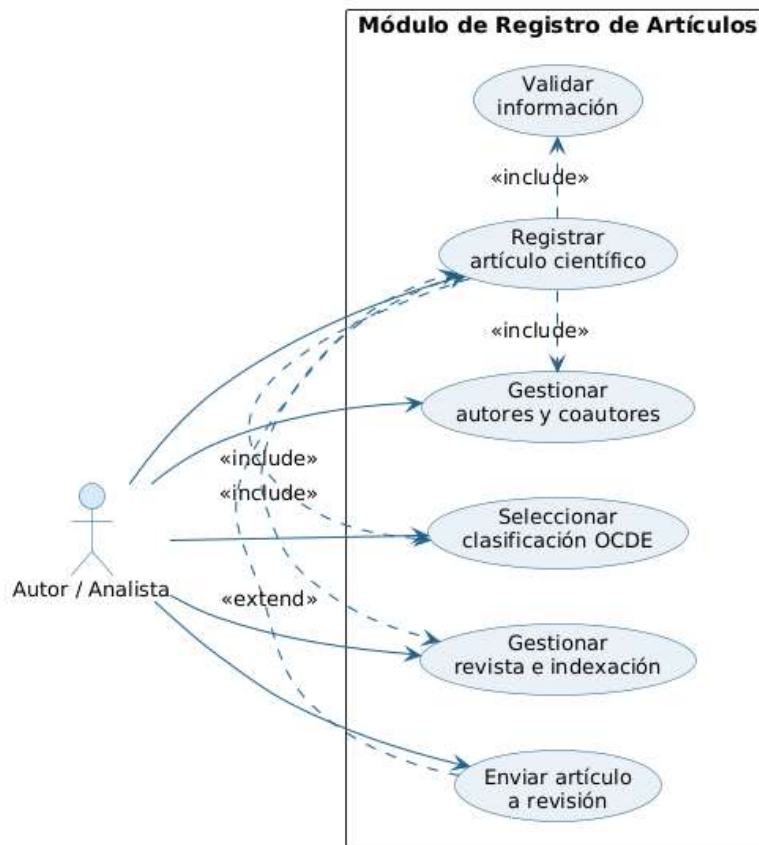


Figura 13. Diagrama de casos de uso del módulo de registro de artículos científicos.

El diagrama resume las funcionalidades principales del registro de artículos científicos, desde el ingreso de información bibliográfica hasta la validación inicial de

los datos. Este módulo permite mantener información estructurada y disponible para procesos posteriores de revisión, reportería institucional y análisis mediante BI.

Módulo de carga masiva de información

El módulo de carga masiva permite importar registros científicos mediante archivos estructurados, facilitando el procesamiento de grandes volúmenes de información académica. Como se observa en la Figura 14, el proceso contempla la descarga de una plantilla institucional, carga del archivo, validación de filas, corrección de inconsistencias y procesamiento de los registros válidos.

CU-04 Cargar información masiva desde archivos estructurados

Actor principal: Analista DIDE / Administrador

Descripción:

Permite importar artículos científicos y registros académicos mediante archivos estructurados, validando la información y gestionando procesos de carga institucional por lotes.

Precondiciones:

- Usuario autenticado.
- Archivo válido.
- Plantilla institucional disponible.
- Permisos habilitados para carga masiva.

Flujo principal:

- El usuario accede al módulo de carga masiva.
- El sistema permite descargar la plantilla institucional.
- El usuario selecciona el archivo de carga.
- El sistema crea un lote de importación.
- El sistema valida filas y estructura de datos.
- El sistema detecta errores e inconsistencias.
- El usuario corrige filas inválidas si existen errores.
- El usuario confirma el procesamiento del lote.
- El sistema registra la información válida.

- El sistema actualiza el estado del lote.

Flujos alternos:

- Si el archivo posee errores estructurales, el sistema bloquea el procesamiento.
- Si existen filas inválidas, el sistema permite corregirlas antes de procesar el lote.
- Si el lote ya fue procesado, el sistema impide una nueva ejecución.

Postcondiciones:

- Información almacenada correctamente.
- Lote procesado y registrado.
- Datos disponibles para workflow y reportería institucional.

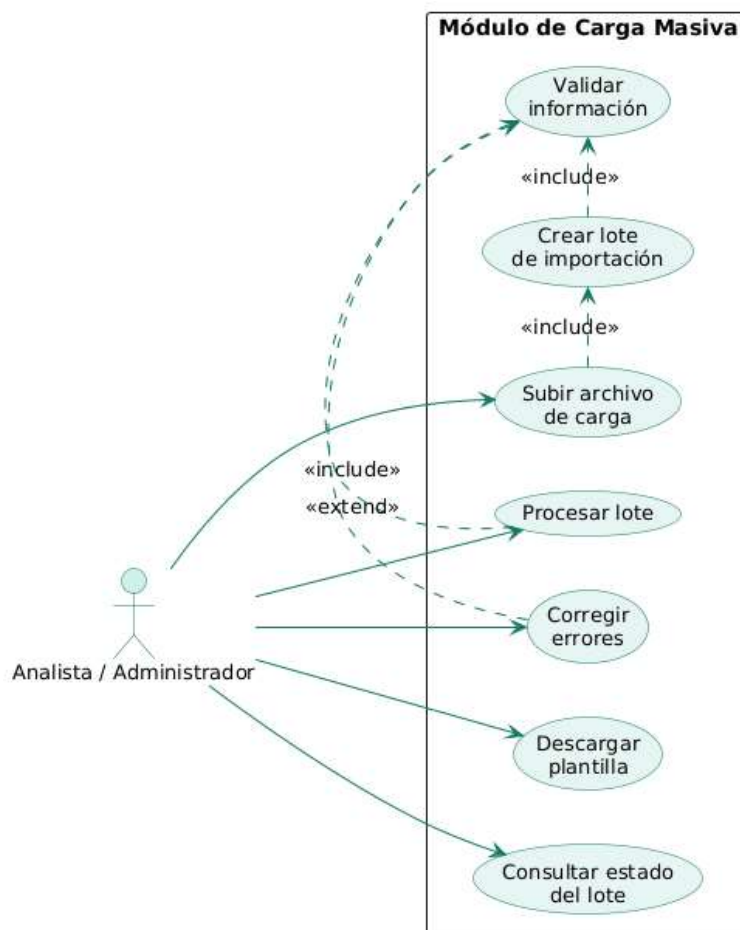


Figura 14. Diagrama de casos de uso del módulo de carga masiva de información.

El diagrama resume el proceso de carga masiva desde la selección del archivo hasta la validación y consolidación de registros válidos. Este módulo contribuye a reducir el ingreso manual de información y permite mantener control sobre los lotes importados antes de su uso en workflow y reportería institucional.

Módulo de workflow y revisión institucional

El módulo de workflow institucional permite gestionar la revisión, validación, aprobación o devolución de la información científica registrada en el sistema. Como se observa en la Figura 15, este módulo incorpora control de estados, observaciones y trazabilidad para administrar el flujo de revisión entre los actores responsables.

CU-05 Gestionar workflow institucional de revisión

Actor principal: Revisor institucional / Área Técnica / Analista DIDE

Descripción:

Permite administrar el flujo institucional de revisión y validación de artículos científicos y lotes de información importados dentro del sistema.

Precondiciones:

- Usuario autenticado.
- Existencia de artículos o lotes enviados a revisión.
- Permisos habilitados según rol institucional.

Flujo principal:

- El usuario accede al módulo de workflow institucional.
- El sistema muestra la bandeja de revisión disponible.
- El revisor selecciona un artículo o lote.
- El sistema muestra el detalle de la información.
- El revisor analiza los datos registrados.
- El usuario registra observaciones si existen inconsistencias.
- El usuario aprueba o devuelve la información.
- El sistema actualiza el estado del flujo.
- El sistema registra la trazabilidad del proceso.

Flujos alternos:

- Si la información presenta inconsistencias, el sistema permite devolver el registro al autor.
- Si el lote ya fue tomado por otro revisor, el sistema bloquea la asignación.
- Si la validación es aprobada, el sistema continúa con la siguiente etapa del flujo institucional.

Postcondiciones:

- Estado de revisión actualizado.
- Historial de validación registrado.
- Información disponible para reportería y procesos institucionales posteriores.

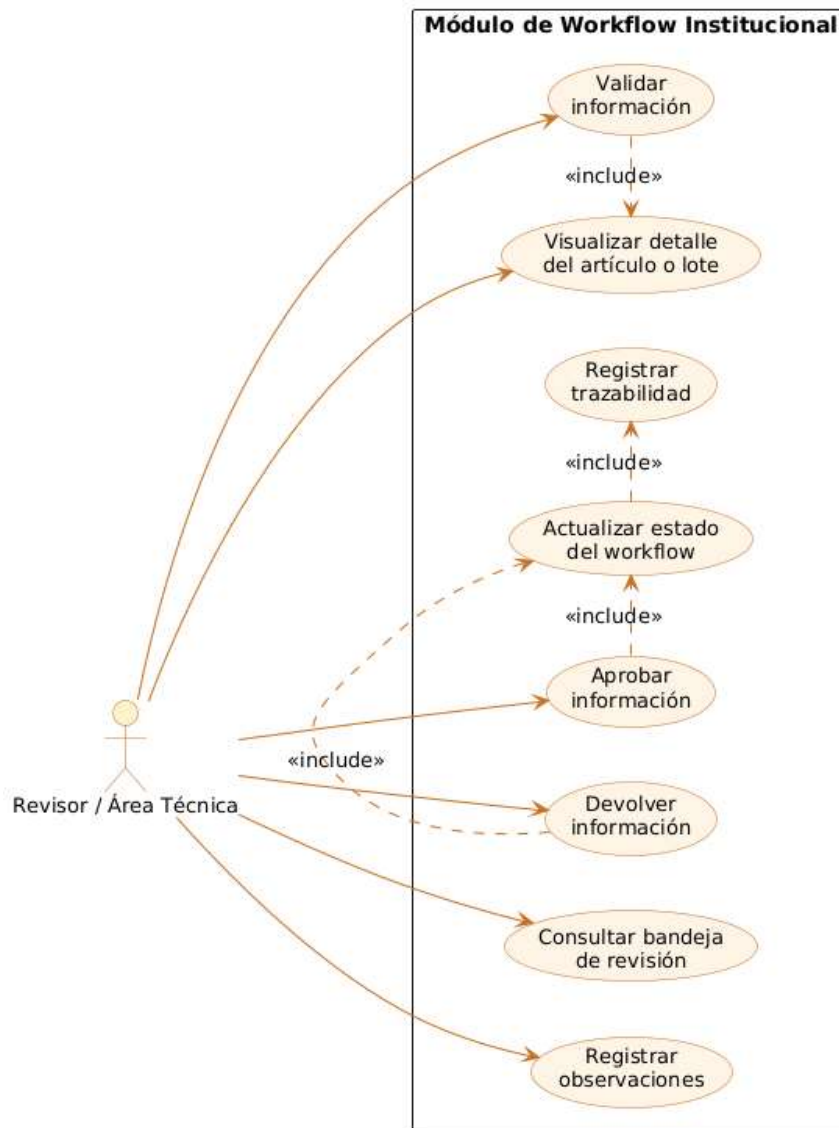


Figura 15. Diagrama de casos de uso del módulo de workflow y revisión institucional.

El diagrama representa el proceso de revisión institucional aplicado a artículos científicos o lotes de información. A través de este módulo, los revisores pueden analizar registros, emitir observaciones, aprobar información o devolverla para corrección, manteniendo trazabilidad sobre cada acción realizada.

Módulo de integración con APIs externas

El módulo de integración con APIs externas permite consultar información científica desde plataformas académicas como Scopus, Crossref, OpenAlex y Semantic Scholar.

Como se observa en la Figura 16, este proceso facilita la búsqueda por DOI, autor, afiliación o criterios textuales, permitiendo enriquecer y validar registros institucionales.

CU-06 Consultar información científica mediante APIs externas

Actor principal: Analista DIDE / Autor / Administrador

Descripción:

Permite consultar información científica desde APIs académicas externas para obtener metadatos bibliográficos, validar registros institucionales y enriquecer información científica dentro del sistema.

Precondiciones:

- Usuario autenticado.
- Conexión disponible con APIs externas.
- Proveedor de consulta configurado.

Flujo principal:

- El usuario accede al módulo de integración externa.
- El usuario selecciona el proveedor de información.
- El usuario ingresa criterios de búsqueda.
- El sistema construye la consulta externa.
- El sistema realiza la petición a la API seleccionada.
- El sistema recibe y normaliza los resultados.
- El usuario visualiza la información obtenida.
- El usuario selecciona registros para utilizar o importar.
- El sistema almacena la información seleccionada.

Flujos alternos:

- Si la API externa no responde, el sistema muestra un mensaje de error.
- Si no existen resultados, el sistema informa la ausencia de coincidencias.
- Si la información presenta inconsistencias, el sistema solicita validación manual.

Postcondiciones:

- Información científica consultada correctamente.
- Datos disponibles para registro, validación y reportería institucional.

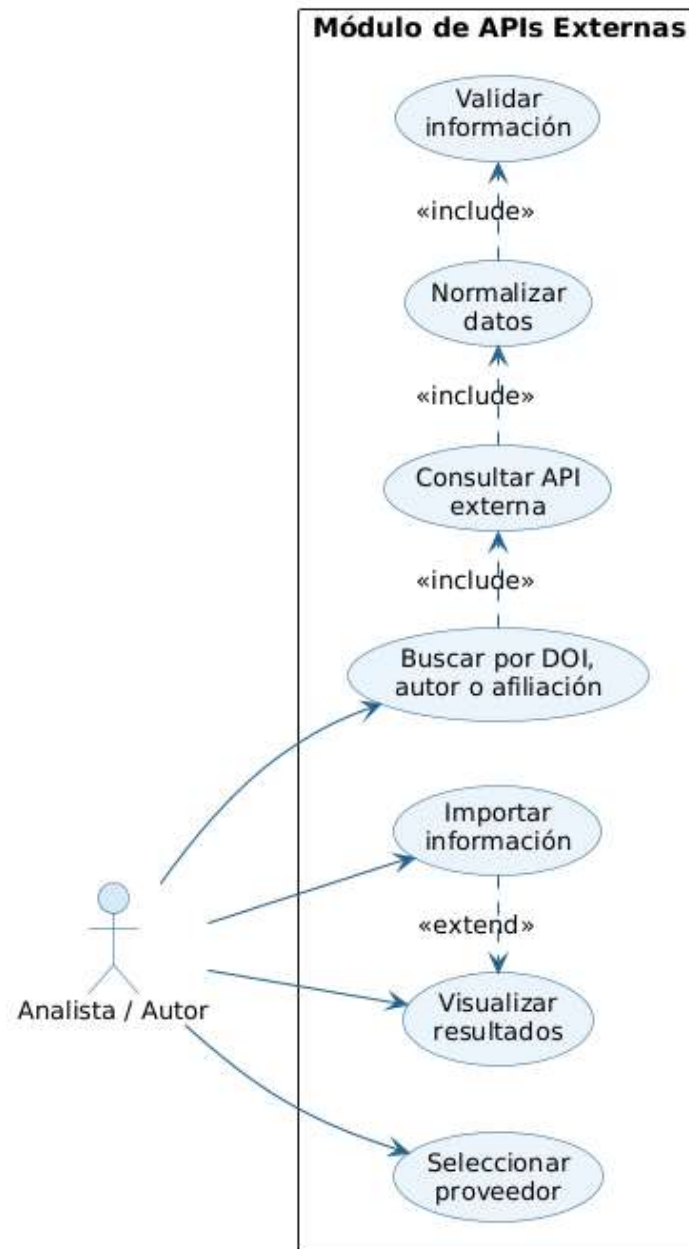


Figura 16. Diagrama de casos de uso del módulo de integración con APIs externas.

El diagrama resume la interacción del usuario con los proveedores externos de información científica. Este módulo permite obtener metadatos académicos,

normalizar resultados y seleccionar información relevante para su almacenamiento dentro del sistema institucional.

Módulo de reportería y Business Intelligence

El módulo de reportería y Business Intelligence permite visualizar información analítica mediante dashboards, KPIs, gráficos y reportes institucionales. Como se observa en la Figura 17, este módulo consulta información consolidada desde el modelo analítico y permite aplicar filtros, generar indicadores y exportar reportes en formatos PDF o Excel.

CU-07 Visualizar reportería analítica institucional

Actor principal: Analista DIDE / Autoridad institucional / Administrador

Descripción:

Permite visualizar información analítica institucional mediante dashboards interactivos, indicadores de producción científica y reportes dinámicos basados en el modelo analítico institucional.

Precondiciones:

- Usuario autenticado.
- Modelo analítico disponible.
- Procesos ETL ejecutados correctamente.

Flujo principal:

- El usuario accede al módulo de reportería institucional.
- El sistema consulta el estado del modelo analítico.
- El sistema carga dashboards e indicadores institucionales.
- El usuario aplica filtros de análisis.
- El sistema consulta información agregada del Data Warehouse.
- El sistema genera KPIs, gráficos y tablas analíticas.
- El usuario visualiza los resultados.
- El usuario exporta reportes si lo requiere.

Flujos alternos:

- Si el modelo analítico no está disponible, el sistema informa la inconsistencia.
- Si no existen datos para el filtro seleccionado, el sistema muestra una notificación.
- Si ocurre un error durante la exportación, el sistema cancela la generación del archivo.

Postcondiciones:

- Información analítica visualizada correctamente.
- Reportes institucionales disponibles para exportación.
- Datos listos para análisis y toma de decisiones.

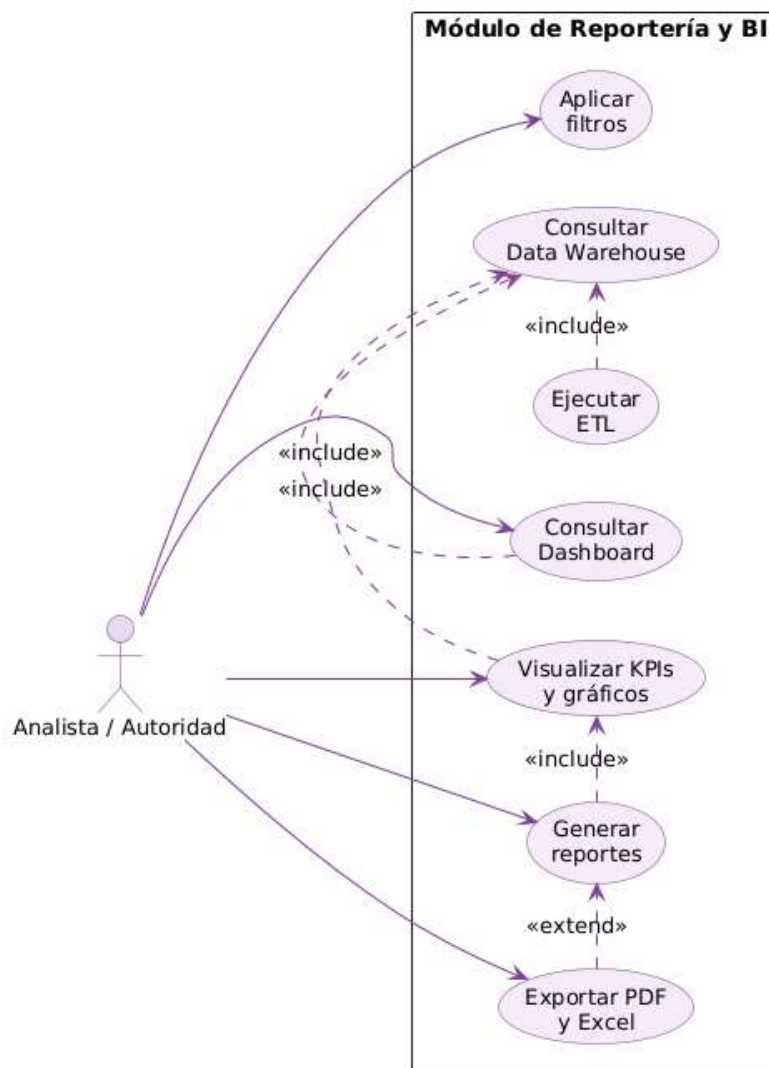


Figura 17. Diagrama de casos de uso del módulo de reportería y Business Intelligence.

El diagrama representa las funcionalidades principales del módulo de reportería, desde la consulta del modelo analítico hasta la visualización y exportación de resultados. Este componente permite convertir la información consolidada en indicadores útiles para el seguimiento institucional y la toma de decisiones.

Módulo de administración y seguridad

El módulo de administración y seguridad permite gestionar usuarios, roles, permisos y configuraciones generales del sistema. Como se observa en la Figura 18, este componente controla el acceso a la plataforma y registra acciones administrativas para fortalecer la seguridad y trazabilidad institucional.

CU-08 Gestionar usuarios, roles y seguridad del sistema

Actor principal: Administrador

Descripción:

Permite administrar usuarios, roles, permisos y configuraciones de seguridad dentro del sistema institucional.

Precondiciones:

- Administrador autenticado.
- Permisos administrativos habilitados.

Flujo principal:

- El administrador accede al módulo de administración.
- El sistema muestra la gestión de usuarios y roles.
- El administrador registra, actualiza o desactiva usuarios.
- El administrador asigna roles y permisos.
- El sistema valida las configuraciones de acceso.
- El sistema almacena los cambios realizados.
- El sistema registra la trazabilidad de acciones administrativas.

Flujos alternos:

- Si el usuario ya existe, el sistema muestra una advertencia.
- Si existen inconsistencias en permisos, el sistema bloquea la asignación.

- Si ocurre un error durante el guardado, el sistema cancela el proceso.

Postcondiciones:

- Usuarios y permisos actualizados correctamente.
- Configuraciones de seguridad almacenadas.
- Historial administrativo registrado.

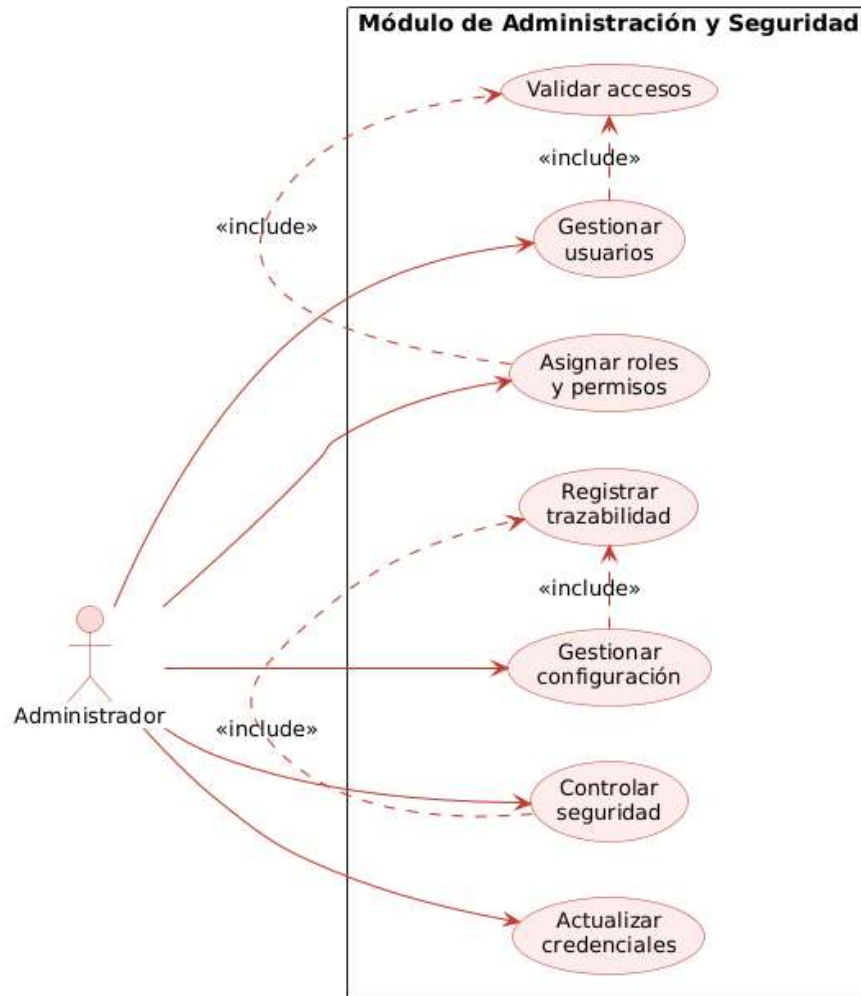


Figura 18. Diagrama de casos de uso del módulo de administración y seguridad.

El diagrama resume las acciones principales del administrador dentro del sistema, incluyendo la gestión de usuarios, asignación de roles, control de permisos y registro de trazabilidad. Este módulo garantiza que el acceso a la información institucional se realice de forma controlada y segura.

Módulo de Inteligencia Artificial para recomendación y predicción

El módulo de Inteligencia Artificial contempla funcionalidades orientadas al análisis predictivo y generación de recomendaciones institucionales a partir de información histórica consolidada. Como se observa en la Figura 19, este proceso permite seleccionar un tipo de análisis, consultar datos procesados, ejecutar modelos de IA y presentar resultados al usuario.

CU-09 Obtener recomendaciones y predicciones institucionales mediante IA

Actor principal: Analista DIDE / Autoridad institucional

Descripción:

Permite generar recomendaciones estratégicas y predicciones institucionales mediante modelos de Inteligencia Artificial entrenados con información histórica de producción científica.

Precondiciones:

- Usuario autenticado.
- Modelo analítico disponible.
- Datos históricos procesados.
- Modelos de IA configurados o entrenados.

Flujo principal:

- El usuario accede al módulo de Inteligencia Artificial.
- El sistema carga los modelos analíticos disponibles.
- El usuario selecciona el tipo de análisis.
- El usuario define criterios y parámetros de predicción.
- El sistema procesa la información histórica institucional.
- El sistema ejecuta los modelos de recomendación o predicción.
- El sistema genera resultados analíticos.
- El usuario visualiza recomendaciones, tendencias y predicciones institucionales.

Flujos alternos:

- Si no existen suficientes datos históricos, el sistema informa la limitación analítica.
- Si el modelo no se encuentra disponible, el sistema cancela el proceso.
- Si ocurre un error durante el procesamiento, el sistema notifica la inconsistencia.

Postcondiciones:

- Resultados predictivos generados correctamente.
- Recomendaciones institucionales disponibles para análisis estratégico.

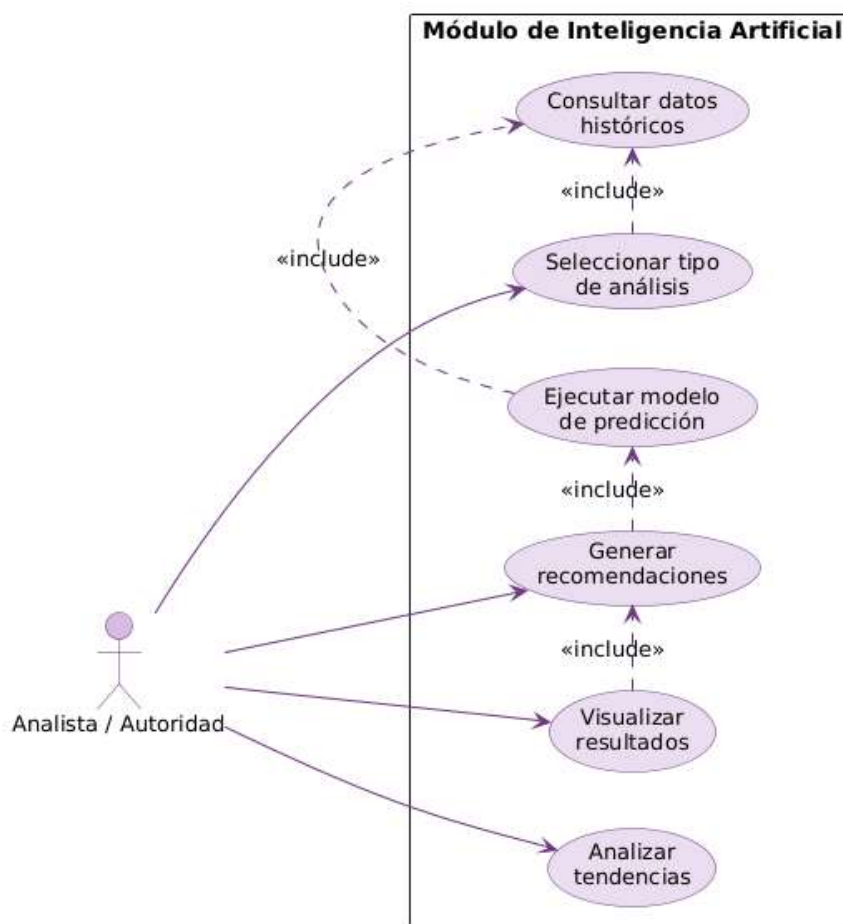


Figura 19. Diagrama de casos de uso del módulo de Inteligencia Artificial para recomendación y predicción.

El diagrama representa el flujo previsto para la generación de recomendaciones y predicciones institucionales. Este módulo utiliza información histórica del modelo analítico para identificar tendencias y apoyar la toma de decisiones relacionadas con la producción científica de la DIDE-UTA.

c. Diagrama de actividades

Los diagramas de actividades fueron elaborados para representar el flujo operativo de los procesos principales del sistema propuesto para la DIDE-UTA. Estos modelos permiten visualizar la secuencia de acciones, decisiones y validaciones realizadas en módulos como registro de artículos, carga masiva, workflow institucional, integración externa, reportería e Inteligencia Artificial. Su uso facilitó la comprensión de la lógica funcional del sistema y permitió identificar puntos de control necesarios para el desarrollo e implementación de la plataforma.

Diagrama de actividades del registro de artículos científicos

El proceso de registro de artículos científicos permite ingresar, validar y clasificar información bibliográfica, académica e institucional relacionada con la producción científica de la universidad. Como se observa en la Figura 20, el flujo contempla el ingreso de datos, la asociación de autores, la selección de clasificaciones académicas y el envío del registro al workflow institucional.

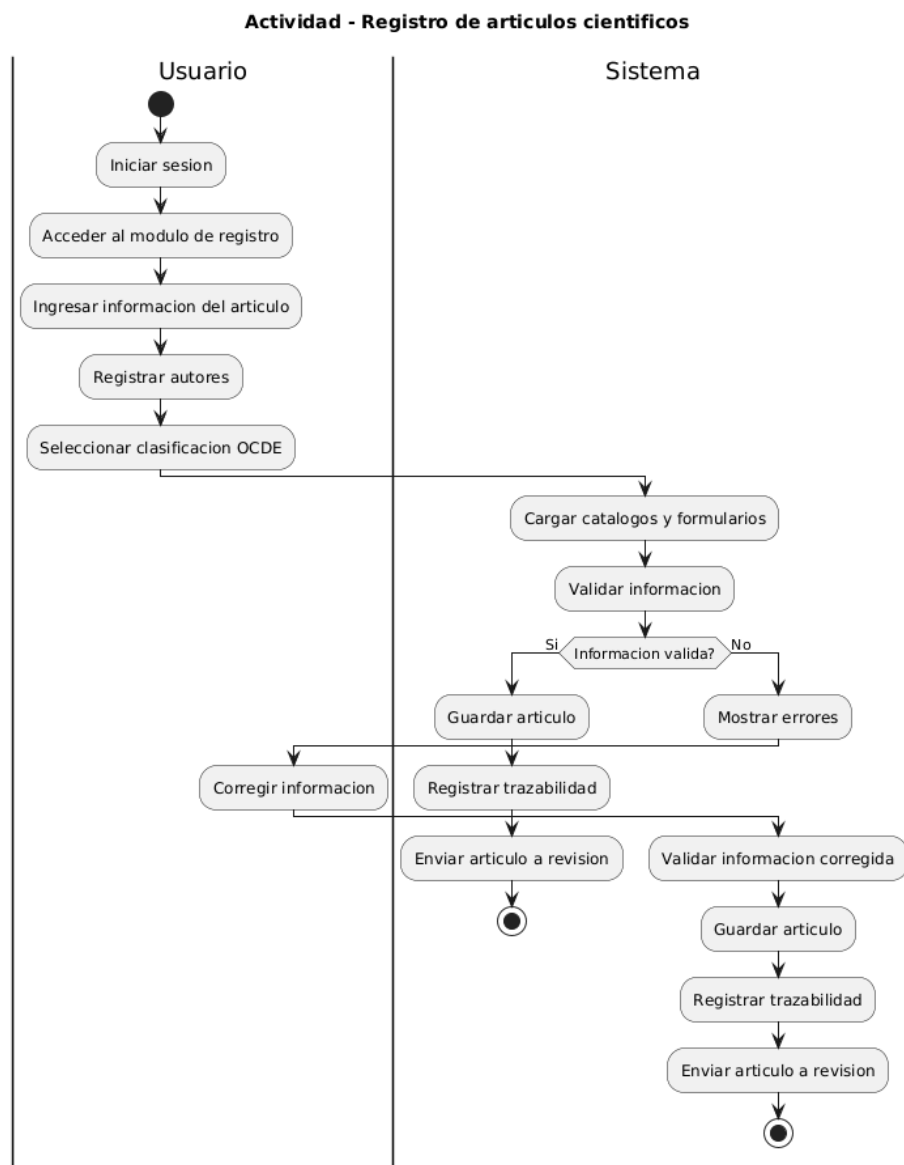


Figura 20. Diagrama de actividades del proceso de registro de artículos científicos.

El diagrama muestra la secuencia de actividades necesarias para registrar un artículo científico dentro del sistema. Este flujo permite controlar la información ingresada antes de su revisión institucional, garantizando mayor orden y consistencia en los datos registrados.

Diagrama de actividades del proceso de carga masiva

El proceso de carga masiva permite importar registros científicos mediante archivos estructurados, reduciendo el ingreso manual de información. Como se observa en la Figura 21, el flujo contempla la descarga de la plantilla, carga del archivo, validación de filas, corrección de inconsistencias y procesamiento de los registros válidos.

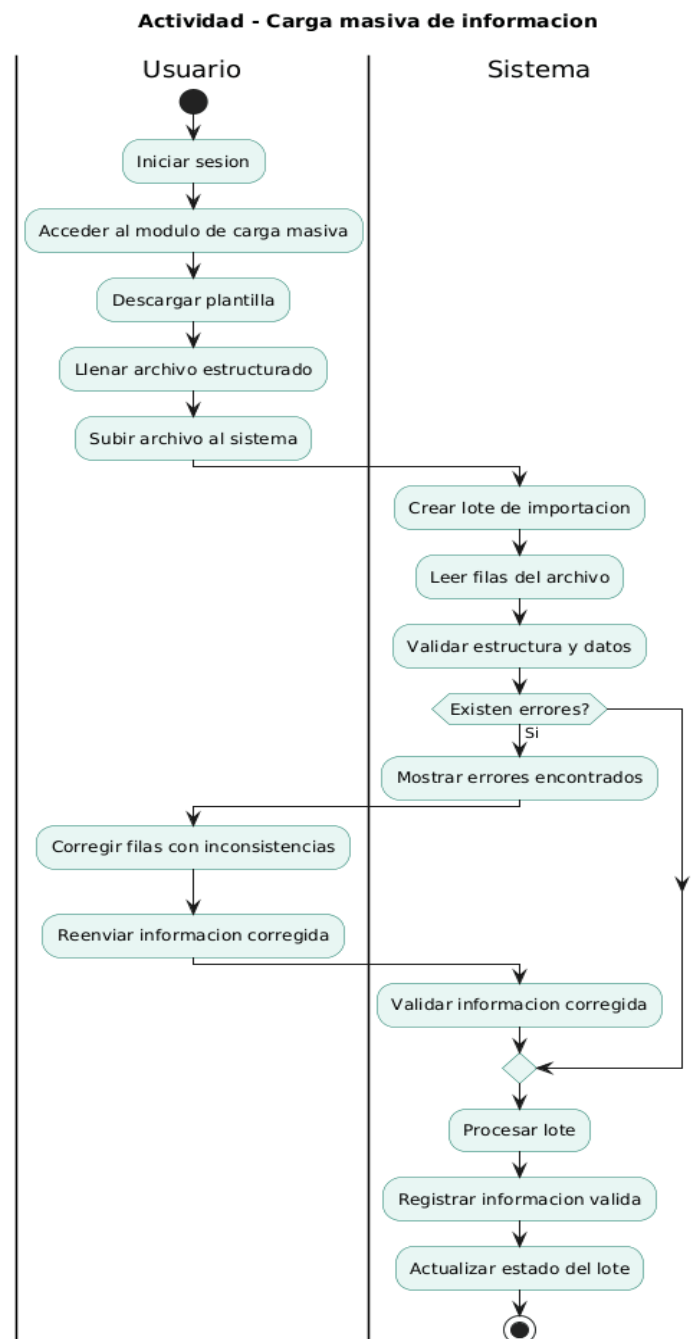


Figura 21. Diagrama de actividades del proceso de carga masiva de información.

El diagrama resume el flujo de procesamiento de archivos de carga masiva, desde la selección del archivo hasta la consolidación de la información válida. Este proceso permite controlar errores, gestionar lotes y mejorar la eficiencia del registro institucional de artículos científicos.

Diagrama de actividades del workflow de revisión

El workflow de revisión permite controlar la validación institucional de artículos o lotes de información registrados en el sistema. Como se observa en la Figura 22, el flujo contempla la recepción del registro, revisión por parte del responsable, aprobación, devolución con observaciones y actualización del estado correspondiente.

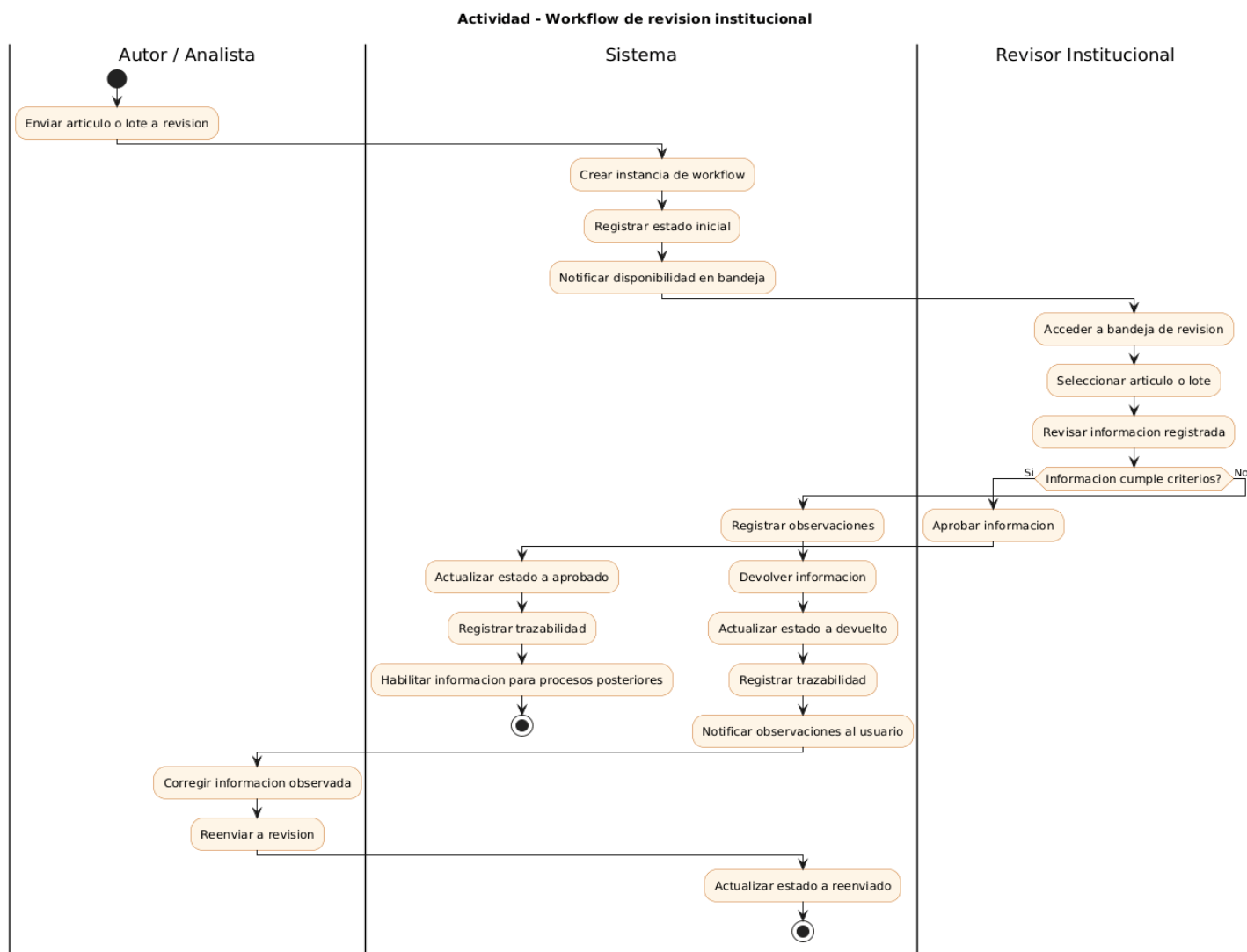


Figura 22. Diagrama de actividades del workflow de revisión institucional.

El diagrama representa el flujo de revisión institucional aplicado a la información científica registrada. A través de este proceso, el revisor puede aprobar registros válidos o devolver aquellos que presentan inconsistencias, manteniendo trazabilidad sobre cada decisión realizada.

Diagrama de actividades de consulta mediante APIs externas

El proceso de consulta mediante APIs externas permite obtener información científica desde plataformas académicas para complementar y validar registros institucionales. Como se observa en la Figura 23, el flujo incluye la selección del proveedor, ejecución de la consulta, normalización de resultados y almacenamiento de la información seleccionada.

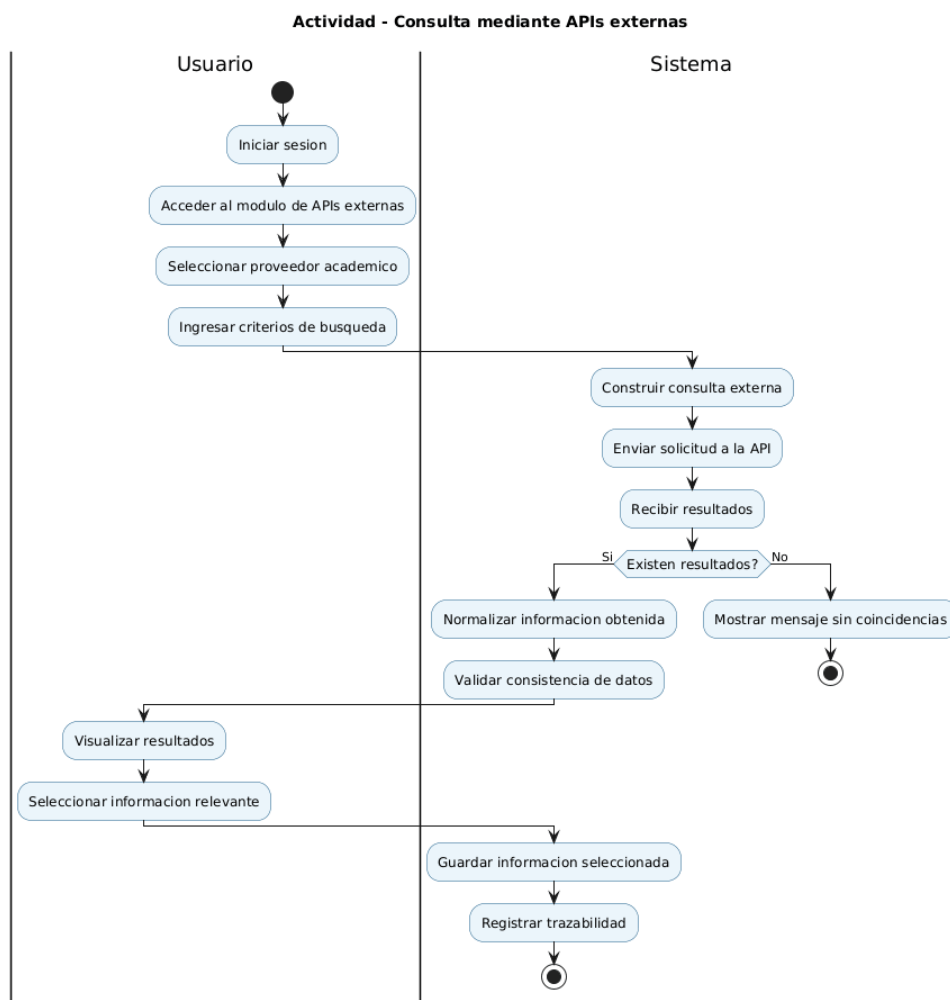


Figura 23. Diagrama de actividades del proceso de consulta mediante APIs externas.

El diagrama muestra el flujo de consulta externa utilizado para obtener metadatos científicos mediante criterios como DOI, autor, afiliación o búsqueda textual. Este proceso permite enriquecer los registros institucionales y mejorar la calidad de la información almacenada en la plataforma.

Diagrama de actividades de generación de reportes y ETL

El proceso de generación de reportes y ejecución ETL permite actualizar el modelo analítico institucional y visualizar información consolidada mediante dashboards, KPIs y reportes exportables. Como se observa en la Figura 24, el flujo contempla la verificación del Data Warehouse, la ejecución del ETL cuando sea necesario, la carga de indicadores y la generación de reportes institucionales.

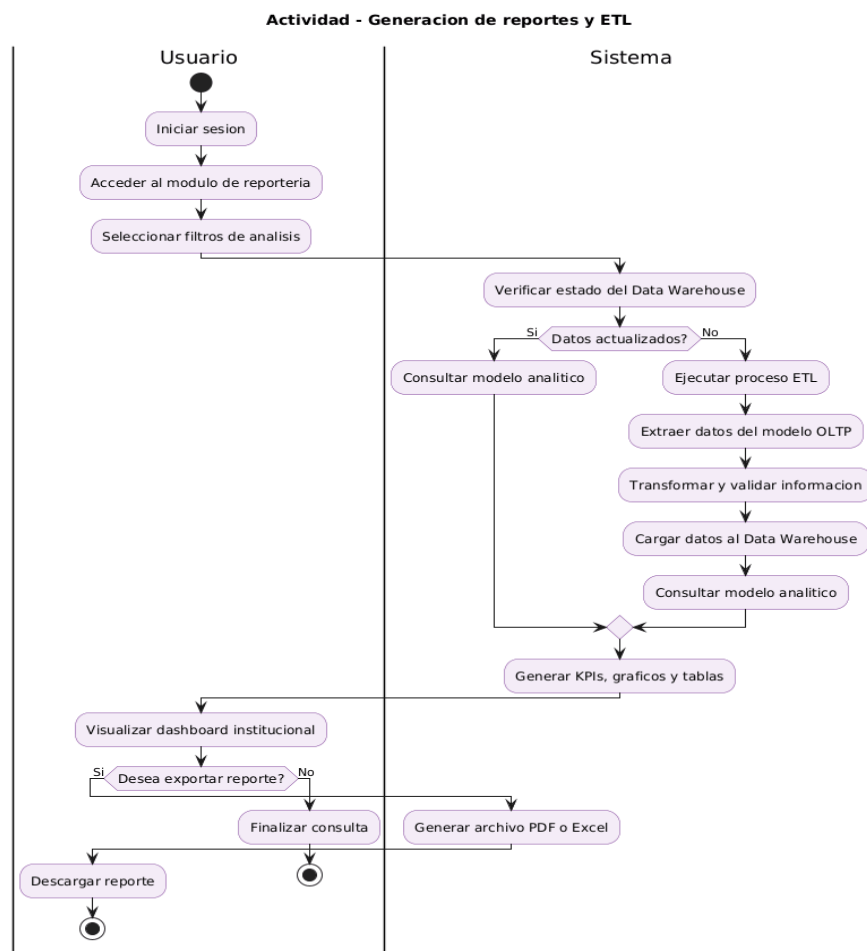


Figura 24. Diagrama de actividades del proceso de generación de reportes y ETL.

El diagrama representa el flujo de actualización y consulta del modelo analítico institucional. Este proceso permite transformar datos operativos en información consolidada para dashboards, filtros analíticos y reportes utilizados en el seguimiento de la producción científica.

Diagrama de actividades de recomendación y predicción mediante IA

El proceso de recomendación y predicción mediante Inteligencia Artificial utiliza información histórica y analítica para generar resultados de apoyo a la toma de decisiones institucionales. Como se observa en la Figura 25, el flujo contempla la selección del tipo de análisis, consulta de datos consolidados, ejecución del modelo y visualización de recomendaciones o predicciones.

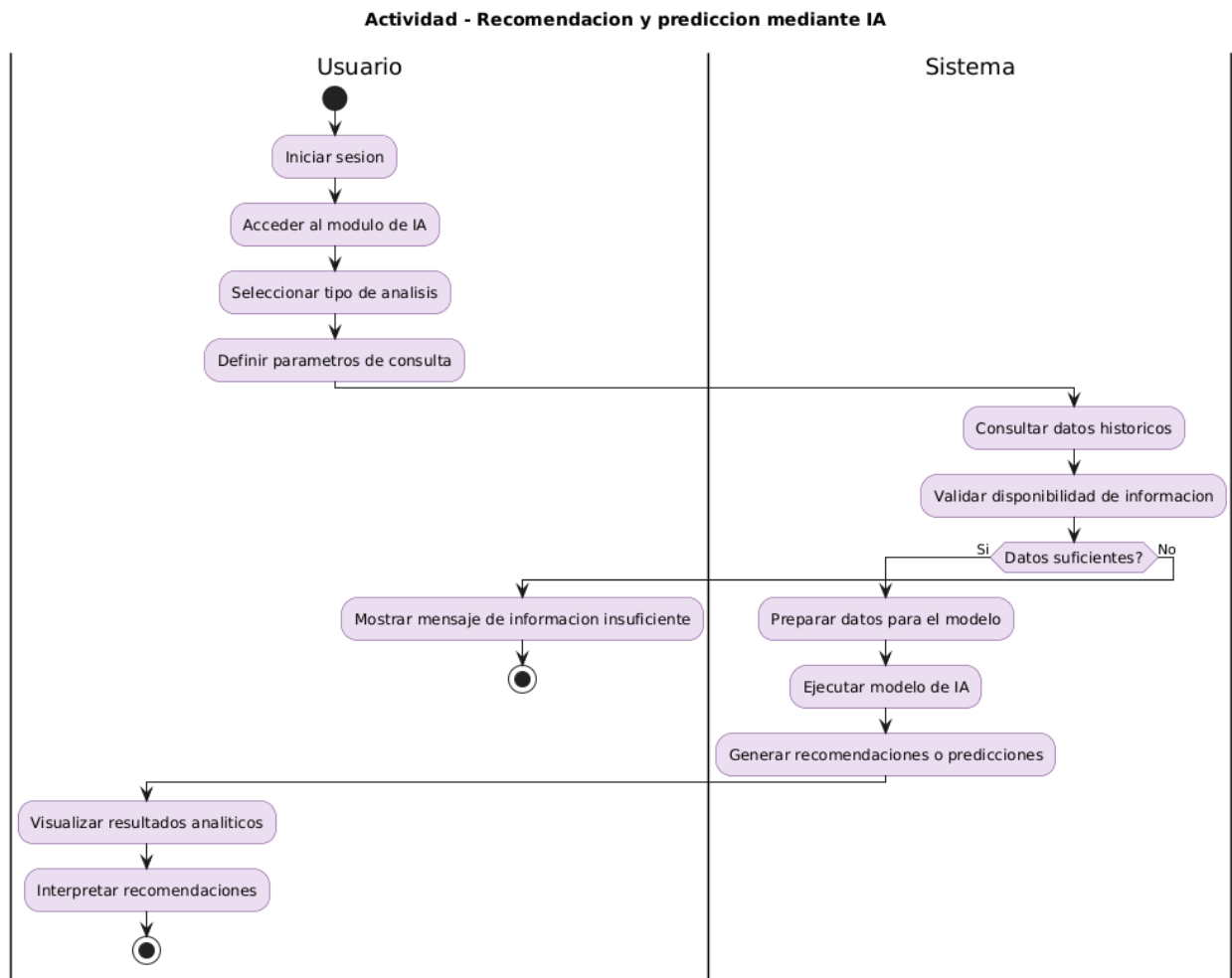


Figura 25. Diagrama de actividades del proceso de recomendación y predicción mediante Inteligencia Artificial.

El diagrama representa el flujo previsto para ejecutar procesos de análisis avanzado mediante Inteligencia Artificial. Este procedimiento valida la disponibilidad de datos, procesa la información histórica y presenta resultados orientados a identificar tendencias o recomendaciones institucionales.

d. Diagramas de secuencia

Los diagramas de secuencia fueron elaborados para representar la interacción temporal entre los actores, la interfaz del sistema, los servicios internos y las fuentes de datos que intervienen en los procesos principales de la plataforma. Estos modelos permiten observar el intercambio de mensajes entre componentes como frontend, backend, base de datos transaccional, Data Warehouse, servicios ETL, APIs externas y módulos de Inteligencia Artificial. Su aplicación facilitó la comprensión técnica del comportamiento interno del sistema y complementó los diagramas de actividades mediante una visión orientada a la comunicación entre capas.

Diagrama de secuencia del registro de artículos científicos

El proceso de registro de artículos científicos involucra la interacción entre el usuario, la interfaz del sistema, los servicios de validación y la base de datos transaccional. Como se observa en la Figura 26, el flujo contempla la carga de formularios, envío de información bibliográfica, validación de datos, almacenamiento del registro y actualización de la trazabilidad institucional.

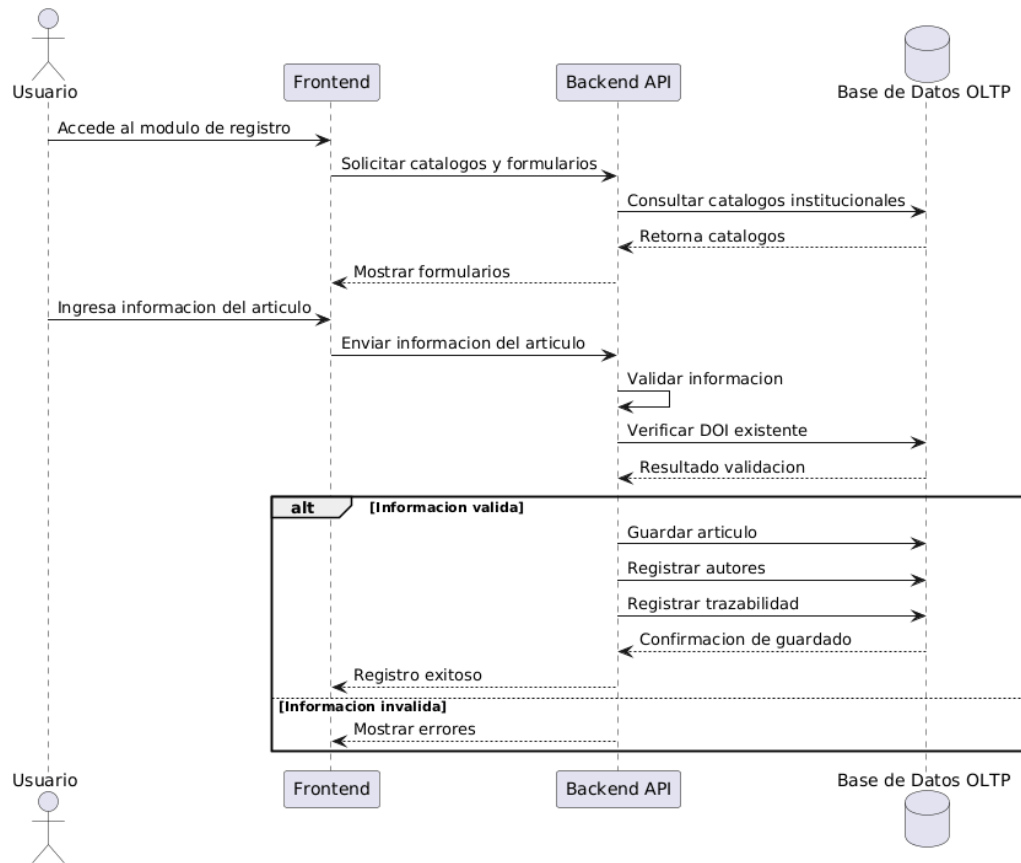


Figura 26. Diagrama de secuencia del proceso de registro de artículos científicos.

El diagrama muestra la comunicación entre el usuario, el frontend, los servicios internos y la base de datos durante el registro de un artículo científico. Esta interacción permite validar la información ingresada, evitar duplicidad de registros y almacenar los datos necesarios para su posterior revisión institucional.

Diagrama de secuencia del proceso de carga masiva de información

El proceso de carga masiva permite importar registros científicos mediante archivos estructurados y validar su contenido antes del procesamiento definitivo. Como se observa en la Figura 27, el flujo integra la interacción entre el usuario, el frontend, el backend, el servicio de validación, el lote de importación y la base de datos OLTP.

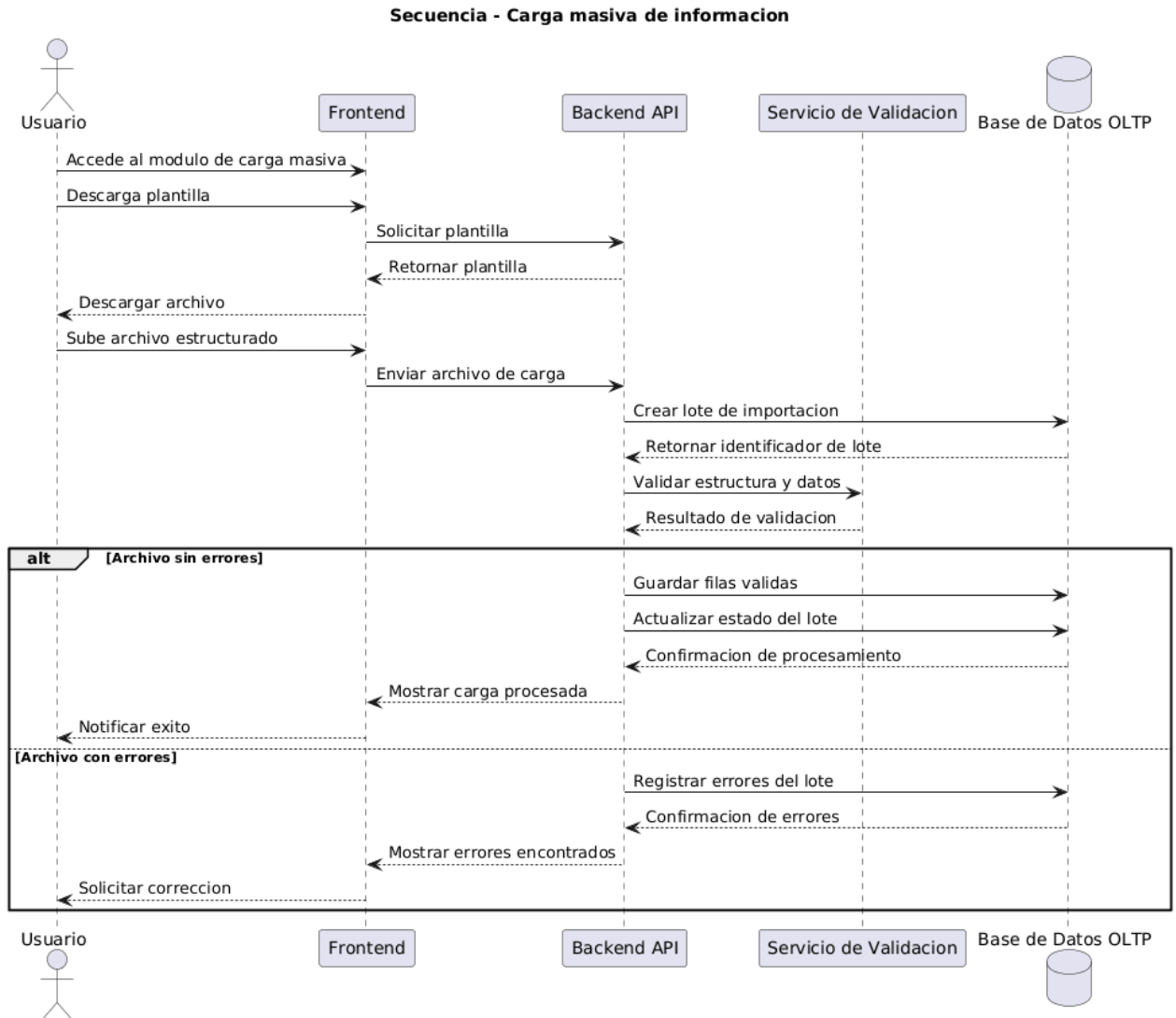


Figura 27. Diagrama de secuencia del proceso de carga masiva de información.

El diagrama representa el intercambio de mensajes necesario para procesar archivos de carga masiva dentro del sistema. En este flujo, el backend coordina la creación del lote, la validación de filas, el registro de errores y la consolidación de los datos válidos en el modelo transaccional.

Diagrama de secuencia del proceso de consulta mediante APIs externas

El proceso de consulta mediante APIs externas permite obtener información científica desde proveedores académicos para enriquecer los registros institucionales. Como se observa en la Figura 28, el flujo integra la interacción entre el usuario, el frontend, el backend, el proveedor externo y la base de datos OLTP.

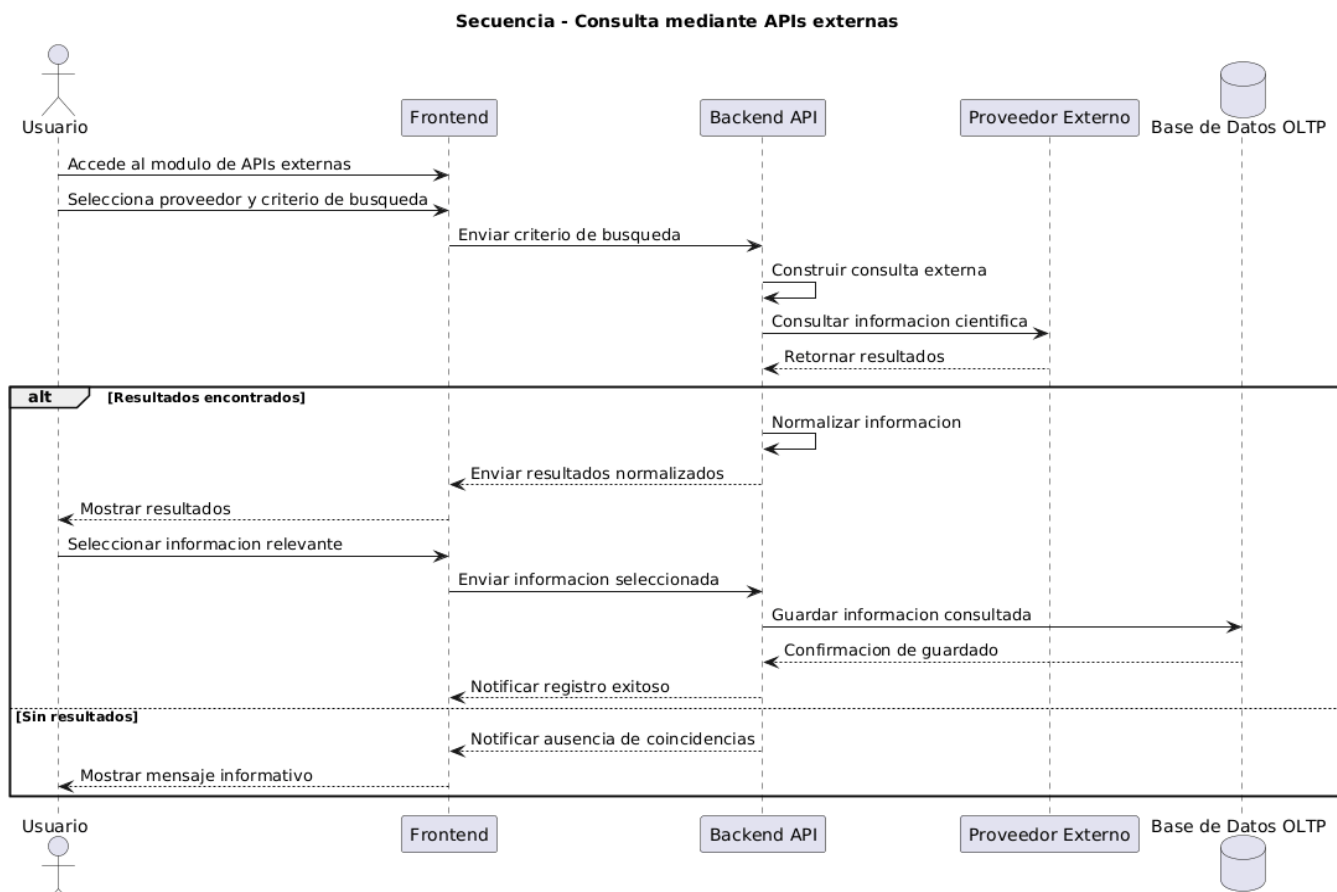


Figura 28. Diagrama de secuencia del proceso de consulta mediante APIs externas.

El diagrama muestra la comunicación entre el sistema institucional y los servicios académicos externos durante la consulta de información científica. Este flujo permite enviar criterios de búsqueda, recibir resultados, normalizar metadatos y almacenar la información seleccionada dentro del modelo transaccional.

Diagrama de secuencia del workflow institucional

El proceso de workflow institucional permite gestionar la revisión, aprobación o devolución de artículos y lotes de información registrados en el sistema. Como se observa en la Figura 29, el flujo integra la interacción entre el usuario solicitante, el sistema, el revisor institucional, los servicios internos y la base de datos OLTP.

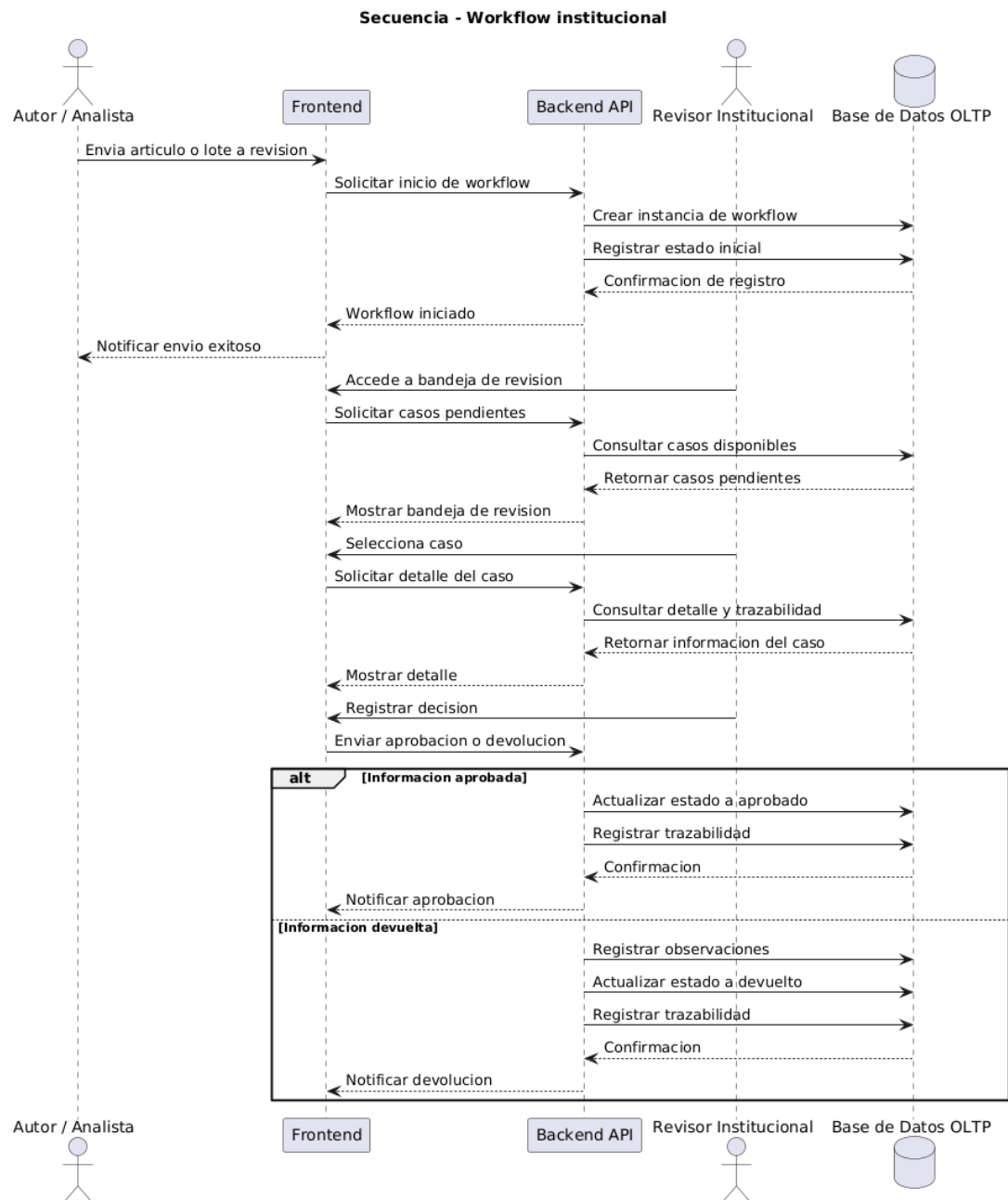


Figura 29. Diagrama de secuencia del workflow institucional.

El diagrama representa la comunicación entre los actores y servicios que intervienen en la revisión institucional de registros científicos. Esta interacción permite crear instancias de revisión, consultar casos pendientes, registrar decisiones y actualizar estados con su respectiva trazabilidad.

Diagrama de secuencia del proceso de generación de dashboards y reportes

El proceso de generación de dashboards y reportes permite consultar información analítica desde el modelo OLAP institucional para visualizar indicadores, gráficos y reportes dinámicos. Como se observa en la Figura 30, el flujo integra la interacción entre el usuario, el frontend, el backend, el proceso ETL y el Data Warehouse institucional

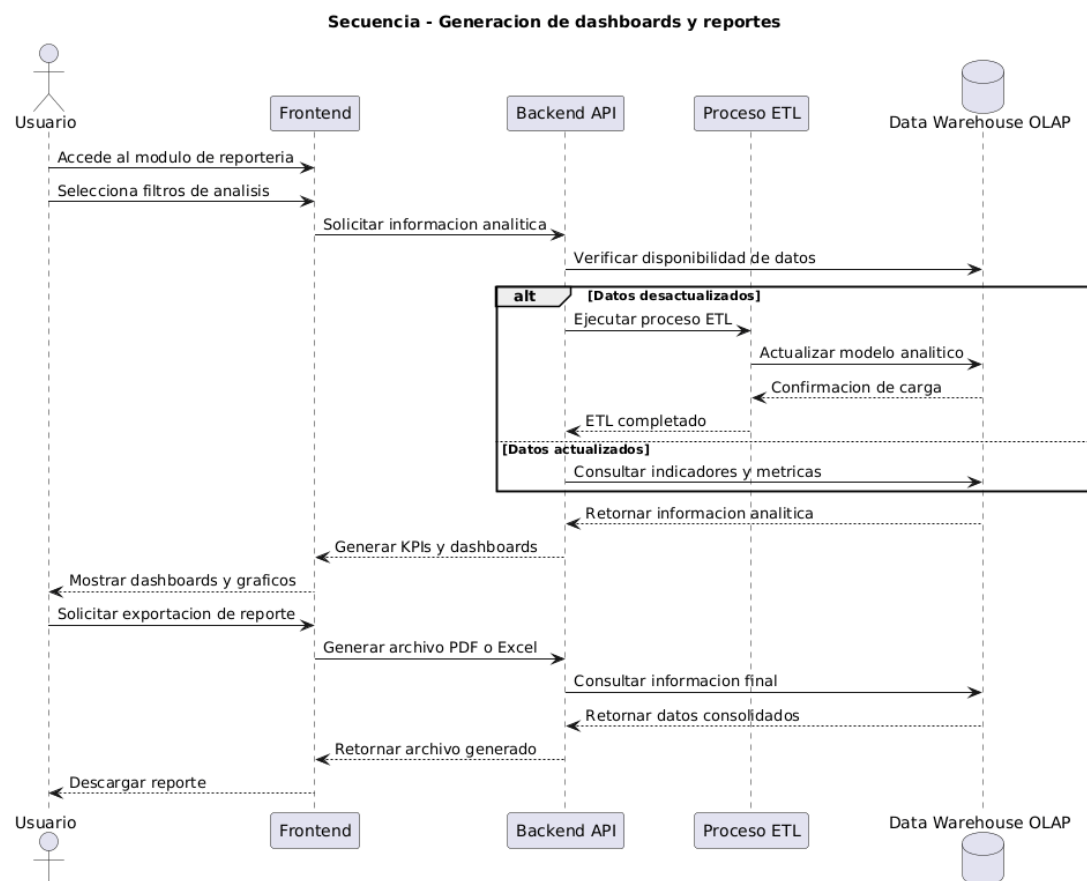


Figura 30. Diagrama de secuencia del proceso de generación de dashboards y reportes.

El diagrama muestra la comunicación entre los componentes analíticos del sistema durante la generación de dashboards y reportes. En este flujo, el backend consulta el

estado del Data Warehouse, ejecuta procesos ETL cuando es necesario y entrega información consolidada al frontend para su visualización o exportación.

Diagrama de secuencia del proceso de recomendación y predicción mediante IA

El proceso de recomendación y predicción mediante Inteligencia Artificial permite consultar información histórica desde el modelo analítico y ejecutar modelos orientados al apoyo de decisiones institucionales. Como se observa en la Figura 31, el flujo integra la interacción entre el usuario, la interfaz, el backend, el Data Warehouse y el componente de IA.

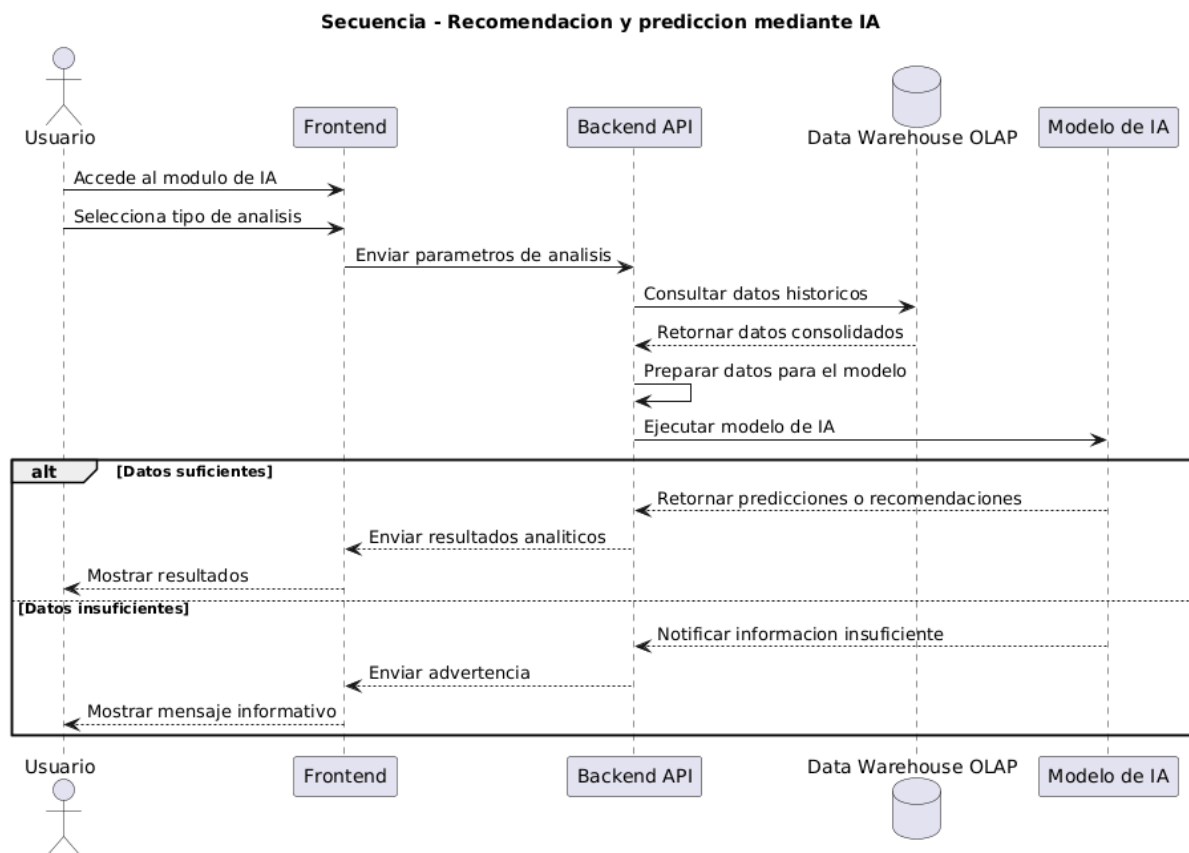


Figura 31. Diagrama de secuencia del proceso de recomendación y predicción mediante Inteligencia Artificial.

El diagrama representa la interacción entre los componentes necesarios para generar recomendaciones o predicciones institucionales. Este flujo contempla la consulta de datos históricos, preparación de variables, ejecución del modelo de IA y presentación de resultados analíticos al usuario.

3.2 Diseño de la propuesta

El diseño de la propuesta tecnológica se desarrolló a partir de los requerimientos funcionales y no funcionales identificados durante la etapa de análisis, considerando las necesidades institucionales relacionadas con la gestión, validación, análisis y reportería de información científica dentro de la DIDE-UTA. En esta fase se definió la arquitectura general del sistema, la organización modular de sus componentes y la integración entre los procesos operativos, analíticos y administrativos de la plataforma.

La propuesta fue estructurada bajo principios de modularidad, separación de responsabilidades y extensibilidad, con el propósito de facilitar el mantenimiento, crecimiento y adaptación del sistema a nuevos requerimientos institucionales. Asimismo, la arquitectura implementada y las herramientas de desarrollo seleccionadas se alinearon con los lineamientos tecnológicos definidos por la Dirección de Tecnologías de Información y Comunicación (DITIC) de la Universidad Técnica de Ambato. Esta decisión permitió mantener compatibilidad con la infraestructura tecnológica existente y favorecer futuros procesos de actualización o ampliación por parte del personal técnico institucional.

De esta manera, el diseño planteado integra módulos relacionados con registro de artículos científicos, carga masiva, workflow institucional, reportería analítica, integración con APIs externas, procesos ETL, Business Intelligence y funcionalidades orientadas a Inteligencia Artificial. Esta organización permite que la plataforma responda a las necesidades actuales de la DIDE-UTA y, al mismo tiempo, conserve una base tecnológica preparada para el crecimiento progresivo del sistema.

3.2.1 Arquitectura general del sistema

La arquitectura general del sistema fue definida considerando los lineamientos tecnológicos y estándares de desarrollo utilizados por la Dirección de Tecnologías de Información y Comunicación (DITIC) de la Universidad Técnica de Ambato. Las herramientas seleccionadas y la arquitectura implementada corresponden a tecnologías compatibles con el entorno institucional, lo que permite facilitar el mantenimiento, escalabilidad y continuidad técnica de la solución dentro de la universidad.

Como se observa en la Figura 32, la propuesta tecnológica fue estructurada mediante una arquitectura multicapa basada en principios de Clean Architecture, integrando componentes de presentación, aplicación, dominio, persistencia e infraestructura. Esta organización permite separar responsabilidades dentro del sistema, evitando dependencias directas entre la interfaz, la lógica de negocio, las reglas de dominio y los mecanismos de acceso a datos o servicios externos.

La capa de presentación gestiona la interacción con los usuarios mediante interfaces web orientadas a registro científico, reportería, administración y visualización analítica. La capa de aplicación centraliza la lógica de negocio, validaciones, servicios internos y procesos relacionados con workflow, ETL y Business Intelligence. La capa de dominio contiene las entidades principales y reglas de negocio del sistema, mientras que las capas de persistencia e infraestructura administran el acceso a bases de datos, autenticación, almacenamiento, reportes e integración con APIs externas.

Esta arquitectura favorece la mantenibilidad del sistema y permite incorporar nuevos módulos sin afectar directamente los componentes principales de la plataforma. Además, facilita la integración progresiva de funcionalidades relacionadas con reportería avanzada, análisis institucional e Inteligencia Artificial, manteniendo coherencia con el ecosistema tecnológico definido por la DITIC.



Figura 32. Arquitectura general del sistema basada en Clean Architecture [48].

Con base en la arquitectura definida, las tecnologías seleccionadas orientadas al desarrollo web, gestión transaccional de datos, reportería institucional, integración externa y procesamiento analítico. La Tabla 10 presenta los principales componentes tecnológicos utilizados en la propuesta, considerando su función dentro del sistema y su relación con los procesos operativos y analíticos implementados.

Tabla 10. Tecnologías utilizadas en la propuesta tecnológica.

Componente	Tecnología utilizada	Descripción
Backend	.NET / C#	Desarrollo de servicios, APIs y lógica de negocio institucional
Frontend	Blazor WebAssembly	Desarrollo de interfaces web interactivas
Base de datos transaccional	SQL Server 2020	Gestión operacional de información científica
Modelo analítico	SQL Server Data Warehouse	Gestión analítica y reportería institucional

Arquitectura	Clean Architecture	Organización modular y separación de responsabilidades
ORM	Entity Framework Core	Acceso y manipulación de datos
APIs externas	REST / JSON	Integración con plataformas científicas externas
Reportería	QuestPDF / ScottPlot	Generación de reportes y gráficos analíticos
Control de versiones	Git / GitHub	Gestión y seguimiento del desarrollo
Inteligencia Artificial	ML.NET	Implementación de funciones de recomendación y predicción institucional

Flujo general de la arquitectura del sistema

El flujo general de la arquitectura del sistema describe la relación entre los componentes operativos, analíticos y de reportería que conforman la plataforma propuesta para la DIDE-UTA. Como se observa en la Figura 33, la información científica se registra inicialmente mediante el frontend y los servicios backend, para luego almacenarse dentro del modelo transaccional OLTP, encargado de soportar los procesos diarios de registro, validación, carga masiva, workflow y administración institucional.

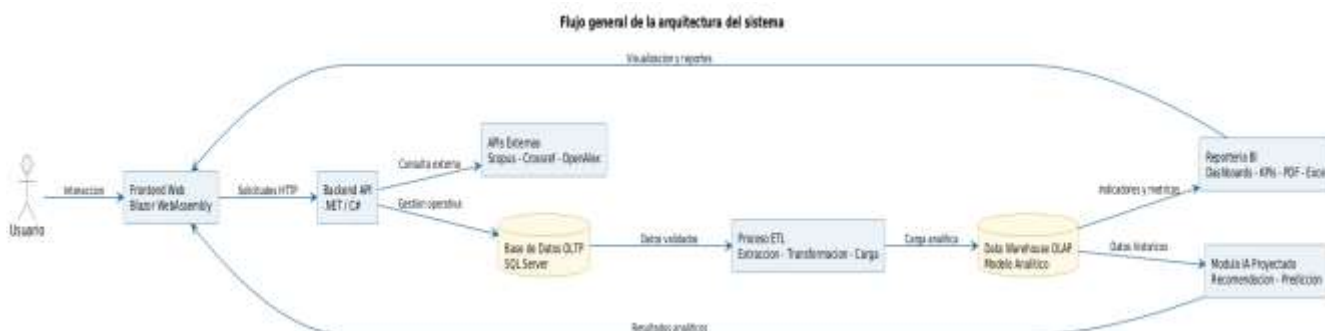


Figura 33. Flujo general de la arquitectura del sistema.

Posteriormente, la información operativa es procesada mediante mecanismos ETL, los cuales extraen, transforman y consolidan los datos hacia el modelo analítico OLAP. Este entorno analítico permite estructurar información histórica para la generación de dashboards, KPIs, reportes institucionales y consultas orientadas a Business Intelligence. La separación entre el modelo transaccional y el modelo analítico permite mantener estabilidad operativa durante los procesos de análisis, evitando que las consultas complejas afecten las operaciones diarias del sistema.

Además, la arquitectura contempla la integración con APIs externas para consultar y enriquecer información científica proveniente de fuentes académicas. De igual manera, el flujo general permite incorporar funcionalidades de Inteligencia Artificial orientadas a recomendación y predicción institucional, utilizando información consolidada del entorno analítico. En conjunto, esta arquitectura permite mantener una plataforma organizada, escalable y alineada con las necesidades operativas y analíticas de la DIDE-UTA.

3.2.2 Diseño del modelo transaccional OLTP

El modelo transaccional OLTP fue diseñado como la base operativa principal del sistema, debido a que soporta los procesos diarios de registro, validación, administración y seguimiento de información científica institucional dentro de la DIDE-UTA. Este modelo permite gestionar artículos científicos, autores, revistas, indexaciones, clasificaciones académicas, carga masiva, workflow institucional, trazabilidad y configuraciones dinámicas utilizadas por la plataforma.

Se lo define como un modelo OLTP extensible regido por metadatos porque no se limita al almacenamiento fijo de artículos científicos, sino que incorpora estructuras configurables para formularios, campos dinámicos, catálogos, matrices de registro y procesos de validación. Esta característica permite adaptar la plataforma a nuevos requerimientos institucionales sin depender completamente de cambios directos sobre la estructura principal de la base de datos.

Asimismo, el modelo fue construido para mantener integridad, trazabilidad y consistencia durante el ciclo de vida de la información, desde el ingreso inicial hasta su revisión, consolidación y posterior uso en procesos analíticos. De esta manera, el modelo OLTP constituye la fuente primaria de datos para los procesos ETL que alimentan el modelo analítico OLAP, la reportería institucional y las funcionalidades relacionadas con Business Intelligence e Inteligencia Artificial.

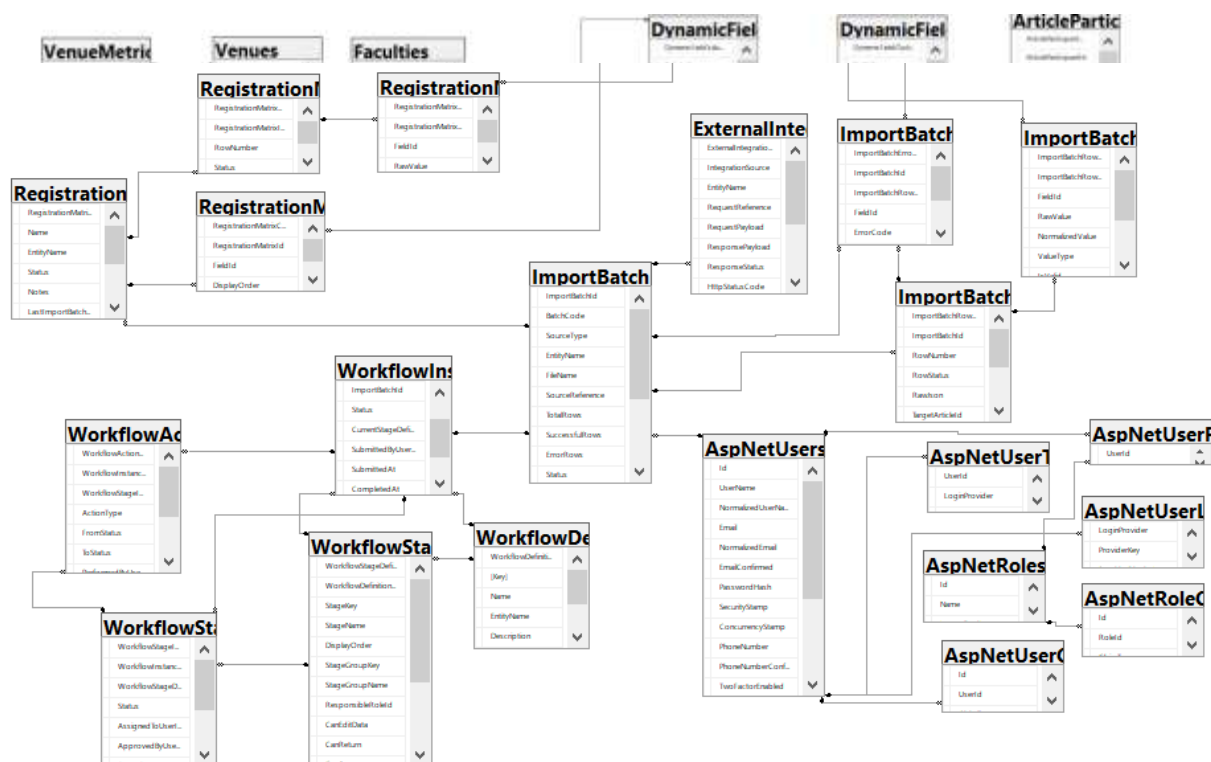


Figura 34. Esquema general del modelo transaccional OLTP.

El modelo entidad-relación fue estructurado para organizar los procesos operativos relacionados con la gestión de información científica institucional. Como se observa en la Figura 34, el diseño integra entidades asociadas a artículos, autores, revistas, métricas editoriales, clasificaciones académicas, formularios dinámicos, carga masiva, workflow y trazabilidad. Esta organización permite centralizar la información operativa, mantener relaciones consistentes entre entidades y preparar datos confiables para procesos posteriores de ETL, reportería y análisis institucional.

Tabla 11. Agrupación funcional de entidades del modelo transaccional OLTP

Grupo funcional	Entidades principales	Funcionalidad dentro del sistema
Gestión de artículos científicos	Article, ArticleParticipant, ArticleIndexing	Permiten registrar artículos científicos, autores, coautores y relaciones con indexaciones académicas institucionales.
Gestión editorial y métricas	Venue, VenueMetric	Permiten administrar revistas, métricas editoriales, cuartiles e información utilizada para reportería institucional.

Clasificación académica	BroadField, SpecificField, DetailedField, ResearchLine	Permiten clasificar artículos científicos según líneas de investigación y estructura OCDE institucional.
Gestión institucional	AcademicTerm, PublicationStatus	Permiten controlar periodos académicos y estados institucionales asociados a la producción científica.
Configuración dinámica y metadatos	DynamicForm, DynamicField, DynamicCatalog	Permiten gestionar formularios configurables, catálogos y estructuras dinámicas orientadas a la extensibilidad del sistema.
Matriz de registro	RegistrationMatrix, MatrixField	Permiten administrar estructuras matriciales para el ingreso y organización de información científica institucional.
Carga masiva y staging	ImportBatch, ImportDetail, StagingArticle	Permiten gestionar procesos de importación masiva, validación de lotes y control temporal de información antes de su consolidación institucional.
Workflow institucional	WorkflowInstance, WorkflowState, WorkflowTrace	Permiten administrar procesos de revisión, validación, observaciones y trazabilidad institucional sobre los registros científicos.
Integración externa	ExternalSource, ApiIntegrationLog	Permiten gestionar consultas y sincronización de información proveniente de APIs académicas externas.
Seguridad y administración	ApplicationUser, ApplicationRole, AuditLog	Permiten controlar autenticación, autorización, trazabilidad y administración institucional del sistema.
Procesos analíticos y BI	EtlExecution, ReportConfiguration	Permiten controlar procesos ETL, reportería y configuración de análisis institucional utilizados en Business Intelligence.

La agrupación funcional de entidades permite comprender la organización del modelo transaccional de acuerdo con los procesos operativos identificados durante el análisis institucional. Esta clasificación separa componentes relacionados con gestión científica, configuración dinámica, carga masiva, workflow, integración externa, seguridad y reportería. Además, facilita identificar qué entidades intervienen en los procesos diarios del sistema y cuáles sirven como fuente para los procesos ETL y el modelo analítico OLAP. De esta manera, el modelo mantiene una estructura modular, extensible y coherente con las necesidades operativas de la DIDE-UTA.

Justificación del modelo transaccional OLTP extensible regido por metadatos

El modelo transaccional implementado fue diseñado bajo un enfoque OLTP, debido a que su función principal consiste en gestionar operaciones institucionales frecuentes relacionadas con el registro, validación, administración y seguimiento de información científica. A diferencia de un modelo transaccional convencional orientado únicamente al almacenamiento de registros estáticos, la estructura desarrollada incorpora mecanismos de extensibilidad y configuración dinámica que permiten adaptar la plataforma conforme evolucionen las necesidades de la DIDE-UTA. Este enfoque favorece la estabilidad operativa del sistema y permite mantener una base organizada para los procesos institucionales diarios.

La extensibilidad del modelo se sustenta en una arquitectura basada en metadatos, mediante la cual se administran formularios dinámicos, campos configurables, catálogos institucionales y matrices de registro. Estos componentes permiten incorporar nuevos elementos funcionales sin modificar de manera permanente la estructura principal de la base de datos. Además, la administración basada en metadatos facilita la reutilización de componentes dentro de diferentes módulos, reduciendo dependencias estructurales y favoreciendo la mantenibilidad de la plataforma.

El modelo también incorpora mecanismos de staging, workflow y trazabilidad para controlar la calidad de la información antes de su consolidación definitiva. Las estructuras de staging permiten validar datos provenientes de registros manuales, cargas masivas e integraciones externas, mientras que el workflow institucional gestiona procesos de revisión, aprobación, devolución y seguimiento. Por su parte, la trazabilidad permite registrar acciones, cambios de estado y modificaciones realizadas durante el ciclo de vida de cada registro. En conjunto, estos mecanismos fortalecen la integridad de los datos y preparan información confiable para los procesos ETL, reportería institucional y análisis mediante Business Intelligence.

La arquitectura funcional del modelo OLTP fue organizada mediante componentes modulares que permiten separar responsabilidades dentro del sistema transaccional. Como se observa en la Figura 35, el modelo integra módulos relacionados con seguridad, registro de artículos, configuración dinámica, carga masiva, workflow, integración externa, trazabilidad y procesos de apoyo analítico. Esta organización facilita la evolución progresiva del sistema y mantiene una estructura operativa preparada para integrarse con los procesos ETL y el entorno de reportería institucional.

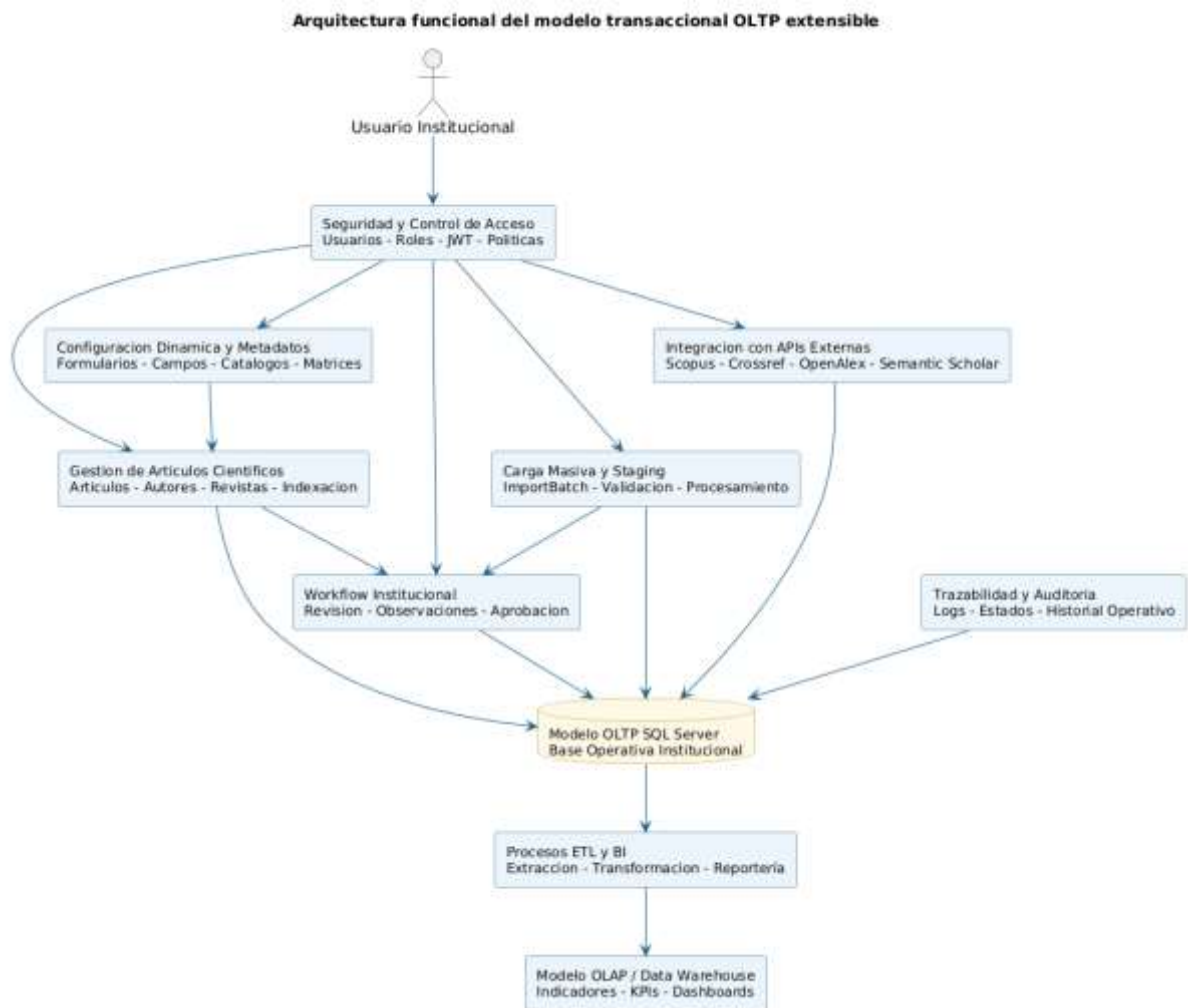


Figura 35. Arquitectura funcional del modelo transaccional OLTP extensible.

Diseño basado en metadatos y componentes dinámicos

Uno de los componentes principales del modelo transaccional corresponde a la estructura basada en metadatos, diseñada para proporcionar flexibilidad a los procesos institucionales gestionados por la plataforma. Este enfoque permite administrar

formularios, campos, catálogos y configuraciones funcionales desde estructuras parametrizables, evitando depender completamente de modificaciones directas sobre la base de datos para incorporar nuevos requerimientos. Como se observa en la Figura 36, la administración basada en metadatos centraliza configuraciones utilizadas por los módulos operativos y facilita la adaptación del sistema a diferentes procesos de registro, validación y análisis.

La estructura dinámica permite configurar formularios, campos y matrices de registro dentro del modelo OLTP, para que posteriormente puedan ser interpretados por los componentes visuales del sistema. De esta manera, el motor de renderizado dinámico puede construir interfaces adaptadas a las configuraciones definidas, favoreciendo la reutilización de elementos y reduciendo dependencias rígidas entre la interfaz y la estructura física de la base de datos. Este enfoque contribuye a mantener una plataforma escalable, adaptable y alineada con las necesidades evolutivas de la DIDE-UTA.

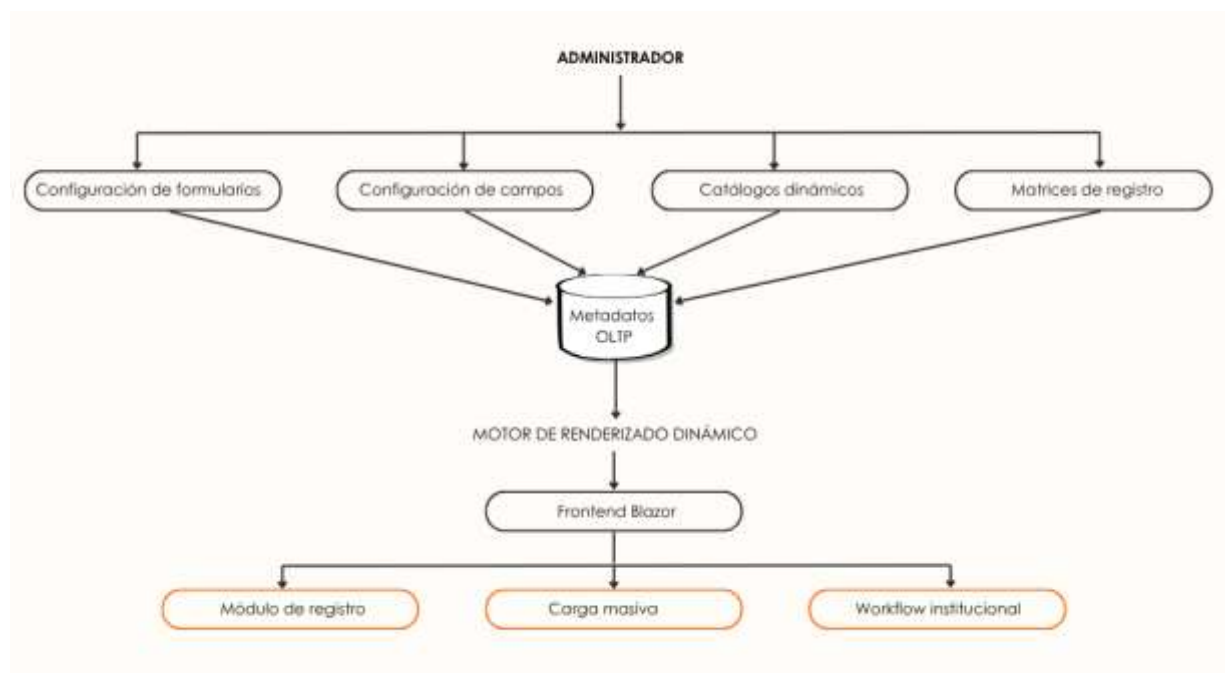


Figura 36. Estructura funcional del modelo basado en metadatos y componentes dinámicos.

Procesos de staging, workflow y trazabilidad institucional

El modelo transaccional incorpora mecanismos de staging, workflow y trazabilidad para controlar el ciclo de vida de la información científica gestionada por la

plataforma. Las estructuras de staging permiten almacenar y validar temporalmente información proveniente de registros manuales, cargas masivas e integraciones externas antes de su consolidación definitiva en el modelo OLTP. Como se observa en la Figura 37, estos componentes actúan como puntos de control para reducir inconsistencias y mejorar la calidad de los datos registrados.

El workflow institucional permite administrar revisiones, observaciones, aprobaciones y devoluciones relacionadas con artículos científicos o lotes de información. A su vez, los mecanismos de trazabilidad registran acciones, cambios de estado y actividades realizadas por los diferentes actores del sistema. Esta integración fortalece la auditoría, seguimiento e integridad de los procesos institucionales, permitiendo disponer de información confiable para reportería y análisis posterior.

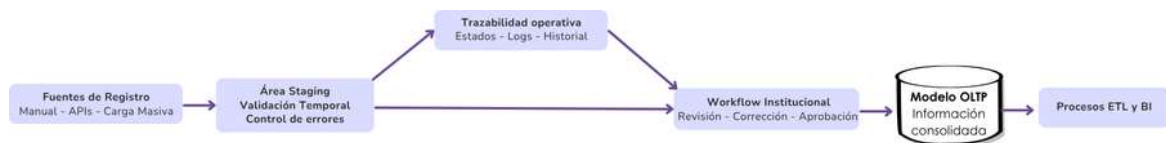


Figura 37. Arquitectura funcional de staging, workflow y trazabilidad institucional.

Integración del modelo OLTP con procesos ETL y Business Intelligence

El modelo transaccional OLTP fue diseñado para integrarse con los procesos ETL y el modelo analítico institucional utilizado en reportería y Business Intelligence. Esta integración permite extraer información operativa proveniente de registros científicos, workflow, carga masiva, integraciones externas y procesos administrativos, para posteriormente transformarla y consolidarla dentro del entorno OLAP. Como se observa en la Figura 38, el flujo entre OLTP, ETL y BI permite separar las operaciones diarias del sistema de los procesos analíticos utilizados para dashboards, KPIs y reportes institucionales.

Los procesos ETL realizan tareas de validación, transformación, limpieza y carga de información hacia el Data Warehouse institucional. Esta separación evita que las consultas analíticas afecten el rendimiento del modelo transaccional y permite mantener información histórica organizada para el análisis institucional. Además, la

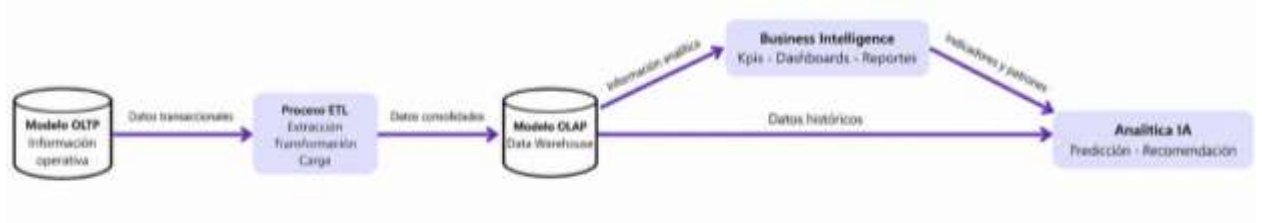


Figura 38. Integración del modelo OLTP con procesos ETL y Business Intelligence.

información consolidada en el entorno analítico constituye la base para funcionalidades de Inteligencia Artificial, recomendación y análisis predictivo.

Estructura relacional y organización de entidades del modelo OLTP

La estructura relacional del modelo transaccional fue diseñada para organizar de manera modular los procesos operativos relacionados con la gestión de información científica institucional. Para ello, se definieron entidades especializadas que permiten administrar artículos, autores, revistas, indexaciones, clasificaciones académicas, formularios dinámicos, carga masiva, workflow, trazabilidad y seguridad. Como se observa en la Figura 39, estas entidades se relacionan mediante una organización funcional orientada a mantener integridad de datos, separación de responsabilidades y escalabilidad operativa.

La entidad Article constituye uno de los componentes principales del modelo, debido a que centraliza información relacionada con la producción científica institucional. A partir de esta entidad se establecen relaciones con autores, revistas, fuentes de indexación, estados de publicación y clasificaciones académicas. De igual manera, las entidades relacionadas con formularios dinámicos, staging, workflow y trazabilidad complementan el proceso operativo, permitiendo administrar información configurable, validar registros y controlar el seguimiento institucional. Esta estructura proporciona una base ordenada para el funcionamiento del sistema y para su posterior integración con procesos ETL y modelos analíticos.

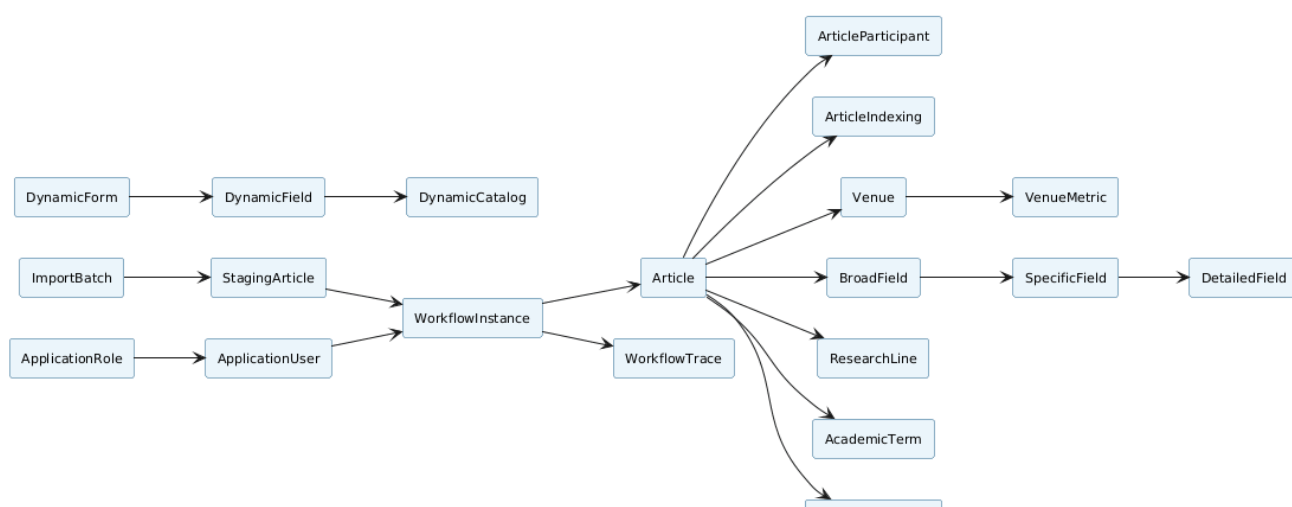


Figura 39. Relación funcional de entidades principales del modelo transaccional OLTP.

Tabla 12. Entidades principales y propósito funcional dentro del modelo OLTP.

Entidad	Propósito funcional
Article	Gestiona el registro principal de artículos científicos institucionales.
ArticleParticipant	Administra autores, coautores y participantes relacionados con cada artículo.
ArticleIndexing	Gestiona la relación entre artículos y fuentes de indexación científica.
Venue	Almacena información editorial de revistas científicas y medios de publicación.
VenueMetric	Gestiona métricas editoriales relacionadas con cuartiles, SJR e indicadores científicos.
BroadField	Administra áreas generales de clasificación académica.
SpecificField	Gestiona subclasificaciones académicas relacionadas con áreas específicas.
DetailedField	Permite organizar clasificaciones detalladas utilizadas en investigación científica.
ResearchLine	Gestiona líneas de investigación institucionales asociadas a los artículos científicos.

AcademicTerm	Controla periodos académicos institucionales utilizados en el sistema.
PublicationStatus	Gestiona estados editoriales y administrativos asociados a los artículos científicos.
DynamicForm	Permite administrar formularios dinámicos configurables dentro de la plataforma.
DynamicField	Gestiona campos dinámicos utilizados por formularios institucionales.
DynamicCatalog	Permite administrar catálogos reutilizables utilizados por estructuras dinámicas.
ImportBatch	Gestiona procesos de carga masiva y agrupación de registros importados.
StagingArticle	Almacena información temporal utilizada durante validaciones y procesamiento previo.
WorkflowInstance	Gestiona procesos de revisión y validación institucional de registros científicos.
WorkflowTrace	Registra trazabilidad y seguimiento de acciones realizadas durante el workflow.
ApplicationUser	Administra usuarios institucionales autenticados dentro de la plataforma.
ApplicationRole	Gestiona roles y permisos utilizados para control de acceso institucional.

La organización funcional de entidades permite comprender el papel que cumple cada componente dentro del modelo OLTP. Las entidades operativas administran información científica, editorial y académica, las entidades dinámicas permiten configurar formularios y campos reutilizables, mientras que las entidades de control gestionan carga masiva, workflow, trazabilidad y seguridad. Esta separación fortalece la integridad del modelo, mejora su mantenibilidad y facilita la incorporación progresiva de nuevas funcionalidades dentro de la plataforma institucional.

Tabla 13. Relaciones funcionales principales del modelo transaccional OLTP.

Relación	Tipo de relación	Descripción funcional
Article → ArticleParticipant	Uno a muchos	Un artículo puede registrar múltiples autores o participantes institucionales.
Article → ArticleIndexing	Uno a muchos	Un artículo puede estar asociado a varias fuentes de indexación científica.
Article → Venue	Muchos a uno	Varios artículos pueden pertenecer a una misma revista o medio científico.
Venue → VenueMetric	Uno a muchos	Una revista puede registrar diferentes métricas científicas por periodo.
BroadField → SpecificField	Uno a muchos	Una clasificación general contiene múltiples clasificaciones específicas.

SpecificField → DetailedField	Uno a muchos	Una clasificación específica agrupa varias clasificaciones detalladas.
DynamicForm → DynamicField	Uno a muchos	Un formulario dinámico puede contener múltiples campos configurables.
ImportBatch → StagingArticle	Uno a muchos	Un lote de carga masiva agrupa múltiples registros temporales.
WorkflowInstance → WorkflowTrace	Uno a muchos	Un proceso de workflow registra múltiples eventos y acciones de trazabilidad.
ApplicationRole → ApplicationUser	Uno a muchos	Un rol institucional puede estar asociado a múltiples usuarios del sistema.

Desde una perspectiva funcional, las relaciones del modelo OLTP permiten representar el ciclo operativo de la información científica institucional, desde el registro del artículo hasta su validación, clasificación, seguimiento y uso posterior en reportería. La estructura relacional favorece la integridad referencial, evita duplicidad innecesaria de datos y permite identificar las entidades fuente utilizadas por los procesos ETL. Además, la combinación de entidades operativas, dinámicas y de control proporciona una base transaccional preparada para soportar procesos administrativos, analíticos y funcionalidades de Inteligencia Artificial.

Estrategias de integridad, escalabilidad y normalización del modelo

El modelo transaccional OLTP fue diseñado considerando principios de integridad, escalabilidad y normalización, con el propósito de garantizar estabilidad operativa y adaptabilidad frente a futuros requerimientos institucionales. La estructura relacional implementa claves primarias, claves foráneas y relaciones controladas entre entidades, permitiendo mantener consistencia durante los procesos de registro, validación y administración de información científica. Asimismo, la separación funcional de entidades evita redundancia innecesaria de datos y mantiene una organización modular orientada a los procesos principales del sistema.

Como se observa en la Figura 40, las estrategias estructurales aplicadas permiten organizar el modelo mediante componentes especializados, relaciones controladas y mecanismos configurables basados en metadatos. Esta organización favorece la reutilización de componentes, la incorporación progresiva de nuevos módulos y la adaptación de configuraciones institucionales sin comprometer la estabilidad general

del sistema. Finalmente, estas estrategias proporcionan una base sólida para la integración con procesos ETL, reportería institucional y funcionalidades relacionadas con Inteligencia Artificial.

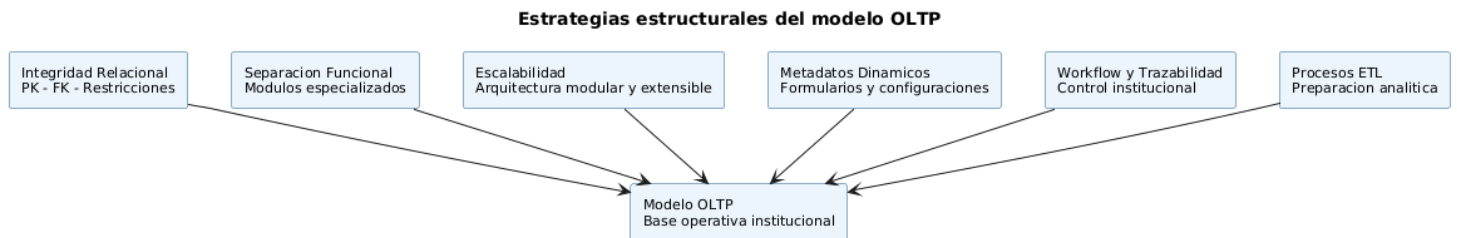


Figura 40. Estrategias estructurales aplicadas en el modelo transaccional OLTP.

3.2.3 Diseño del modelo analítico OLAP

El modelo analítico OLAP fue diseñado con el propósito de transformar la información operativa generada dentro del sistema transaccional en datos estructurados para análisis institucional, reportería y Business Intelligence. Debido a la existencia de varios procesos analíticos dentro de la plataforma, se organizó como una constelación dimensional compuesta por múltiples tablas de hechos y dimensiones compartidas. Esta estructura permite consolidar información histórica relacionada con producción científica, autores, revistas, indexaciones, líneas de investigación, métricas editoriales y procesos institucionales, facilitando la generación de indicadores y dashboards orientados a la toma de decisiones estratégicas dentro de la DIDE-UTA.

Asimismo, el diseño fue construido bajo un enfoque dimensional tipo Data Warehouse, permitiendo organizar la información mediante tablas de hechos especializadas y dimensiones optimizadas para consultas analíticas y generación de métricas institucionales. De igual manera, la separación entre el entorno transaccional OLTP y el entorno analítico OLAP permite mantener la estabilidad operativa del sistema durante la ejecución de procesos de reportería y análisis complejos. La arquitectura implementada incorpora procesos ETL encargados de extraer, transformar y consolidar información proveniente de los módulos operativos hacia la constelación dimensional. Finalmente, este modelo constituye la base analítica utilizada para

reportería institucional, generación de KPIs y funcionalidades relacionadas con Inteligencia Artificial y análisis predictivo.

Fundamentación metodológica del modelo dimensional

El diseño del modelo analítico institucional fue desarrollado bajo un enfoque dimensional basado en la metodología propuesta por Ralph Kimball, debido a que permite estructurar entornos analíticos orientados a reportería, generación de indicadores y Business Intelligence mediante modelos optimizados para consulta. Esta metodología se caracteriza por organizar la información en tablas de hechos y dimensiones, comúnmente representadas mediante esquemas estrella que simplifican las consultas analíticas y mejoran el rendimiento de los procesos de reportería. En el presente proyecto, dicho enfoque fue aplicado mediante una constelación dimensional, compuesta por varias tablas de hechos especializadas que comparten dimensiones comunes. Esta adaptación permitió representar distintos procesos institucionales, como publicaciones científicas, autoría, indexaciones, métricas editoriales, validaciones y workflow, dentro de un mismo entorno analítico organizado y escalable. Además, la constelación dimensional facilita la construcción progresiva del Data Warehouse mediante la incorporación de nuevas dimensiones, métricas y procesos analíticos conforme evolucionen las necesidades institucionales. De igual manera, este enfoque favorece la generación de dashboards, KPIs y consultas multidimensionales utilizadas para apoyar la toma de decisiones estratégicas dentro de la DIDE-UTA. Finalmente, el modelo dimensional implementado proporciona una base sólida para futuros procesos relacionados con minería de datos, Inteligencia Artificial y análisis predictivo institucional.

La decisión de utilizar una arquitectura dimensional tipo constelación se encuentra relacionada con las necesidades institucionales identificadas durante el análisis de requerimientos y reportería científica dentro de la DIDE-UTA. A diferencia de arquitecturas orientadas al enfoque corporativo centralizado propuesto por Bill Inmon, donde los datos suelen organizarse inicialmente bajo modelos normalizados de mayor complejidad, la metodología Kimball permite desarrollar estructuras dimensionales flexibles y orientadas directamente al análisis institucional. En este caso, la constelación dimensional facilita representar varios procesos analíticos dentro de un mismo Data Warehouse, tales como publicaciones científicas, autoría, indexaciones,

workflow, validaciones y métricas editoriales. Esta estructura permite reutilizar dimensiones compartidas, optimizar tiempos de consulta y simplificar la generación de indicadores utilizados dentro de dashboards institucionales y procesos de Business Intelligence. Finalmente, la metodología seleccionada permitió construir un entorno analítico adaptable y alineado con las necesidades tecnológicas de la universidad.

El modelo dimensional implementado fue estructurado mediante tablas de hechos encargadas de almacenar métricas institucionales y tablas dimensionales utilizadas para proporcionar contexto analítico a los procesos de reportería y análisis institucional. Las tablas de hechos permiten consolidar métricas relacionadas con publicaciones científicas, indexaciones, participación de autores y métricas editoriales, mientras que las dimensiones permiten analizar información desde diferentes perspectivas temporales, académicas y científicas. Asimismo, la utilización de dimensiones relacionadas con tiempo, clasificación académica, revistas, líneas de investigación y estados editoriales facilita la generación de análisis multidimensionales utilizados para Business Intelligence institucional. De igual manera, esta arquitectura permite optimizar procesos ETL y mantener estructuras organizadas para la consolidación histórica de información proveniente del modelo transaccional OLTP. La separación entre dimensiones y métricas también favorece la mantenibilidad y crecimiento progresivo del Data Warehouse institucional. Finalmente, el modelo implementado proporciona una arquitectura analítica preparada para soportar reportería avanzada y funcionalidades relacionadas con Inteligencia Artificial institucional.

Arquitectura general del modelo analítico OLAP

La arquitectura del modelo analítico fue estructurada para permitir el procesamiento y consolidación de información proveniente del modelo transaccional institucional mediante procesos ETL orientados al análisis y reportería. Esta organización permite separar las operaciones transaccionales de los procesos analíticos, garantizando estabilidad operativa durante la generación de consultas, indicadores y dashboards institucionales. Asimismo, la arquitectura implementada facilita la integración entre los procesos ETL, el Data Warehouse institucional y los módulos de Business Intelligence utilizados para reportería analítica. De igual manera, el modelo permite almacenar información histórica consolidada para posteriores procesos de análisis,

minería de datos y funcionalidades relacionadas con Inteligencia Artificial. Como se observa en la Figura 41, la arquitectura analítica integra componentes operativos, procesos ETL, almacenamiento dimensional y herramientas de visualización institucional. Finalmente, esta estructura permite mantener una plataforma escalable y preparada para el crecimiento progresivo de los procesos analíticos institucionales.

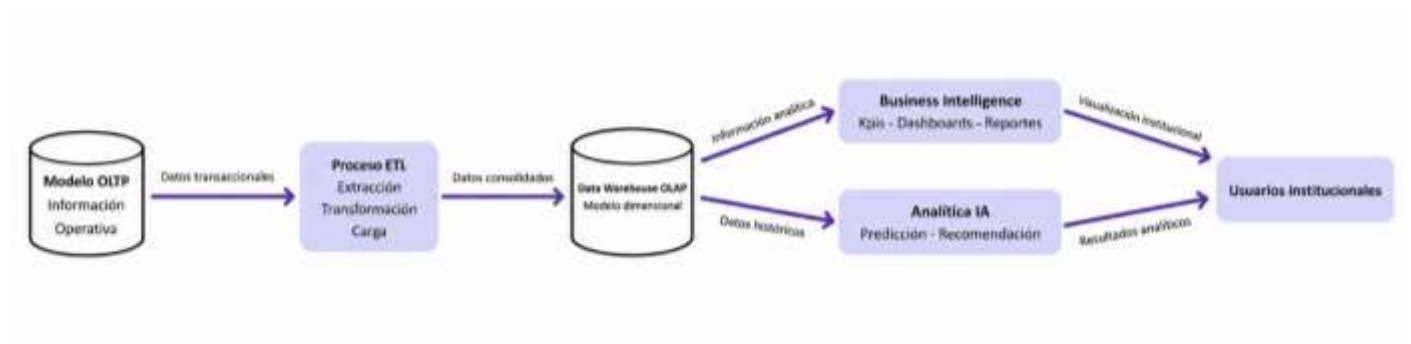


Figura 41. Arquitectura general del modelo analítico OLAP institucional.

La arquitectura general del modelo analítico permite consolidar información histórica proveniente de los diferentes módulos operativos gestionados por el sistema institucional. A través de esta estructura, los procesos ETL extraen y transforman datos provenientes del modelo transaccional OLTP para posteriormente almacenarlos dentro del Data Warehouse institucional. Asimismo, el modelo dimensional implementado permite optimizar consultas analíticas utilizadas para reportería, generación de KPIs y construcción de dashboards institucionales. De igual manera, la arquitectura desarrollada prepara información organizada para funcionalidades relacionadas con análisis predictivo e Inteligencia Artificial institucional. La separación entre los procesos operativos y analíticos también favorece el rendimiento y estabilidad general de la plataforma durante la ejecución de procesos complejos de reportería. Finalmente, esta organización permite mantener una arquitectura escalable y alineada con las necesidades analíticas de la DIDE-UTA.

Arquitectura dimensional tipo constelación del Data Warehouse

El modelo analítico institucional fue estructurado bajo una arquitectura dimensional tipo constelación de hechos, debido a que el Data Warehouse integra varios procesos analíticos relacionados con producción científica, autoría, indexación, métricas editoriales, carga masiva, validaciones y workflow institucional. Esta arquitectura está compuesta por múltiples tablas de hechos especializadas que comparten dimensiones

comunes, lo que permite analizar la información desde diferentes perspectivas institucionales. A diferencia de un esquema estrella simple, la constelación de hechos permite representar varios procesos de negocio dentro de un mismo entorno analítico, manteniendo orden, escalabilidad y reutilización de dimensiones. Esta estructura facilita la consolidación de información histórica proveniente del modelo transaccional OLTP y permite generar consultas analíticas de alto rendimiento para reportería, dashboards e indicadores institucionales. Como se observa en la Figura 42, las tablas de hechos se relacionan con dimensiones analíticas compartidas, permitiendo una visión integrada de los procesos gestionados por la DIDE-UTA. Finalmente, esta arquitectura proporciona una base organizada y extensible para Business Intelligence y funcionalidades relacionadas con Inteligencia Artificial.

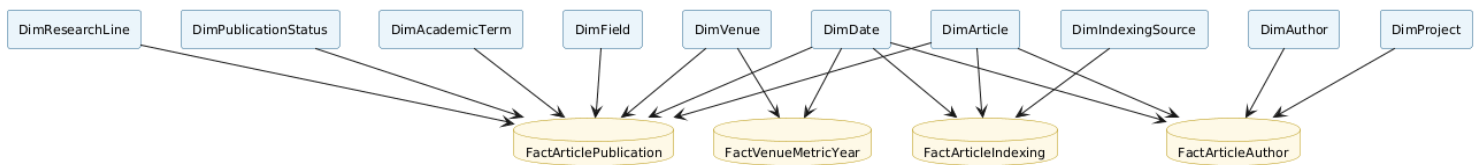


Figura 42. Modelo conceptual tipo constelación del Data Warehouse institucional.

La constelación dimensional implementada permite organizar la información analítica institucional mediante varias tablas de hechos asociadas a procesos específicos del sistema. Cada tabla de hechos concentra métricas de un área determinada, como publicaciones científicas, indexaciones, participación de autores, métricas editoriales, carga masiva, validaciones y workflow institucional. A su vez, estas tablas comparten dimensiones comunes como tiempo, artículo, autor, revista, clasificación académica, línea de investigación, estado de publicación y fuente de indexación. Esta organización permite realizar análisis multidimensionales sin duplicar estructuras innecesarias y facilita la generación de dashboards, reportes e indicadores estratégicos. Además, el modelo puede crecer progresivamente mediante la incorporación de nuevas tablas de hechos, dimensiones o vistas analíticas, sin rediseñar por completo el Data Warehouse institucional.

Presentación del modelo OLAP implementado

Con base en la arquitectura dimensional definida, se implementó un modelo OLAP orientado a consolidar información histórica relacionada con producción científica, métricas institucionales y procesos de reportería analítica. El modelo organiza la información mediante tablas de hechos y dimensiones especializadas que permiten optimizar consultas, filtros, agrupaciones e indicadores utilizados en Business Intelligence. La estructura implementada facilita la integración de datos provenientes del modelo transaccional OLTP mediante procesos ETL encargados de transformar, depurar y consolidar la información institucional. Además, mantiene independencia entre los procesos operativos y analíticos, evitando que las consultas de reportería afecten el funcionamiento transaccional del sistema. Como se observa en el Anexo 14, el modelo OLAP implementado corresponde a una constelación dimensional donde diferentes hechos comparten dimensiones académicas, temporales, editoriales y científicas. Esta organización permite sostener reportería avanzada y preparar información estructurada para funcionalidades de Inteligencia Artificial.

El modelo dimensional implementado fue estructurado mediante tablas de hechos encargadas de almacenar métricas institucionales y tablas dimensionales utilizadas para proporcionar contexto analítico dentro del Data Warehouse. Las dimensiones relacionadas con tiempo, autores, revistas, clasificación académica, indexaciones, estados de publicación y líneas de investigación permiten analizar la información desde diferentes perspectivas. Por su parte, las tablas de hechos consolidan métricas específicas sobre publicaciones científicas, indexaciones, métricas editoriales, participación de autores y procesos institucionales. Esta arquitectura mejora el rendimiento de las consultas analíticas utilizadas en dashboards, KPIs y reportes institucionales. Además, facilita la incorporación de nuevas métricas y dimensiones conforme evolucionen las necesidades de análisis de la DIDE-UTA. En conjunto, el modelo OLAP constituye la base analítica utilizada para Business Intelligence y funciones de análisis predictivo.

La Tabla 14 presenta las principales tablas de hechos y dimensiones que conforman la constelación dimensional del modelo OLAP institucional. Las tablas de hechos representan procesos analíticos específicos, mientras que las dimensiones proporcionan el contexto necesario para realizar consultas, filtros, agrupaciones y

análisis multidimensionales. Esta estructura permite que diferentes hechos compartan dimensiones comunes, característica propia de una arquitectura dimensional tipo constelación.

Tabla 14. Tablas de hechos y dimensiones de la constelación dimensional OLAP institucional.

Tipo	Entidad	Funcionalidad analítica
Tabla de hechos	FactArticlePublication	Almacena métricas relacionadas con publicaciones científicas institucionales.
Tabla de hechos	FactArticleIndexing	Gestiona indicadores relacionados con indexaciones científicas y bases académicas.
Tabla de hechos	FactVenueMetricYear	Almacena métricas editoriales y científicas por revista y periodo.
Tabla de hechos	FactArticleAuthor	Gestiona métricas relacionadas con participación de autores y producción científica.
Dimensión	DimArticle	Proporciona contexto analítico relacionado con artículos científicos.
Dimensión	DimAuthor	Permite analizar información relacionada con autores institucionales.
Dimensión	DimVenue	Gestiona información analítica de revistas y medios científicos.
Dimensión	DimDate	Permite realizar análisis temporales y generación de tendencias históricas.
Dimensión	DimAcademicTerm	Gestiona análisis relacionados con periodos académicos institucionales.
Dimensión	DimField	Proporciona clasificación académica utilizada para reportería institucional.
Dimensión	DimResearchLine	Gestiona análisis por líneas de investigación institucionales.
Dimensión	DimPublicationStatus	Permite analizar estados editoriales y administrativos de publicaciones científicas.
Dimensión	DimIndexingSource	Gestiona información analítica relacionada con bases de indexación científica.
Dimensión	DimProject	Permite relacionar producción científica con proyectos institucionales.

La organización dimensional implementada dentro del modelo OLAP permite estructurar la información institucional mediante componentes analíticos

especializados orientados a reportería y Business Intelligence. La utilización de múltiples tablas de hechos y dimensiones compartidas facilita el análisis multidimensional de producción científica, indexaciones, líneas de investigación, métricas editoriales y procesos institucionales. Esta arquitectura favorece la generación de indicadores históricos y dashboards analíticos utilizados para el seguimiento y evaluación de la gestión científica dentro de la DIDE-UTA. Además, la separación entre hechos y dimensiones permite mantener consultas optimizadas y escalabilidad analítica frente al crecimiento progresivo de la información institucional. La estructura también prepara datos consolidados para futuros procesos relacionados con Inteligencia Artificial y análisis predictivo. Finalmente, la constelación dimensional desarrollada constituye la base principal para la reportería analítica y el entorno BI institucional.

Proceso ETL institucional

El proceso ETL fue diseñado para extraer, transformar y consolidar la información proveniente del modelo transaccional OLTP hacia la constelación dimensional del entorno OLAP institucional. Esta arquitectura permite convertir los datos operativos generados diariamente dentro de la plataforma en información estructurada para reportería, indicadores y procesos de Business Intelligence. El flujo ETL implementado considera procesos de validación, depuración, homologación y trazabilidad sobre la información proveniente de registros manuales, cargas masivas, integraciones externas y procesos de workflow. Además, permite consolidar datos relacionados con artículos científicos, autores, revistas, indexaciones, métricas editoriales y estructuras dinámicas basadas en metadatos. Como se observa en la Figura 43, el flujo ETL establece una conexión organizada entre la extracción



Figura 43. Arquitectura general del proceso ETL institucional.

operacional, la transformación de datos y la carga hacia las tablas de hechos y dimensiones compartidas del Data Warehouse. Finalmente, esta estructura garantiza consistencia, integridad y preparación analítica de los datos utilizados posteriormente para reportería institucional, Business Intelligence e Inteligencia Artificial. La arquitectura ETL implementada permite transformar información operativa proveniente del modelo transaccional en datos estructurados para análisis institucional y reportería analítica. La etapa de extracción obtiene información relacionada con artículos científicos, autores, indexaciones, workflow institucional, carga masiva y estructuras dinámicas configuradas mediante metadatos. Posteriormente, la fase de transformación realiza procesos de validación, homologación, depuración y normalización de datos antes de su consolidación dentro del entorno analítico institucional. La etapa de carga permite poblar dimensiones compartidas y tablas de hechos especializadas que forman parte de la constelación dimensional. De esta manera, el flujo ETL mantiene consistencia histórica y calidad analítica de los datos utilizados dentro del entorno BI institucional. Además, la arquitectura desarrollada prepara información estructurada para futuros procesos relacionados con minería de datos e Inteligencia Artificial. Finalmente, el proceso ETL constituye uno de los componentes fundamentales dentro de la arquitectura analítica propuesta para la DIDE-UTA.

Arquitectura operativa del proceso ETL

El proceso ETL institucional fue desarrollado mediante una arquitectura híbrida compuesta por procedimientos almacenados, procesos internos implementados en .NET y mecanismos de consolidación analítica utilizados para poblar la constelación dimensional del Data Warehouse institucional. Esta arquitectura permite ejecutar procesos de extracción, transformación y carga tanto desde procedimientos SQL Server como desde servicios implementados dentro del backend del sistema. El entorno ETL incorpora mecanismos de validación y control sobre la información proveniente de módulos operativos, staging institucional e integraciones externas. Además, el flujo implementado mantiene separación entre la lógica operacional del sistema OLTP y la consolidación analítica utilizada posteriormente por el entorno BI institucional. Como se observa en la Figura 44, la arquitectura ETL integra componentes de validación, staging, procedimientos de transformación y

consolidación dimensional dentro del Data Warehouse. Finalmente, esta estructura permite mantener procesos analíticos organizados, escalables y preparados para futuras ampliaciones institucionales.

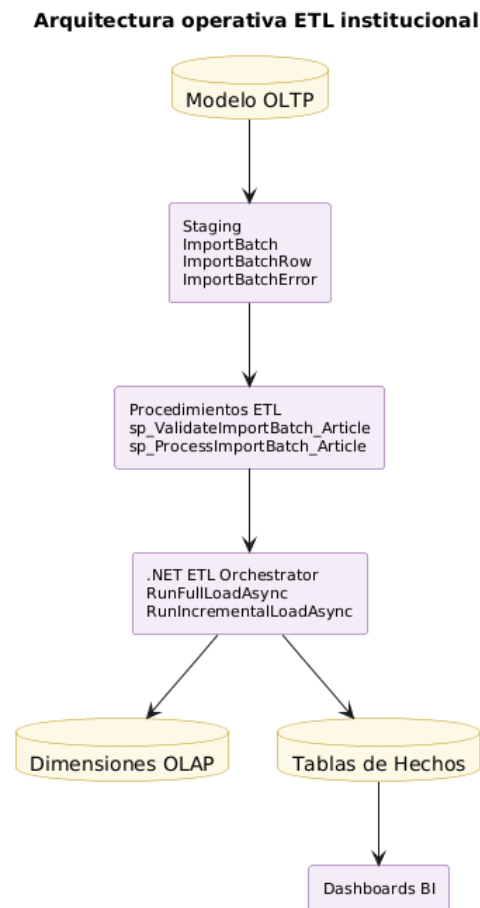


Figura 44. Arquitectura operativa del proceso ETL institucional.

La arquitectura operativa implementada permite controlar el flujo de información desde los procesos operativos institucionales hasta la consolidación analítica utilizada por el entorno BI. Las estructuras de staging permiten almacenar temporalmente información proveniente de registros manuales, cargas masivas e integraciones externas antes de su procesamiento definitivo. Los procedimientos almacenados realizan procesos de validación y control sobre la información registrada dentro del sistema institucional. Además, el orquestador ETL desarrollado en .NET permite ejecutar procesos de carga completa e incremental hacia las dimensiones compartidas y tablas de hechos de la constelación OLAP. Esta arquitectura facilita poblar los

componentes analíticos utilizados posteriormente para reportería, dashboards y generación de indicadores institucionales. La separación entre staging, validación y consolidación analítica fortalece la integridad y consistencia de la información utilizada dentro del entorno BI. Finalmente, esta estructura proporciona una arquitectura ETL organizada y preparada para el crecimiento progresivo del entorno analítico institucional.

Tabla 15. Procedimientos y componentes utilizados dentro del proceso ETL institucional.

Componente ETL	Funcionalidad
sp_ValidateImportBatch_Article	Valida registros importados y detecta errores de estructura y obligatoriedad.
sp_ProcessImportBatch_Article	Procesa lotes de artículos científicos desde staging hacia estructuras operativas.
sp_ProcessImportBatchRow_Article	Procesa individualmente registros importados relacionados con artículos científicos.
sp_ValidateImportBatch_ArticleParticipant	Valida participantes y coautores relacionados con artículos científicos.
sp_ProcessImportBatchRow_ArticleParticipant	Procesa información relacionada con autores y participantes institucionales.
sp_RegisterExternalIntegrationResponse	Registra respuestas provenientes de integraciones externas y APIs académicas.
sp_TransformExternalArticleJsonToStaging	Transforma información externa en estructuras compatibles con staging institucional.
RunFullLoadAsync	Ejecuta procesos completos de carga hacia el Data Warehouse institucional.
RunIncrementalLoadAsync	Ejecuta cargas incrementales y actualización progresiva de información analítica.
EtlOrchestrator	Coordina procesos ETL y consolidación de dimensiones y hechos analíticos.

Tabla 16. Procesos de transformación y validación aplicados durante el ETL.

Proceso	Descripción funcional
Validación de obligatoriedad	Verifica campos requeridos antes de consolidar registros institucionales.

Homologación de datos	Normaliza nombres, estados y estructuras académicas institucionales.
Validación referencial	Controla relaciones entre entidades operativas y dimensiones analíticas.
Eliminación de duplicados	Evita redundancia de registros provenientes de múltiples fuentes operativas.
Conversión de formatos	Transforma datos temporales y estructuras utilizadas dentro del OLTP.
Generación de claves sustitutas	Permite administrar dimensiones compartidas dentro de la constelación dimensional institucional.
Consolidación histórica	Mantiene información organizada para reportería y análisis temporal.
Carga incremental	Permite actualizar información analítica sin afectar históricos existentes.
Registro de errores	Controla inconsistencias detectadas durante validación y transformación.
Trazabilidad ETL	Registra procesos ejecutados mediante estructuras de auditoría ETL.

Los procesos ETL implementados permiten garantizar calidad, consistencia y preparación analítica de la información utilizada posteriormente dentro del entorno BI institucional. La incorporación de estructuras de staging, validación y homologación facilita controlar inconsistencias antes de consolidar la información dentro de la constelación dimensional OLAP. Además, los procedimientos almacenados implementados permiten automatizar procesos de validación y procesamiento relacionados con artículos científicos, autores, cargas masivas e integraciones externas. El orquestador ETL desarrollado en .NET permite ejecutar procesos completos e incrementales orientados a poblar dimensiones compartidas y tablas de hechos especializadas utilizadas para reportería institucional. La arquitectura implementada también fortalece la trazabilidad analítica mediante mecanismos de auditoría y control sobre los procesos ETL ejecutados dentro de la plataforma. Finalmente, esta estructura proporciona una base sólida para Business Intelligence institucional y funcionalidades relacionadas con Inteligencia Artificial y análisis predictivo.

Vistas analíticas institucionales y explotación de información

La arquitectura BI implementada incorpora vistas analíticas especializadas orientadas a consolidar información científica, operacional y administrativa utilizada para reportería y análisis institucional dentro de la DIDE-UTA. Estas vistas permiten encapsular consultas complejas sobre la constelación dimensional del Data

Warehouse, facilitando el consumo de información por parte de dashboards, KPIs y reportes institucionales. La utilización de vistas analíticas permite abstraer la complejidad interna del modelo dimensional, evitando que los componentes de reportería dependan directamente de consultas extensas sobre múltiples tablas de hechos y dimensiones. Además, las vistas implementadas integran métricas relacionadas con publicaciones científicas, indexaciones, workflow institucional, validaciones, carga masiva y estructuras dinámicas provenientes del modelo transaccional extensible. Como se observa en la Tabla 17, el entorno BI institucional integra diferentes vistas analíticas utilizadas para reportería estratégica y generación de indicadores multidimensionales. Finalmente, esta arquitectura permite mantener procesos analíticos organizados y preparados para futuras ampliaciones institucionales.

Tabla 17. Vistas analíticas implementadas para reportería institucional.

Vista analítica	Funcionalidad institucional
vw_ArticlePublicationMetrics	Consolida métricas relacionadas con publicaciones científicas institucionales.
vw_ArticleIndexingAnalysis	Permite analizar indexaciones científicas y presencia institucional en bases académicas.
vw_VenueMetricsAnalysis	Gestiona métricas editoriales relacionadas con revistas científicas y cuartiles.
vw_AuthorScientificProduction	Analiza participación institucional de autores e investigadores científicos.
vw_WorkflowPerformance	Permite evaluar tiempos y estados relacionados con workflow institucional.
vw_ValidationErrorAnalysis	Consolida métricas relacionadas con errores y validaciones institucionales.
vw_RegistrationBatchAnalysis	Analiza rendimiento y comportamiento de procesos de carga masiva institucional.
vw_DynamicAttributeAnalysis	Permite analizar atributos dinámicos provenientes de estructuras configurables.
vw_ResearchLineIndicators	Consolida indicadores relacionados con líneas de investigación institucionales.
vw_OpenAccessIndicators	Permite analizar publicaciones relacionadas con acceso abierto institucional.

La utilización de vistas analíticas especializadas permite optimizar procesos de reportería y generación de dashboards dentro del entorno BI institucional. La consolidación previa de información científica y operacional facilita reducir la complejidad de las consultas utilizadas por indicadores y componentes visuales del sistema. Además, las vistas implementadas permiten integrar información proveniente de múltiples tablas de hechos y dimensiones compartidas relacionadas con workflow, validaciones, métricas editoriales, carga masiva y producción científica institucional. Esta arquitectura favorece la mantenibilidad y escalabilidad del entorno analítico, debido a que permite modificar estructuras internas del Data Warehouse sin afectar directamente los componentes de reportería institucional. La incorporación de vistas relacionadas con atributos dinámicos también fortalece la capacidad analítica sobre estructuras configurables provenientes del modelo OLTP extensible regido por metadatos. Finalmente, esta organización proporciona una arquitectura BI robusta, flexible y preparada para funcionalidades relacionadas con Inteligencia Artificial institucional.

Tabla 18. Indicadores institucionales soportados por el entorno BI.

Indicador institucional	Descripción analítica
Producción científica institucional	Permite visualizar el total de publicaciones científicas registradas institucionalmente.
Producción científica por periodo	Analiza tendencias históricas de producción académica institucional.
Artículos indexados en Scopus	Permite evaluar presencia institucional en bases científicas internacionales.
Distribución por cuartiles Q1-Q4	Analiza calidad editorial de publicaciones científicas institucionales.
Indicadores Open Access	Permite evaluar publicaciones relacionadas con acceso abierto institucional.
Producción por línea de investigación	Analiza actividad científica relacionada con líneas institucionales.
Rendimiento de workflow institucional	Evalúa tiempos y estados relacionados con procesos de validación institucional.
Errores de validación institucional	Permite identificar inconsistencias y problemas operativos dentro de registros científicos.
Rendimiento de carga masiva	Analiza comportamiento de procesos institucionales de importación y procesamiento de lotes.

Uso de atributos dinámicos	Permite evaluar comportamiento de estructuras configurables y metadatos institucionales.
Participación institucional de autores	Analiza producción científica relacionada con investigadores institucionales.
Métricas editoriales institucionales	Permite visualizar indicadores SJR, indexaciones y métricas científicas.

La implementación de indicadores institucionales permite transformar información operacional y científica en métricas estratégicas utilizadas para procesos de evaluación, seguimiento y toma de decisiones dentro de la DIDE-UTA. La arquitectura BI desarrollada facilita integrar indicadores relacionados con producción científica, procesos administrativos, validaciones, carga masiva, workflow y métricas editoriales gestionadas por la plataforma institucional. La utilización de una constelación dimensional permite mantener escalabilidad y separación funcional entre diferentes áreas de análisis, evitando concentrar todos los indicadores en una única tabla de hechos. Además, la integración de componentes dinámicos provenientes del modelo OLTP extensible fortalece la adaptabilidad del entorno analítico frente a nuevos requerimientos institucionales. Esta organización también reduce riesgos asociados al mantenimiento del Data Warehouse, debido a que permite incorporar nuevas dimensiones, hechos y vistas analíticas sin afectar directamente la arquitectura general del entorno BI. Finalmente, el modelo implementado constituye una plataforma analítica avanzada preparada para reportería estratégica, Business Intelligence institucional e Inteligencia Artificial.

3.2.4 Diseño visual del sistema

El diseño visual del sistema fue desarrollado considerando criterios de organización funcional, experiencia de usuario, accesibilidad institucional y navegación modular dentro de la plataforma propuesta. La estructura visual implementada permite integrar módulos relacionados con registro científico, carga masiva, workflow institucional, reportería analítica y administración dinámica mediante interfaces organizadas y orientadas a procesos institucionales específicos. Asimismo, el sistema incorpora componentes visuales relacionados con formularios dinámicos, dashboards analíticos, paneles administrativos y mecanismos de validación institucional contruidos sobre la

arquitectura tecnológica implementada. De igual manera, la organización visual desarrollada facilita la interacción entre usuarios institucionales y procesos operativos relacionados con gestión científica y BI. Debido a la amplitud funcional del sistema y la cantidad de interfaces implementadas, las capturas correspondientes al diseño visual y funcionamiento de los módulos desarrollados fueron organizadas dentro del Anexo 11 a fin de evidenciar de forma complementaria el sistema implementado.

3.3 Desarrollo de la propuesta

El desarrollo de la propuesta tecnológica se realizó a partir de los requerimientos funcionales y no funcionales definidos durante la etapa de análisis y del diseño arquitectónico establecido para el sistema institucional. Esta fase tuvo como finalidad construir una plataforma orientada a la gestión de información científica, reportería analítica, visualización de indicadores y generación de funciones de recomendación y predicción mediante técnicas de BI e IA. Para ello, se consideró una organización modular que permitió desarrollar componentes relacionados con autenticación, gestión de artículos científicos, carga masiva, workflow institucional, integración con fuentes externas, procesos ETL, dashboards, reportes y análisis predictivo.

El desarrollo del sistema se apoyó en una metodología ágil seleccionada mediante un análisis comparativo de enfoques aplicables al contexto institucional. Esta selección permitió organizar el trabajo de forma iterativa, controlar los avances del sistema, validar incrementos funcionales y mantener trazabilidad sobre los módulos desarrollados. Además, el enfoque adoptado permitió equilibrar la flexibilidad requerida durante el desarrollo con la necesidad de documentación, pruebas y control formal propias de un entorno universitario.

3.3.1 Selección de la metodología de desarrollo

Para el desarrollo de la propuesta tecnológica se realizó un análisis comparativo de metodologías ágiles, con el propósito de seleccionar el enfoque más adecuado para las características del proyecto y el contexto institucional de la DIDE-UTA. Debido a que el sistema contempla módulos relacionados con gestión de artículos científicos, carga masiva, workflow institucional, reportería analítica, procesos ETL y funciones de

recomendación y predicción, fue necesario considerar una metodología que permitiera organizar el trabajo de manera iterativa, controlar cambios, mantener documentación y asegurar calidad técnica durante el desarrollo. En este sentido, se evaluaron metodologías ágiles como Scrum, Kanban, Scrumban, Extreme Programming, Lean Software Development, DSDM/AgilePM, FDD y Scrum + XP.

La evaluación metodológica se realizó mediante un análisis cualitativo y una matriz cuantitativa ponderada, considerando criterios relacionados con la cadencia de trabajo, gestión de cambios, documentación institucional, elaboración de artefactos, planificación por módulos, pruebas y control del flujo de actividades. El documento completo de evaluación metodológica, donde se incluye la comparación cualitativa detallada, la matriz de puntuación por criterio y los cálculos realizados para cada metodología, se presenta en el Anexo 10. En el cuerpo principal de la tesis se incluyen únicamente los criterios de evaluación y los resultados finales de la matriz ponderada, con el fin de mantener una explicación clara y directamente relacionada con la selección metodológica aplicada al desarrollo del sistema.

Tabla 19. Criterios de evaluación para la selección de la metodología de desarrollo.

Código	Criterio	Peso
C1	Entregas semanales y cadencia de reuniones	20 %
C2	Gestión de cambios constantes	20 %
C3	Documentación y aprobaciones institucionales	15 %
C4	Diseño y artefactos, como mockups, casos y diagramas	10 %
C5	Metas mensuales por módulos y releases	15 %
C6	Planificación de pruebas y despliegue	10 %
C7	Visibilidad y control del flujo mediante tablero y límites WIP	10 %
	Total	100 %

Los criterios definidos permitieron valorar cada metodología de acuerdo con las necesidades reales del proyecto. Se asignó mayor peso a las entregas semanales y a la gestión de cambios, debido a que el desarrollo del sistema requería avances progresivos y ajustes derivados de la retroalimentación institucional. También se otorgó relevancia a la documentación y aprobaciones, considerando que el proyecto se desarrolló dentro de un entorno universitario donde se requiere evidencia formal del proceso. Asimismo, la planificación de pruebas y despliegue fue considerada un

criterio importante, debido a la necesidad de validar módulos críticos relacionados con reportería, procesamiento de datos, ETL e Inteligencia Artificial.

Tabla 20. Resultados de la matriz ponderada de metodologías ágiles.

Metodología	Puntaje	Porcentaje
Scrumban	4.75	95 %
Scrum + XP	4.75	95 %
DSDM / AgilePM	4.60	92 %
Scrum	4.35	87 %
XP	4.05	81 %
FDD	3.80	76 %
Kanban	3.60	72 %
Lean	3.50	70 %

Los resultados de la matriz ponderada muestran que Scrumban y Scrum + XP obtuvieron el mayor nivel de adecuación al proyecto, alcanzando ambos un 95 %. Sin embargo, se seleccionó Scrum + XP como metodología de desarrollo, debido a que presentó mejor desempeño en el criterio relacionado con planificación de pruebas y despliegue. Este aspecto fue considerado fundamental para el proyecto, ya que la solución desarrollada integra componentes de procesamiento de datos, reportería analítica, procesos ETL y funciones de recomendación y predicción, los cuales requieren validación técnica constante y control de calidad durante su implementación.

La selección de Scrum + XP permitió combinar la organización iterativa de Scrum con las prácticas técnicas de ingeniería propuestas por Extreme Programming. Scrum aportó la planificación por incrementos, organización de tareas, revisión de avances y control del desarrollo por módulos. Por su parte, XP complementó el proceso mediante prácticas orientadas a mejorar la calidad del software, como pruebas frecuentes, revisión continua del código, refactorización, integración progresiva y validación técnica de funcionalidades. Esta combinación resultó adecuada para el contexto del proyecto, ya que permitió mantener una planificación estructurada sin perder flexibilidad frente a los ajustes derivados de la retroalimentación institucional.

En consecuencia, la metodología Scrum + XP fue considerada pertinente para el desarrollo del sistema de reportería y visualización de datos con funciones de

recomendación y predicción. Su aplicación permitió organizar el proyecto por módulos funcionales, controlar entregas progresivas, validar incrementos del sistema y mantener trazabilidad sobre las actividades desarrolladas. Además, este enfoque favoreció la calidad técnica de la solución y proporcionó una base metodológica adecuada para las fases posteriores de pruebas, resultados e implantación del sistema.

3.3.2 Aplicación de la metodología

La metodología híbrida Scrum + XP fue aplicada durante el desarrollo del sistema con el propósito de organizar el trabajo de forma iterativa, controlar los avances funcionales y asegurar la calidad técnica de los módulos implementados. Su aplicación permitió estructurar el desarrollo en incrementos progresivos, priorizando las funcionalidades definidas en los requerimientos del sistema y organizándolas mediante historias de usuario, sprints, criterios de aceptación y revisiones técnicas. Aunque los documentos metodológicos presentan la planificación general del desarrollo, su aplicación se orientó directamente a los módulos definidos previamente en el análisis y diseño de la propuesta, tales como gestión de artículos científicos, carga masiva, workflow institucional, integración con APIs externas, reportería analítica, procesos ETL, Business Intelligence e Inteligencia Artificial.

El proceso metodológico se organizó mediante sprints semanales, cada uno orientado al desarrollo de un conjunto específico de funcionalidades. Esta organización permitió avanzar desde los componentes base del sistema, como autenticación, administración de usuarios y registro de artículos, hacia módulos más especializados como carga masiva de información científica, validación de datos, reportería institucional, dashboards, exportación de reportes, procesamiento analítico y funciones de recomendación y predicción. El plan completo de ejecución metodológica, donde se describen los roles adaptados, artefactos utilizados, prácticas XP y métricas de seguimiento, se presenta en el Anexo 5.

Adaptación de roles metodológicos

Debido a que el desarrollo fue realizado por un único desarrollador, los roles tradicionales de Scrum fueron adaptados al contexto académico del proyecto. El tesista asumió las funciones relacionadas con el desarrollo técnico, la organización del

proceso y la gestión de tareas. El tutor participó como responsable de la validación técnica, mientras que el responsable institucional de la DIDE intervino en la validación funcional de los avances del sistema. Esta adaptación permitió mantener una estructura metodológica organizada, sin perder la participación de los actores clave durante el proceso de desarrollo.

Tabla 21. Roles adaptados para la aplicación de Scrum + XP.

Rol metodológico	Responsable	Función dentro del proyecto
Development Team	Tesista	Desarrollo e implementación del sistema
Scrum Master técnico	Tesista	Organización del proceso, seguimiento de tareas y control del avance
Product Owner técnico	Tutor de tesis	Validación técnica de funcionalidades y revisión de avances
Product Owner funcional	Responsable DIDE	Validación funcional de módulos y retroalimentación institucional

La adaptación de roles permitió mantener control sobre el proceso de desarrollo y asegurar que las funcionalidades implementadas respondieran tanto a criterios técnicos como a necesidades institucionales. Además, esta organización facilitó la comunicación entre el desarrollo del sistema, la revisión académica y la validación funcional por parte de los usuarios involucrados.

Organización del desarrollo por sprints

El desarrollo fue organizado en cuatro sprints, cada uno con un enfoque funcional específico. Esta planificación permitió distribuir el trabajo en etapas progresivas, iniciando con la infraestructura base del sistema y continuando con los módulos de gestión científica, reportería, procesamiento de datos, Inteligencia Artificial e integración final. El plan de sprints completo, junto con la matriz de trazabilidad entre casos de uso, historias de usuario y sprints, se presenta en el Anexo 5.

Tabla 22. Plan resumido de sprints aplicados al desarrollo del sistema.

Sprint	Enfoque principal	Funcionalidades desarrolladas	Entregable
Sprint 1	Infraestructura base y gestión operativa	Autenticación, gestión de usuarios, registro de artículos científicos, carga masiva inicial y	Sistema base funcional con usuarios, registro y

		estructura base del modelo transaccional	almacenamiento operativo
Sprint 2	Reportería y visualización institucional	Dashboards, generación de reportes, filtros analíticos, exportación de información y consultas hacia el modelo analítico	Sistema con reportería institucional, indicadores y visualización analítica
Sprint 3	Procesamiento de datos e Inteligencia Artificial	Limpieza y transformación de datos, preparación del modelo analítico, entrenamiento de modelo predictivo, generación de predicciones y recomendaciones	Sistema con funciones de recomendación y predicción basadas en datos
Sprint 4	Integración, trazabilidad y cierre funcional	Clasificación OCDE, auditoría, ajustes de workflow, validación de módulos, integración final y pruebas generales	Versión final del sistema lista para evaluación

La planificación por sprints permitió mantener una secuencia lógica de desarrollo. En el primer sprint se priorizaron las funcionalidades base y los procesos operativos iniciales, en el segundo se incorporaron componentes de reportería y visualización, en el tercero se trabajó sobre procesamiento de datos e Inteligencia Artificial, y en el cuarto se consolidaron los procesos de integración, auditoría, trazabilidad y validación final. Esta distribución permitió desarrollar el sistema de forma incremental y mantener coherencia entre los requerimientos, los módulos diseñados y las funcionalidades implementadas.

Aplicación de prácticas XP

Como complemento a Scrum, se aplicaron prácticas de Extreme Programming orientadas a fortalecer la calidad técnica del sistema. Entre las prácticas adoptadas se consideraron la validación técnica frecuente en componentes críticos, integración continua mediante repositorio GitHub, refactorización progresiva del código, entregas frecuentes y control de versiones. Estas prácticas permitieron mantener un desarrollo más ordenado, reducir errores, mejorar la mantenibilidad del código y validar de forma progresiva las funcionalidades implementadas. El plan de aseguramiento de calidad basado en XP se presenta completo en el Anexo 7.

Tabla 23. Prácticas XP aplicadas durante el desarrollo.

Práctica XP	Aplicación dentro del proyecto
Validación técnica frecuente	Validación de componentes críticos relacionados con lógica de negocio, reportería, predicción y recomendación mediante pruebas funcionales, integración y revisión técnica.
Integración continua	Uso de GitHub para integrar cambios y mantener trazabilidad del código
Refactorización continua	Mejora progresiva de la estructura interna del código sin alterar su comportamiento funcional
Entregas frecuentes	Presentación de incrementos funcionales durante cada sprint
Control de versiones	Registro histórico de cambios realizados durante el desarrollo

La incorporación de prácticas XP permitió complementar la gestión iterativa de Scrum con mecanismos técnicos de calidad. Esto fue especialmente importante debido a que el sistema incluyó procesos de validación de datos, generación de reportes, integración con el modelo analítico y funcionalidades de recomendación y predicción. Por ello, las prácticas de prueba, integración y refactorización aportaron mayor estabilidad al desarrollo.

Definition of Done aplicada

Para determinar cuándo una historia de usuario o incremento podía considerarse finalizado, se definió una Definition of Done. Esta incluyó criterios relacionados con implementación funcional, cumplimiento de criterios de aceptación, integración al repositorio, compilación correcta, ejecución de pruebas, ausencia de errores críticos, revisión técnica y validación funcional. El documento completo de Definition of Done se presenta en el Anexo 6.

Tabla 24. Criterios resumidos de Definition of Done aplicados al proyecto.

Criterio DoD	Aplicación dentro del proyecto
Implementación funcional	La funcionalidad debía estar desarrollada según la historia de usuario
Criterios de aceptación	La funcionalidad debía cumplir las condiciones definidas en el Product Backlog
Integración al sistema	El código debía integrarse correctamente al repositorio del proyecto
Compilación sin errores	El sistema debía ejecutarse sin errores críticos de compilación

Validación mediante pruebas	Se debían ejecutar pruebas funcionales, pruebas de integración o validación técnica básica según el tipo de funcionalidad.
Revisión técnica	La funcionalidad debía ser revisada técnicamente con el tutor
Validación funcional	Los módulos orientados al usuario debían ser validados mediante demostración funcional

La Tabla 24 muestra, de forma resumida la aplicación de la Definition of Done quien permitió asegurar que las funcionalidades no fueran consideradas terminadas únicamente por estar programadas, sino por cumplir condiciones mínimas de calidad, integración, prueba y validación. Esto fortaleció la trazabilidad del proceso y permitió mantener una relación directa entre las historias de usuario, los incrementos desarrollados y las evidencias de validación utilizadas posteriormente en las pruebas del sistema.

Cierre de la aplicación metodológica

La aplicación de Scrum + XP permitió desarrollar el sistema de forma ordenada, incremental y verificable. Scrum facilitó la organización del trabajo mediante sprints, backlog, planificación y revisión de avances, mientras que XP aportó prácticas técnicas orientadas a mejorar la calidad del software. En conjunto, esta metodología permitió controlar el avance del proyecto, validar funcionalidades con los actores involucrados y mantener evidencia metodológica del proceso de implementación. Esta aplicación sirvió como base para el desarrollo modular del sistema y para la posterior ejecución de pruebas, evaluación de resultados e implantación de la solución tecnológica en el contexto de la DIDE-UTA.

3.3.3 Selección de herramientas tecnológicas

Las herramientas tecnológicas utilizadas durante el desarrollo del sistema fueron seleccionadas conforme a los lineamientos técnicos propuestos por la Dirección de Tecnologías de Información y Comunicación (DITIC) de la Universidad Técnica de Ambato, ver Tabla 25. Debido a que estas tecnologías ya fueron descritas en la etapa de diseño de la propuesta, en este apartado se presentan únicamente como soporte directo para la implementación de los módulos funcionales del sistema.

Tabla 25. Herramientas tecnológicas utilizadas durante el desarrollo del sistema.

Herramienta / tecnología	Aplicación dentro del desarrollo
.NET / C#	Desarrollo del backend, APIs, servicios y lógica de negocio
Blazor WebAssembly	Desarrollo de la interfaz web interactiva
SQL Server	Gestión del modelo transaccional OLTP y del modelo analítico OLAP
Entity Framework Core	Acceso a datos, migraciones y administración de entidades
REST / JSON	Integración con APIs externas y servicios académicos
QuestPDF	Generación de reportes institucionales exportables
ScottPlot	Generación de gráficos analíticos para reportes
Git / GitHub	Control de versiones y trazabilidad del código
ML.NET	Implementación de funciones de recomendación y predicción

La utilización de estas herramientas permitió implementar los módulos operativos, analíticos y predictivos definidos en el diseño del sistema, manteniendo coherencia con la arquitectura institucional propuesta por DITIC y con los requerimientos funcionales de la DIDE-UTA.

3.3.4 Desarrollo modular del sistema

El desarrollo del sistema se realizó mediante una organización modular, tomando como base los requerimientos funcionales definidos durante la etapa de análisis, los diagramas UML elaborados y la planificación establecida en el Product Backlog. Esta organización permitió dividir la solución en componentes funcionales relacionados con seguridad, gestión de artículos científicos, carga masiva, workflow institucional, integración con fuentes externas, reportería analítica, procesos ETL, Business Intelligence e Inteligencia Artificial. Cada módulo fue implementado de forma progresiva durante los sprints definidos, permitiendo validar avances parciales y mantener trazabilidad entre las historias de usuario y las funcionalidades desarrolladas, ver Tabla 26.

La construcción modular permitió separar responsabilidades dentro del sistema y facilitar el mantenimiento de la plataforma. Los módulos operativos se orientaron a la gestión diaria de información científica, mientras que los módulos analíticos permitieron consolidar datos, generar indicadores y visualizar información institucional. De igual manera, los componentes de Inteligencia Artificial fueron

incorporados como parte del análisis avanzado, con el propósito de apoyar procesos de recomendación y predicción sobre la información histórica disponible. Esta estructura permitió integrar los procesos de gestión, validación, reportería y análisis dentro de una misma solución tecnológica.

Tabla 26. Módulos desarrollados dentro del sistema.

Módulo	Funcionalidades implementadas	Relación con el sistema
Seguridad y administración	Inicio de sesión, gestión de usuarios, roles y permisos	Controla el acceso a la plataforma y protege la información institucional
Gestión de artículos científicos	Registro, edición, clasificación OCDE, autores, revistas e indexaciones	Administra la información principal de producción científica
Carga masiva de información	Importación de archivos, validación de estructura, procesamiento de lotes y control de errores	Reduce el ingreso manual de información y permite cargar registros por lotes
Workflow institucional	Revisión, aprobación, devolución, observaciones y cambios de estado	Controla el proceso institucional de validación de registros científicos
Integración con APIs externas	Consulta de información científica mediante servicios externos y normalización de resultados	Permite enriquecer y validar datos científicos institucionales
Modelo transaccional OLTP	Almacenamiento operativo de artículos, autores, revistas, catálogos, workflow y trazabilidad	Soporta las operaciones diarias del sistema
Procesos ETL	Extracción, transformación, depuración y carga hacia el modelo analítico	Conecta el entorno operativo con el Data Warehouse
Modelo analítico OLAP	Consolidación de información histórica mediante constelación dimensional	Proporciona la base para reportería, KPIs y análisis institucional
Reportería y Business Intelligence	Dashboards, gráficos, filtros, KPIs y reportes exportables	Permite visualizar información estratégica para la toma de decisiones
Inteligencia Artificial	Procesamiento de datos, generación de predicciones y recomendaciones	Apoya el análisis avanzado de producción científica institucional
Auditoría y trazabilidad	Registro de acciones, eventos y cambios realizados por los usuarios	Fortalece el seguimiento y control de los procesos del sistema

El módulo de seguridad y administración permitió controlar el acceso al sistema mediante usuarios, roles y permisos. Este componente fue necesario para garantizar

que cada actor institucional pueda acceder únicamente a las funcionalidades correspondientes a su perfil. Además, la administración de usuarios permitió organizar el uso de la plataforma y mantener un control básico sobre las acciones realizadas dentro del sistema.

El módulo de gestión de artículos científicos constituyó uno de los componentes principales de la plataforma, debido a que permitió registrar y administrar la información relacionada con la producción científica institucional. Este módulo integró datos bibliográficos, autores, revistas, indexaciones, estados de publicación, clasificación OCDE y otros metadatos necesarios para la gestión y reportería posterior. A partir de este componente se alimentaron otros procesos como workflow, reportería, ETL y análisis institucional.

El módulo de carga masiva fue desarrollado para facilitar el ingreso de grandes volúmenes de información científica mediante archivos estructurados. Este componente permitió validar la estructura de los datos, identificar errores, procesar lotes y consolidar registros válidos dentro del modelo transaccional. Su implementación contribuyó a reducir la dependencia de procesos manuales y a mejorar el control sobre la calidad de la información ingresada.

El módulo de workflow institucional permitió gestionar el proceso de revisión, aprobación o devolución de artículos y registros cargados en el sistema. A través de este componente se administraron estados, observaciones y acciones realizadas por los usuarios responsables de la validación institucional. Además, el workflow se complementó con mecanismos de trazabilidad que permitieron registrar cambios y dar seguimiento al ciclo de vida de la información científica.

El módulo de integración con APIs externas permitió consultar información científica desde fuentes académicas externas, con el propósito de enriquecer y validar los registros institucionales. Este componente fue diseñado para trabajar mediante servicios REST y estructuras JSON, facilitando la obtención de metadatos científicos y su posterior almacenamiento dentro del sistema.

Los módulos analíticos se desarrollaron mediante la integración entre el modelo OLTP, los procesos ETL y el modelo OLAP institucional. El entorno transaccional permitió almacenar la información operativa, mientras que los procesos ETL se

encargaron de transformar y consolidar los datos hacia el Data Warehouse. A partir de esta información analítica se implementaron dashboards, KPIs, gráficos y reportes orientados a la toma de decisiones dentro de la DIDE-UTA.

Finalmente, el módulo de Inteligencia Artificial fue incorporado para apoyar procesos de recomendación y predicción institucional. Este componente utilizó información histórica y consolidada del entorno analítico para generar resultados orientados al análisis avanzado de la producción científica. De esta manera, el desarrollo modular permitió cumplir con el propósito de implementar una plataforma que no solo registre información, sino que también facilite su análisis, visualización y aprovechamiento mediante técnicas de Business Intelligence e Inteligencia Artificial.

3.3.5 Implementación de funciones de recomendación y predicción

El módulo de Inteligencia Artificial fue implementado con el propósito de apoyar la toma de decisiones institucionales a partir de la información histórica relacionada con la producción científica registrada en el sistema. Su finalidad no consiste en reemplazar el criterio técnico o administrativo de los usuarios de la DIDE-UTA, sino en transformar los datos consolidados en señales analíticas útiles para la planificación, seguimiento y mejora continua. Para ello, el módulo utiliza información proveniente del modelo analítico institucional, principalmente datos asociados a artículos, autores, indexaciones, facultades, líneas de investigación, métricas editoriales y procesos de workflow. De esta manera, la Inteligencia Artificial se integra al entorno de Business Intelligence como una capa adicional de análisis predictivo y recomendación institucional.

El módulo combina modelos predictivos de series temporales con reglas explicables de recomendación. Esta decisión permitió incorporar funciones de predicción sobre la producción científica institucional y, al mismo tiempo, generar recomendaciones comprensibles para el usuario. El informe técnico completo del módulo de Inteligencia Artificial y el informe base, donde se detallan servicios implementados, endpoints, métricas, reglas, fórmulas, escenarios, limitaciones y mejoras futuras, se presentan en los anexos correspondientes.

Funciones principales del módulo de Inteligencia Artificial

El sistema realiza predicciones orientadas a estimar la producción científica institucional esperada durante los siguientes seis meses. La variable objetivo corresponde al total de artículos por mes, tomando como fuente los datos históricos consolidados en la reportería institucional y el Data Warehouse. Como resultado, el módulo genera el mes proyectado, el valor esperado de artículos, límites inferior y superior de confianza, algoritmo utilizado, porcentaje de confianza e interpretación del resultado. Esta información permite anticipar si la producción científica esperada tiende a subir, bajar o mantenerse estable, facilitando la comparación entre la producción real y la producción estimada.

Además de la predicción general, el módulo genera análisis predictivos segmentados por facultad y por la línea de investigación. En el caso de las facultades, el sistema calcula artículos históricos, meses disponibles, producción esperada para los próximos seis meses, tendencia, confianza estimada y recomendación asociada. Esta funcionalidad permite identificar unidades académicas con mayor o menor ritmo de producción y priorizar acciones de seguimiento institucional. De forma similar, la predicción por línea de investigación permite reconocer áreas activas, estables o en descenso, así como contrastar la producción científica con las áreas estratégicas definidas por la universidad.

Tabla 27. Funciones principales del módulo de Inteligencia Artificial.

Función IA	Descripción	Uso institucional
Predicción mensual de producción científica	Estima el total esperado de artículos para los próximos seis meses.	Anticipar tendencias generales de producción científica.
Predicción por facultad	Calcula producción esperada, tendencia y confianza por unidad académica.	Identificar facultades que requieren seguimiento o acompañamiento.
Predicción por línea de investigación	Estima el comportamiento futuro de la producción por área temática.	Apoyar la planificación investigativa institucional.

Recomendaciones de calidad de datos	Sugiere completar autores, filiación, ORCID, indexación y cuartiles.	Mejorar la confiabilidad del dataset institucional.
Recomendaciones editoriales	Analiza indexaciones, cuartiles, acceso abierto y revistas.	Orientar estrategias de publicación y normalización editorial.
Recomendaciones de colaboración	Identifica autores, coautorías y señales de trazabilidad institucional.	Fortalecer redes académicas y colaboración científica.

El componente de recomendación fue desarrollado mediante reglas explicables basadas en indicadores, calidad de datos y resultados predictivos. Estas recomendaciones se orientan a mejorar la interpretación de la información generada por el sistema y evitar que los usuarios asuman las predicciones como valores absolutos. Por ejemplo, cuando el sistema detecta baja confianza, alta volatilidad mensual o pocos datos históricos, recomienda utilizar el pronóstico como alerta temprana y no como una meta cerrada. También genera recomendaciones relacionadas con calidad del dataset, completitud de autores, filiación, ORCID, indexación, cuartiles editoriales y cobertura de información científica, se puede verificar las mismas en Tabla 27.

Algoritmos utilizados en el módulo IA

Desde el punto de vista técnico, el sistema compara diferentes algoritmos o modelos candidatos para seleccionar el de mejor desempeño. Entre los modelos considerados se encuentran la persistencia del último valor, el promedio móvil de tres meses, la regresión lineal temporal y el modelo ForecastBySsa de ML.NET. La persistencia del último valor se utiliza como línea base mínima, mientras que el promedio móvil permite suavizar fluctuaciones en series pequeñas. La regresión lineal temporal permite identificar tendencias crecientes o descendentes, y ForecastBySsa se emplea como modelo principal de pronóstico de series temporales cuando existe suficiente información histórica, se puede evidenciar estos algoritmos en Tabla 28.

Tabla 28. Algoritmos utilizados en el módulo IA.

Algoritmo	Tipo	Aplicación
Persistencia del último valor	Línea base	Usa el último mes observado como predicción inicial.
Promedio móvil de tres meses	Línea base interpretable	Suaviza fluctuaciones recientes de producción científica.
Regresión lineal temporal	Regresión	Identifica tendencias crecientes o descendentes en la serie histórica.
ForecastBySsa de ML.NET	Pronóstico de series temporales	Genera predicciones a seis meses con intervalos de confianza.

El modelo ForecastBySsa de ML.NET se basa en el análisis de series temporales mediante Singular Spectrum Analysis. En el sistema, este modelo se utiliza como alternativa principal cuando la serie histórica cuenta con suficientes meses disponibles, debido a que permite generar predicciones con horizonte de seis meses e intervalos de confianza. No obstante, el sistema no depende de un único algoritmo, ya que compara modelos candidatos y selecciona aquel que presenta mejor desempeño según las métricas de evaluación definidas.

Entrenamiento, validación y selección del modelo

El entrenamiento del modelo se ejecuta desde el backend del sistema a partir de la producción científica mensual. Para evitar evaluaciones incorrectas en datos temporales, el sistema utiliza validación temporal y no mezcla aleatoria de registros. En este proceso, los primeros meses se destinan al entrenamiento y los últimos meses a la validación, respetando el orden cronológico de la serie. La cantidad de meses destinados a validación depende del volumen histórico disponible: si existen 18 meses o más se reservan seis meses para validación, si existen entre 12 y 17 meses se reservan cuatro, si existen entre 8 y 11 meses se reservan tres, y si existen menos de ocho meses el sistema marca el escenario como insuficiente para validación.

Tabla 29. Métricas de evaluación del modelo predictivo.

Métrica	Descripción	Uso dentro del sistema
MAE	Error absoluto medio.	Métrica principal para seleccionar el mejor modelo.
RMSE	Raíz del error cuadrático medio.	Penaliza errores grandes y complementa la evaluación.
MAPE	Error porcentual medio.	Permite interpretar el error relativo, con precaución en meses de baja producción.
Score	Aproximación de desempeño.	Estima el desempeño general a partir del error porcentual.

En Tabla 29 se muestra la evaluación de los algoritmos que se realiza mediante métricas de error, principalmente MAE, RMSE y MAPE. El MAE se utiliza como métrica principal para seleccionar el modelo, debido a que representa el error absoluto medio entre los valores reales y los valores estimados. El RMSE se utiliza como métrica complementaria porque penaliza errores grandes, mientras que el MAPE permite observar el error porcentual medio, aunque su interpretación puede ser inestable cuando existen meses con pocos artículos. El sistema selecciona el modelo con menor MAE y, en caso de empate, considera el menor RMSE. El modelo ganador se registra como mejor algoritmo y puede ser promovido como versión activa dentro del sistema.

Base del módulo IA

El módulo de Inteligencia Artificial incorpora una base orientada a evaluar la preparación de los datos, generar predicciones, calcular niveles de confianza y emitir recomendaciones explicables. Para ello, el sistema utiliza indicadores ponderados, modelos de series temporales, métricas de error y reglas de decisión basadas en umbrales institucionales. Esta base permite que las predicciones y recomendaciones no se presenten como resultados aislados, sino como salidas analíticas respaldadas por indicadores medibles y criterios de evaluación.

Uno de los indicadores principales corresponde al OverallScore, utilizado para estimar el nivel de preparación del sistema para ejecutar predicciones y recomendaciones confiables, estos indicadores se puede evidencias en Tabla 30. Este indicador combina cinco dimensiones: volumen de datos, trazabilidad de autores, cobertura de indexación, calidad de carga y señales de workflow. La ponderación asigna mayor

peso al volumen de datos, debido a que la cantidad de registros históricos influye directamente en la estabilidad de los modelos predictivos. También se considera la trazabilidad de autores y la cobertura de indexación, debido a su importancia para generar recomendaciones relacionadas con colaboración, filiación, visibilidad científica y estrategias editoriales.

Tabla 30. Indicadores utilizados para calcular la preparación del módulo IA.

Indicador	Peso	Propósito
Volumen de datos	30 %	Evalúa si existen suficientes artículos para análisis predictivo.
Trazabilidad de autores	22 %	Mide la disponibilidad de información asociada a autores.
Cobertura de indexación	20 %	Evalúa la existencia de datos sobre fuentes, bases y cuartiles.
Calidad de carga	14 %	Mide la proporción de registros procesados correctamente.
Señales de workflow	14 %	Evalúa evidencia operativa del proceso de revisión y devolución.
Total	100 %	

De manera general, el cálculo del indicador de preparación se expresa mediante una combinación ponderada de los puntajes parciales. Esta estructura permite clasificar el estado del dataset en niveles de preparación alta, media, inicial o baja. Con ello, el sistema no solo ejecuta predicciones, sino que también informa al usuario si la base histórica tiene suficiente calidad y volumen para interpretar los resultados con mayor confianza.

Tabla 31. Fórmulas principales utilizadas en el módulo IA.

Componente	Fórmula general	Interpretación
OverallScore	$ \begin{aligned} & (DatasetVolumeScore \times 0.30) \\ & + (AuthorTraceScore \times 0.22) \\ & + (IndexingCoverageScore \times 0.20) \\ & + (\max(LoadQualityScore, 55) \\ & \times 0.14) \\ & + (\max(WorkflowSignalScore, 45) \\ & \times 0.14) \end{aligned} $	Estima el nivel general de preparación para IA.
Trazabilidad de autores	$ \begin{aligned} & (Artículos\ con\ trazabilidad / Total\ de\ artículos) \\ & \times 100 \end{aligned} $	Mide completitud de información autoral.
Calidad de carga	$ (Filas\ exitosas / Total\ de\ filas\ procesadas) \times 100 $	Evalúa calidad de importación o staging.
Cobertura de indexación	$ \begin{aligned} & (Artículos\ con\ indexación / Total\ de\ artículos) \\ & \times 100 \end{aligned} $	Mide disponibilidad de información editorial.
Promedio móvil 3 meses	$ (Valor(t) + Valor(t - 1) + Valor(t - 2)) / 3 $	Suaviza fluctuaciones recientes.
MAE	$\sum$	Real - Predicho
RMSE	$ \sqrt{(\sum (Real - Predicho)^2 / n)} $	Penaliza errores grandes.
MAPE	$\sum ($	Real - Predicho
Score	$ \max(0, 100 - MAPE) $	Indicador complementario de desempeño.

Para la predicción de producción científica, el sistema utiliza una serie temporal mensual donde la variable objetivo corresponde al total de artículos registrados por mes. Sobre esta serie se aplican modelos base, regresión lineal temporal y el modelo ForecastBySsa de ML.NET. El promedio móvil de tres meses permite suavizar fluctuaciones recientes, mientras que la regresión lineal temporal permite identificar tendencias crecientes, descendentes o estables. Por su parte, ForecastBySsa se utiliza

como modelo principal de pronóstico cuando existe suficiente historial, generando valores esperados e intervalos de confianza para los seis meses posteriores.

Además, el sistema calcula un puntaje de confianza para la predicción considerando la cantidad de meses históricos disponibles y la volatilidad de la serie mensual. Una serie más larga incrementa la confianza del pronóstico, mientras que una alta variabilidad reduce dicho puntaje. Esta lógica evita presentar las predicciones como valores absolutos y permite comunicar al usuario que el resultado debe interpretarse como una estimación sujeta a incertidumbre.

Las recomendaciones institucionales se generan mediante reglas explicables basadas en umbrales. Por ejemplo, si el volumen de artículos es bajo, el sistema recomienda ampliar la base histórica, si la confianza del pronóstico es baja, sugiere utilizar la predicción como alerta temprana, si la trazabilidad de autores es insuficiente, recomienda completar información autoral, y si la cobertura de indexación es limitada, propone normalizar fuentes, cuartiles y métricas editoriales. Este enfoque permite que cada recomendación tenga una razón comprensible para el usuario y esté respaldada por indicadores medibles.

Arquitectura funcional del módulo IA

La arquitectura funcional del módulo se integra con el flujo general del sistema. Los datos se registran inicialmente en el modelo transaccional OLTP, posteriormente, son procesados mediante ETL y consolidados en el Data Warehouse institucional. A partir de esta información, la reportería institucional alimenta el servicio de Inteligencia Artificial, el cual genera predicciones, recomendaciones y diagnósticos de preparación de datos, ver arquitectura en Figura 45. Finalmente, los resultados son presentados al usuario mediante el dashboard IA, permitiendo visualizar tendencias, alertas, recomendaciones y estados de confianza del análisis.

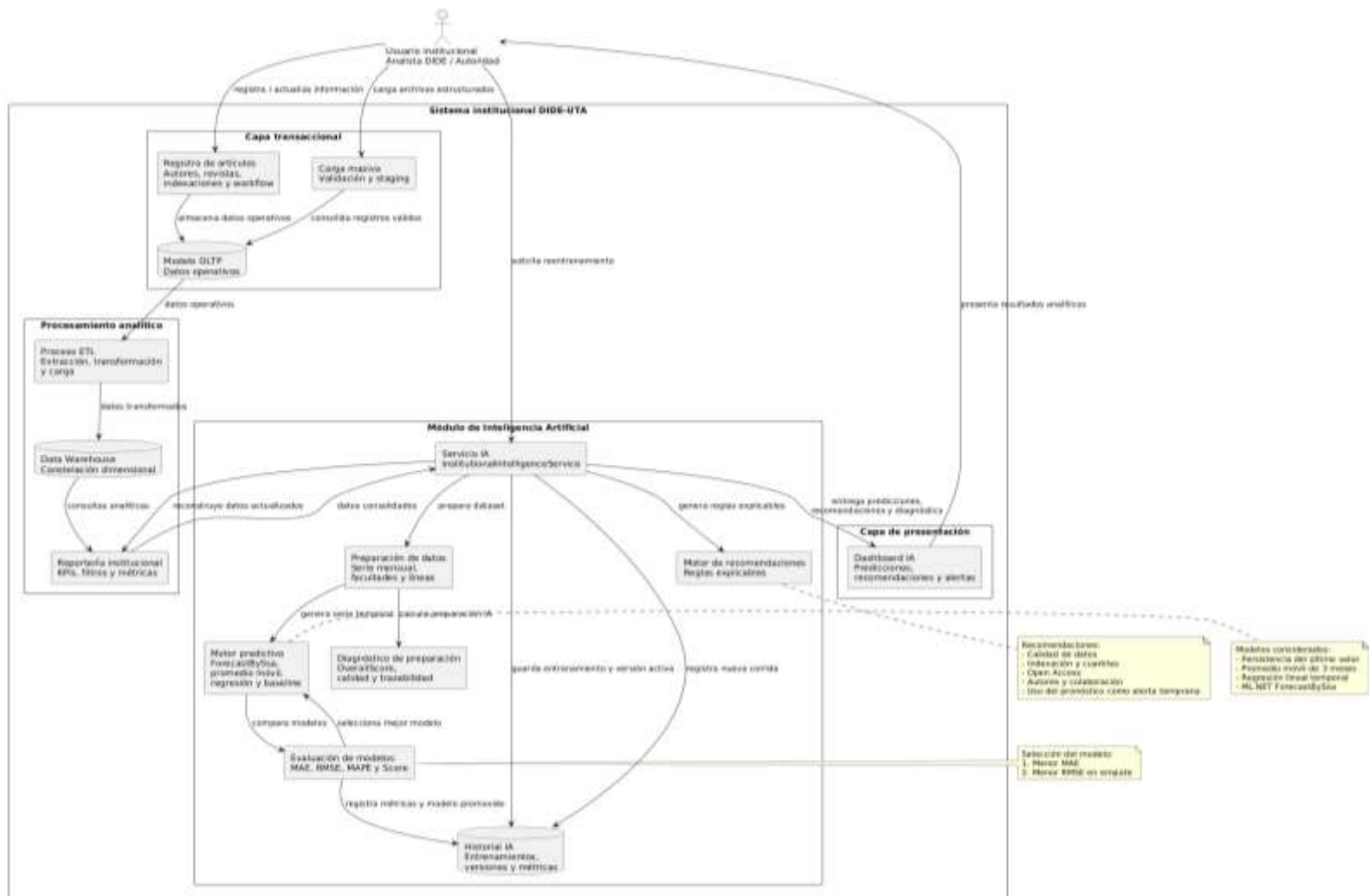


Figura 45. Arquitectura funcional del módulo de Inteligencia Artificial.

El módulo IA también incorpora mecanismos de reentrenamiento, historial y auditoría. El reentrenamiento puede ejecutarse desde el módulo de Inteligencia Artificial o mediante el endpoint correspondiente del backend. La política definida consiste en reentrenar después de cada ETL completo o cuando ingresen nuevos meses con producción confirmada. Durante este proceso, el sistema reconstruye la serie mensual, separa entrenamiento y validación, evalúa los algoritmos candidatos, selecciona el modelo con menor error, genera una versión de modelo, guarda el historial de entrenamiento e invalida la caché del dashboard IA para mostrar los resultados actualizados.

Para garantizar trazabilidad, cada corrida de entrenamiento queda registrada en la base de datos, incluyendo fecha de inicio, fecha de finalización, estado, algoritmo promovido, versión activa, razón de selección, usuario que ejecutó el proceso, dataset utilizado, filas de entrenamiento, filas de validación y métricas obtenidas por algoritmo. Esta característica permite auditar el comportamiento del módulo de Inteligencia Artificial y proporciona evidencia técnica sobre la forma en que el sistema selecciona y actualiza sus modelos predictivos.

Limitaciones y mejoras futuras del módulo IA

El módulo presenta algunas limitaciones propias del estado actual de implementación. La predicción principal trabaja con una serie mensual agregada, por lo que su precisión depende de la cantidad de meses disponibles y de la completitud de los datos históricos. Además, las recomendaciones editoriales se basan en reglas explicables y no todavía en un modelo avanzado de ranking. De igual manera, el modelo de riesgo de demora en workflow queda planteado como una mejora futura. No obstante, estas limitaciones no reducen el aporte del módulo, ya que el sistema implementa una base funcional de análisis predictivo y recomendación institucional, con trazabilidad, comparación de algoritmos, validación temporal e integración directa con el entorno BI.

En conjunto, el módulo de Inteligencia Artificial permite pasar de una reportería descriptiva hacia una analítica predictiva y prescriptiva inicial. Su integración con el Data Warehouse y los dashboards institucionales aporta valor al sistema, debido a que permite anticipar tendencias de producción científica, generar recomendaciones

comprensibles y apoyar decisiones estratégicas dentro de la DIDE-UTA. De esta manera, el componente IA contribuye directamente al cumplimiento del objetivo de desarrollar un sistema de reportería y visualización de datos con funciones de recomendación y predicción basadas en técnicas de Business Intelligence e Inteligencia Artificial.

3.4 Pruebas del sistema

Las pruebas del sistema fueron realizadas durante y después del ciclo de desarrollo con el propósito de validar el funcionamiento de los módulos implementados, verificar la integración entre componentes y comprobar la estabilidad de los flujos críticos de la plataforma. Estas pruebas se enfocaron principalmente en los procesos de registro de artículos científicos, carga masiva de información, validación de datos, reportería institucional, exportación de reportes, ejecución de procesos ETL y funcionamiento general del sistema. Asimismo, permitieron verificar que los módulos desarrollados respondieran a los requerimientos funcionales definidos previamente y que la información procesada mantuviera consistencia antes de alimentar el entorno analítico y los procesos de Business Intelligence.

La estrategia de pruebas se apoyó en evidencias operativas, scripts de validación, datasets de prueba y casos documentados. De acuerdo con el reporte de pruebas prioritarias, la validación se concentró en pruebas de integración, pruebas smoke, pruebas funcionales de carga, pruebas con errores controlados y pruebas de reportería y exportación. Si bien dentro del plan de aseguramiento de calidad se consideraron pruebas unitarias como parte de las prácticas XP, en la versión documentada del sistema no se integró un proyecto formal de pruebas unitarias con herramientas como xUnit, NUnit o MSTest. Por esta razón, la evaluación se centró en pruebas operativas y de integración aplicadas sobre flujos reales del sistema.

Los documentos completos de pruebas, incluyendo casos detallados, scripts utilizados, datasets, resultados esperados y evidencias de ejecución, se presentan en los anexos correspondientes. En el cuerpo principal de la tesis se incluyen únicamente las pruebas más relevantes para demostrar confiabilidad funcional, integración entre módulos y soporte operativo de la solución tecnológica desarrollada que puede ser evidenciada en Tabla 32.

3.4.1 Estrategia de pruebas

La estrategia de pruebas fue definida considerando los módulos críticos del sistema y su relación con los objetivos de la investigación. Se priorizaron los procesos que intervienen directamente en la gestión de información científica, la generación de reportería institucional y la preparación de datos para análisis mediante BI e IA. Por ello, las pruebas se enfocaron en validar que el sistema permita registrar artículos, procesar cargas masivas, controlar errores, actualizar el modelo analítico, generar dashboards y exportar reportes institucionales.

Las pruebas aplicadas permitieron revisar diferentes niveles del sistema. Las pruebas de integración validaron la comunicación entre API, base de datos, staging, reportería y exportación. Las pruebas smoke permitieron comprobar rápidamente que los servicios principales se encontraran operativos después de cambios o antes de una demostración. Las pruebas de carga funcional permitieron evaluar el comportamiento del sistema con datasets pequeños, medianos y extensos. Finalmente, las pruebas con errores controlados permitieron verificar que el sistema rechace información inválida antes de consolidarla en el modelo transaccional

Tabla 32. Tipos de pruebas aplicadas al sistema.

Tipo de prueba	Alcance	Evidencia generada
Pruebas smoke / end-to-end parcial	Validación rápida de backend, autenticación, reportería, exportación y ETL	Scripts PowerShell y archivos generados
Pruebas de integración de registro	Registro individual de artículos, participantes, revistas y métricas	Scripts contra API y base de datos
Pruebas de carga masiva	Subida, validación, staging y procesamiento de lotes	Datasets CSV/XLSX y scripts de carga
Pruebas con errores controlados	Validación de campos obligatorios, DOI duplicado, FK inválida y datos incompletos	Datasets con errores esperados
Pruebas de reportería y exportación	Dashboards, filtros, PDF, Excel y actualización posterior al ETL	Archivos PDF/Excel y resultados de smoke
Pruebas de carga funcional	Evaluación con lotes pequeños, medianos y extensos hasta 1000 registros	Datasets de volumen y tiempos de procesamiento

La aplicación de estos tipos de prueba permitió revisar el comportamiento del sistema desde diferentes perspectivas. De esta manera, no solo se verificó que los módulos funcionaran individualmente, sino también que pudieran integrarse dentro del flujo general de la plataforma.

3.4.2 Pruebas de carga masiva e integridad de datos

Las pruebas de carga masiva fueron desarrolladas para validar el comportamiento del sistema durante la importación, validación y consolidación de información científica. Estas pruebas permitieron comprobar que el modelo transaccional pueda registrar artículos con diferentes niveles de complejidad, desde registros básicos hasta registros con catálogos completos, campos dinámicos, autores, coautores e información proveniente de integración externa. Además, se validaron escenarios con errores controlados para comprobar que el sistema pueda detectar inconsistencias antes de consolidar la información dentro del modelo OLTP, se puede evidenciar en Tabla 33.

Los casos de prueba documentados permitieron verificar que el modelo soporte registros básicos sin depender obligatoriamente de catálogos adicionales, participantes o campos dinámicos. También se obtuvo una percepción favorable sobre la integridad referencial mediante artículos asociados a revistas, periodos académicos, estados de publicación, líneas de investigación y clasificación OCDE. De igual manera, se validó que los campos dinámicos puedan almacenarse sin modificar la estructura física de la tabla principal de artículos, demostrando la extensibilidad del modelo basado en metadatos.

Tabla 33. Casos principales de pruebas de carga masiva e integridad.

Caso	Descripción	Resultado
CI-LOAD-01	Artículo con solo campos básicos	Aprobado
CI-LOAD-02	Artículo con catálogos completos	Aprobado
CI-LOAD-03	Artículo con campos dinámicos	Aprobado
CI-LOAD-04	Artículo con un autor	Aprobado
CI-LOAD-05	Artículo con varios autores/coautores ordenados	Aprobado
CI-LOAD-06	Artículo con datos provenientes de integración externa	Aprobado
CI-LOAD-07	Carga masiva con registros válidos	Aprobado
CI-LOAD-08	Carga masiva con errores: DOI duplicado	Aprobado
CI-LOAD-09	Carga masiva con errores: FK inválida	Aprobado

CI-LOAD-10	Carga masiva con errores: campo requerido vacío	Aprobado
------------	---	----------

Los resultados obtenidos evidencian que el sistema permitió registrar información científica en escenarios básicos y completos, manteniendo consistencia en los datos almacenados. Asimismo, las pruebas con errores controlados permitieron comprobar que los registros inválidos no se consolidan automáticamente en el modelo definitivo, lo cual fortalece la calidad de los datos antes de su uso en reportería, ETL y análisis institucional.

3.4.3 Pruebas de integración del registro individual

Las pruebas de integración del registro individual tuvieron como objetivo validar que el endpoint de registro permita crear artículos científicos junto con sus elementos asociados, como revista, métricas editoriales, participantes y campos base. Este tipo de prueba permitió comprobar la comunicación entre el frontend, backend, API, lógica de negocio y persistencia en la base de datos. Además, permitió verificar que el sistema pueda resolver o crear registros relacionados, como revistas o métricas, sin generar inconsistencias en el modelo transaccional.

De acuerdo con el reporte de pruebas prioritarias, estas pruebas consideraron escenarios de registro mínimo válido, creación repetida con DOI único, persistencia de revista, registro de métricas editoriales, medición de tiempos y control de fallos devueltos por la API, ver Tabla 34. Esto permitió validar tanto la funcionalidad principal de registro como el comportamiento del sistema frente a múltiples ejecuciones controladas.

Tabla 34. Pruebas de integración del registro individual.

Caso	Validación realizada	Resultado esperado
REG-01	Registro mínimo válido	Se crea artículo y participante principal
REG-02	Registro repetido por lote de prueba	Se crean múltiples artículos con DOI único
REG-03	Persistencia de revista	La revista se crea o se resuelve correctamente
REG-04	Persistencia de métrica de revista	Se registra SJR o cuartil cuando aplica

REG-05	Medición de tiempos	Se obtiene promedio, mínimo y máximo por registro
REG-06	Control de fallos	Se reportan errores devueltos por la API

Estas pruebas permitieron comprobar que el flujo de registro individual funciona de manera integrada y que la información registrada puede servir posteriormente como fuente para procesos de workflow, reportería y ETL.

3.4.4 Pruebas smoke y end-to-end parcial

Las pruebas smoke fueron utilizadas para verificar rápidamente que los componentes principales del sistema se encontraran operativos. Estas pruebas permitieron validar el estado general del backend, autenticación, recuperación de información del usuario, salud de reportería, consulta de dashboard institucional, analítica de autores, filtros, generación de archivos PDF y Excel, así como la ejecución opcional del proceso ETL. Su finalidad fue comprobar que los flujos esenciales del sistema funcionaran después de cambios o antes de una demostración funcional.

El reporte de pruebas prioritarias identifica estas pruebas como pruebas de integración, smoke y end-to-end parcial, debido a que validan varios componentes trabajando en conjunto, ver Tabla 35. Esta cobertura resulta importante para la tesis porque demuestra que autenticación, reportería, Data Warehouse, exportación y consultas principales pueden operar dentro de un mismo flujo funcional.

Tabla 35. Pruebas smoke aplicadas al sistema.

Caso	Validación	Resultado esperado
SMK-01	Ping del backend	El endpoint responde correctamente
SMK-02	Login administrador	Se obtiene token de acceso válido
SMK-03	Usuario actual	Se recuperan datos y roles del usuario autenticado
SMK-04	Salud de reportería	El Data Warehouse responde y muestra estado de conexión
SMK-05	Dashboard institucional	Se obtiene resumen de artículos, indicadores y facultades
SMK-06	Analítica de autores	Se validan autores, publicaciones y coautorías
SMK-07	Filtro por mes	El dashboard filtrado no devuelve más registros que la vista general

SMK-08	PDF institucional	Se genera archivo PDF no vacío
SMK-09	Excel institucional	Se genera archivo Excel no vacío
SMK-10	ETL opcional	Se ejecuta refresco ETL y se vuelve a consultar el dashboard

Los resultados esperados de estas pruebas permiten comprobar que los módulos principales del sistema no funcionan de manera aislada, sino integrados dentro de un flujo operativo y analítico. Esto es especialmente relevante para validar la arquitectura propuesta, donde la información fluye desde el modelo OLTP hacia el entorno OLAP y los componentes de reportería.

3.4.5 Pruebas de reportería, ETL y exportación

Las pruebas de reportería y exportación tuvieron como finalidad validar el funcionamiento de los componentes relacionados con Business Intelligence. Estas pruebas verificaron la conexión con el Data Warehouse, la carga de KPIs, la consulta de dashboards, la aplicación de filtros, la generación de reportes y la exportación de información en formatos PDF y Excel. Además, se consideró la ejecución del proceso ETL para comprobar que la reportería pueda actualizarse después de procesar nuevos datos, ver Tabla 36.

La validación de reportería permitió revisar que los indicadores institucionales se generen a partir de información consolidada y que los reportes exportables contengan datos válidos. También se verificó que los filtros aplicados sobre el dashboard no incrementen incorrectamente el universo de registros, sino que reduzcan o mantengan el conjunto de datos consultado. Esto permite demostrar consistencia en el comportamiento de los filtros y confiabilidad en los resultados presentados al usuario.

Tabla 36. Pruebas de reportería y exportación.

Caso	Validación	Resultado esperado
REP-01	Salud del Data Warehouse	Conexión activa con base analítica
REP-02	Dashboard institucional	KPIs y conteos cargan correctamente
REP-03	Filtro por mes	El filtro reduce o mantiene el universo de datos
REP-04	Analítica de autores	Autores y publicaciones no se truncan
REP-05	Exportación PDF	Archivo generado con contenido
REP-06	Exportación Excel	Archivo generado con contenido
REP-07	ETL completo	La reportería se actualiza después del procesamiento

Estas pruebas se relacionan directamente con el objetivo de la tesis, ya que permiten validar que el sistema no solo almacene información científica, sino que también la transforme en indicadores, reportes y visualizaciones útiles para la toma de decisiones institucionales.

3.4.6 Pruebas de carga funcional y rendimiento operativo

Las pruebas de carga funcional fueron aplicadas para evaluar el comportamiento del sistema frente a diferentes volúmenes de información. Para ello, se utilizaron datasets pequeños, medianos y extensos, incluyendo archivos con registros mínimos válidos, registros completos, catálogos, revistas, errores controlados y lotes extensos de hasta 1000 registros. Estas pruebas permitieron analizar tiempos de carga, validación y procesamiento, así como cantidad de filas válidas, filas con error, artículos creados, participantes creados, revistas resueltas y posibles fallos o timeouts.

Tabla 37. Datasets utilizados en pruebas de carga funcional.

Dataset	Registros	Propósito
Carga mínima válida	20	Validar carga mínima correcta
Carga completa válida	25	Validar campos completos, catálogos y datos extendidos
Carga con catálogos válidos	18	Validar resolución de catálogos
Carga de revistas válidas	15	Validar revistas y venues
Carga con errores controlados	16	Validar errores esperados
Carga completa extensa A	1000	Validar volumen alto prudente
Carga completa extensa B	1000	Repetir volumen alto con dataset alternativo

En Tabla 37 se puede evidenciar la utilización de los datasets que permitió comprobar el comportamiento del sistema en escenarios de volumen controlado. Además, permitió validar que la carga masiva pueda operar con registros válidos y con registros erróneos, manteniendo mecanismos de control antes de la consolidación definitiva de la información. Estas pruebas son importantes porque evidencian la capacidad del sistema para reducir procesos manuales y soportar escenarios de importación institucional.

3.4.7 Consideraciones sobre pruebas unitarias

Dentro del enfoque Scrum + XP, las pruebas unitarias fueron consideradas como una práctica importante para validar componentes críticos del sistema, especialmente validadores, reglas de negocio, generación de reportes, predicción y recomendación. Sin embargo, en la versión documentada del sistema no se integró un proyecto formal de pruebas unitarias con herramientas como xUnit, NUnit o MSTest. Por esta razón, las pruebas unitarias se mantienen como una mejora técnica recomendada para fortalecer la automatización del aseguramiento de calidad.

Entre las pruebas unitarias recomendadas se encuentran validaciones de campos dinámicos, detección de DOI duplicado, control de participantes, resolución de revistas, validación de filas inválidas en staging y reglas de workflow. Estas pruebas permitirían mejorar la cobertura técnica del sistema y facilitar futuras fases de mantenimiento.

3.4.8 Resultados generales de las pruebas

Los resultados generales de las pruebas evidenciaron que los flujos principales del sistema fueron validados satisfactoriamente en escenarios funcionales y de integración. Las pruebas de carga masiva demostraron que el sistema puede registrar artículos con datos básicos, catálogos completos, campos dinámicos, autores, coautores e información proveniente de integración externa. Asimismo, los casos con errores controlados permitieron verificar que el sistema identifica inconsistencias como DOI duplicado, claves foráneas inválidas y campos requeridos vacíos antes de consolidar la información definitiva.

Las pruebas smoke permitieron comprobar el funcionamiento general del backend, autenticación, reportería, Data Warehouse, exportación PDF/Excel y actualización mediante ETL. Por su parte, las pruebas de carga funcional con datasets de diferentes tamaños permitieron validar la estabilidad operativa del sistema frente a escenarios de importación controlada. En conjunto, estas pruebas aportan evidencia sobre la confiabilidad funcional de la plataforma y su capacidad para soportar los procesos principales de gestión, análisis y reportería institucional.

Tabla 38. Resumen general de resultados de pruebas.

Grupo de prueba	Resultado general	Observación
Carga masiva e integridad	Aprobado	Se validaron registros básicos, completos, dinámicos y con errores controlados
Registro individual	Validado	Se obtuvo una persistencia favorable de artículos, participantes, revistas y métricas
Smoke / end-to-end parcial	Validado	Se verificaron backend, autenticación, reportería, exportación y ETL
Reportería y exportación	Validado	Se generaron dashboards, filtros, PDF y Excel
Carga funcional	Validado	Se utilizaron datasets pequeños, medianos y extensos
Pruebas unitarias automatizadas	Pendiente como mejora	No existe proyecto formal de unit tests integrado en la versión documentada

En conclusión, las pruebas realizadas permitieron validar los componentes fundamentales del sistema y comprobar que los principales flujos institucionales funcionan de forma integrada. Además, evidencian que la plataforma mantiene mecanismos de control de calidad de datos antes de alimentar los procesos de reportería, ETL y análisis institucional. Como mejora técnica futura, se recomienda implementar un proyecto formal de pruebas unitarias automatizadas para ampliar la cobertura sobre validadores, reglas de negocio, servicios críticos y componentes relacionados con Inteligencia Artificial.

3.5 Resultados obtenidos

Los resultados obtenidos se presentan en función de los objetivos específicos definidos para la investigación. Esta organización permite evidenciar la relación entre el análisis inicial de la problemática, el desarrollo del sistema propuesto y la evaluación del desempeño alcanzado mediante pruebas funcionales, integración de módulos y mediciones de reportería. En este sentido, los resultados se estructuran considerando tres aspectos principales: el análisis del estado actual de la gestión informativa en la DIDE-UTA, el desarrollo del sistema de reportería y visualización con funciones de recomendación y predicción, y la evaluación del desempeño del sistema mediante indicadores técnicos y operativos.

3.5.1 Resultados relacionados con el análisis del estado actual de la gestión informativa

El primer objetivo específico estuvo orientado a analizar el estado actual de la gestión informativa en la DIDE, con el propósito de identificar las principales limitaciones en la generación, análisis y presentación de los datos. Como resultado de este proceso, se identificaron problemáticas relacionadas con la gestión manual de información científica, dispersión de datos en múltiples archivos, ausencia de reportería centralizada, limitada trazabilidad de procesos, generación manual de reportes y falta de integración con fuentes externas de información científica.

El levantamiento de información permitió evidenciar que los procesos de reportería dependían principalmente de archivos Excel, filtros manuales, consolidación de datos disponibles en el momento y revisión individual de registros. Esta situación ocasionaba demoras en la generación de reportes, dificultad para consultar información histórica y mayor dependencia de tareas manuales. Asimismo, el análisis permitió identificar la necesidad de contar con una plataforma centralizada que integre registro, validación, carga masiva, workflow, reportería analítica, Business Intelligence e Inteligencia Artificial.

Tabla 39. Resultados asociados al análisis del contexto institucional.

Resultado identificado	Evidencia generada	Aporte al proyecto
Gestión manual de información científica	Entrevistas y análisis del proceso actual	Justificó la necesidad de centralización
Dispersión de datos en archivos y formatos	Problemáticas institucionales identificadas	Evidenció la necesidad de una base integrada
Generación manual de reportes	Encuesta y levantamiento de información	Sustentó la automatización de reportería
Limitada trazabilidad institucional	Análisis de necesidades operativas	Justificó workflow y auditoría
Necesidad de dashboards y KPIs	Requerimientos funcionales y no funcionales	Orientó el diseño BI
Interés en recomendación y predicción	Requerimientos de IA	Sustentó el módulo de Inteligencia Artificial

En conjunto, los resultados del análisis permitieron definir los requerimientos funcionales y no funcionales del sistema. Estos requerimientos sirvieron como base

para estructurar los módulos de gestión de artículos científicos, carga masiva, workflow institucional, integración externa, reportería, procesos ETL, Business Intelligence e Inteligencia Artificial. Por tanto, fue posible identificar las limitaciones principales del proceso actual y convertirlas en necesidades funcionales concretas para el desarrollo de la propuesta tecnológica.

3.5.2 Resultados relacionados con el desarrollo del sistema BI e IA

El segundo objetivo específico estuvo orientado a desarrollar un sistema de reportería y visualización interactiva con funciones de recomendación y predicción, basado en técnicas de Business Intelligence e Inteligencia Artificial. Como resultado, se implementó una plataforma modular que integra componentes operativos, analíticos y predictivos, alineados con los requerimientos identificados durante la etapa de análisis.

El sistema desarrollado permitió centralizar la gestión de información científica institucional mediante un modelo transaccional OLTP extensible, orientado al registro, validación, administración y trazabilidad de artículos científicos. Además, se implementaron procesos de carga masiva, validación por lotes, integración con APIs externas, workflow institucional y administración de usuarios mediante roles y permisos. Estos módulos permitieron estructurar la información operativa necesaria para alimentar los procesos posteriores de reportería y análisis.

Desde el enfoque analítico, se diseñó e implementó un modelo OLAP basado en una constelación dimensional, compuesto por tablas de hechos y dimensiones compartidas. Esta estructura permitió consolidar información histórica relacionada con artículos científicos, autores, revistas, indexaciones, líneas de investigación, métricas editoriales y estados institucionales. A partir de este modelo se desarrollaron procesos ETL, dashboards, KPIs, reportes institucionales y mecanismos de exportación de información.

Tabla 40. Componentes desarrollados como resultado del sistema propuesto.

Componente desarrollado	Resultado obtenido	Relación con el objetivo
Modelo OLTP	Registro operativo de artículos, autores, revistas, catálogos, workflow y trazabilidad	Centralización de información científica

Carga masiva	Importación, validación y procesamiento de registros por lotes	Reducción de ingreso manual
Workflow institucional	Revisión, aprobación, devolución y seguimiento de registros	Control del proceso institucional
Integración externa	Consulta y normalización de información científica desde APIs	Enriquecimiento de datos
Modelo OLAP	Constelación dimensional para análisis histórico	Base para BI y KPIs
Procesos ETL	Extracción, transformación y carga hacia el Data Warehouse	Consolidación analítica
Reportería BI	Dashboards, filtros, gráficos, KPIs y reportes exportables	Visualización interactiva
Inteligencia Artificial	Predicción de producción científica y recomendaciones explicables	Apoyo predictivo y prescriptivo inicial

El módulo de Inteligencia Artificial permitió incorporar funciones de predicción y recomendación institucional. La predicción principal estima la producción científica esperada para los próximos seis meses a partir de una serie temporal mensual de artículos. Además, el sistema genera predicciones segmentadas por facultad y línea de investigación, así como recomendaciones relacionadas con calidad de datos, trazabilidad de autores, indexación, acceso abierto, colaboración científica y uso de predicciones como alertas tempranas.

Tabla 41. Evidencia funcional del módulo de Inteligencia Artificial.

Componente evaluado	Evidencia generada	Resultado mostrado por el sistema
Dataset utilizado	Serie mensual de producción científica consolidada en el Data Warehouse	Mes, año y total de artículos por periodo
Modelos comparados	Persistencia, promedio móvil de tres meses, regresión lineal temporal y ForecastBySsa	Comparación de candidatos para predicción
Métricas calculadas	MAE, RMSE y MAPE	Evaluación del error por modelo
Selección del modelo	Menor MAE y menor RMSE en caso de empate	Modelo promovido como mejor alternativa
Predicción mensual	Horizonte de seis meses	Valor esperado, límite inferior, límite superior y confianza

Recomendaciones generadas	Reglas explicables por indicadores	Calidad de datos, ORCID, filiación, indexación, cuartiles y tendencias
Historial de entrenamiento	Registro de corrida IA	Fecha, algoritmo, métricas, dataset y usuario ejecutor

Desde el punto de vista técnico, el módulo IA compara modelos como persistencia del último valor, promedio móvil de tres meses, regresión lineal temporal y ForecastBySsa de ML.NET. La selección del mejor modelo se realiza mediante validación temporal y métricas como MAE, RMSE y MAPE. Asimismo, el sistema incorpora un diagnóstico de preparación para IA mediante indicadores ponderados como volumen de datos, trazabilidad de autores, cobertura de indexación, calidad de carga y señales de workflow.

Para evidenciar el funcionamiento real del módulo IA, se verificó que el sistema generara predicciones a partir de la serie mensual consolidada en el Data Warehouse. Durante este proceso se compararon modelos candidatos mediante métricas de error como MAE, RMSE y MAPE, seleccionando como mejor alternativa el modelo con menor MAE y, en caso de empate, el menor RMSE. Además, el sistema generó recomendaciones explicables relacionadas con calidad de datos, trazabilidad autoral, cobertura de indexación, acceso abierto, cuartiles y comportamiento de la producción científica. Esta evidencia demuestra que el módulo IA no se limita a una descripción conceptual, sino que ejecuta procesos de predicción, comparación de modelos, selección del algoritmo y generación de recomendaciones dentro del entorno BI.

3.5.3 Resultados relacionados con la evaluación del desempeño del sistema

El tercer objetivo específico estuvo orientado a evaluar el desempeño del sistema desarrollado mediante métricas relacionadas con reducción de tiempos de reportería, adopción institucional y utilidad percibida por los usuarios. Para su cumplimiento, se consideraron los resultados obtenidos durante las pruebas funcionales, pruebas de integración, pruebas de carga masiva, pruebas smoke, pruebas de reportería, exportación de reportes, mediciones de tiempos registradas por el sistema y evaluación post-implantación basada en dimensiones CSUQ.

Las pruebas aplicadas permitieron validar los principales flujos funcionales de la plataforma. Entre ellos se incluyeron el registro individual de artículos, carga masiva

mediante staging, validación de errores controlados, procesamiento de lotes, reportería institucional, exportación en PDF y Excel, ejecución de procesos ETL y funcionamiento general del sistema. Estas pruebas evidenciaron que los módulos operativos y analíticos trabajan de forma integrada, permitiendo que la información registrada pueda ser procesada, consolidada y utilizada en reportes institucionales.

Tabla 42. Resultados generales de pruebas aplicadas al sistema.

Grupo de prueba	Resultado general	Interpretación
Carga masiva e integridad	Aprobado	Se validaron registros básicos, completos, dinámicos y con errores controlados
Registro individual	Validado	Se comprobó persistencia de artículos, participantes, revistas y métricas
Smoke / end-to-end parcial	Validado	Se verificaron backend, autenticación, reportería, exportación y ETL
Reportería y exportación	Validado	Se generaron dashboards, filtros, PDF y Excel
Carga funcional	Validado	Se utilizaron datasets pequeños, medianos y extensos
Pruebas unitarias automatizadas	Pendiente como mejora técnica	No existe proyecto formal de unit tests integrado en la versión documentada

En cuanto a la reducción de tiempos de reportería, el sistema registró 80 operaciones en la tabla ReportingPerformanceMetrics, correspondientes a carga del dashboard institucional, analítica de autores, ejecución de ETL y generación de reportes PDF/Excel. Todas las operaciones registradas finalizaron exitosamente, obteniendo una tasa de éxito del 100 % en la muestra observada.

Como línea base previa al sistema, se consideró el proceso manual descrito por los usuarios de la DIDE-UTA, en el cual la generación de reportes podía tomar aproximadamente medio día de trabajo o un tiempo cercano, dependiendo de la disponibilidad de datos, consolidación de archivos Excel, aplicación de filtros, revisión de información histórica y preparación del reporte final. Debido a que antes del desarrollo de la plataforma no existía un sistema que registrara automáticamente los tiempos de reportería, esta línea base fue tratada como una estimación operativa del proceso manual.

Tabla 41. Resultados generales de tiempos de reportería.

Indicador	Valor observado	Interpretación
Operaciones medidas	80	Muestra inicial de operaciones registradas por el sistema
Operaciones exitosas	80	No se observaron fallos en las operaciones medidas
Operaciones fallidas	0	La reportería se mantuvo estable durante la muestra
Tasa de éxito	100 %	Indicador favorable de estabilidad funcional
Tiempo promedio real del sistema	10,53 segundos	Tiempo promedio adecuado para consultas institucionales
Tiempo real acumulado	14,04 minutos	Las 80 operaciones se resolvieron en menos de 15 minutos acumulados
Tiempo manual estimado acumulado	68,17 horas	Línea base estimada para tareas manuales equivalentes
Ahorro estimado acumulado	67,93 horas	Evidencia del potencial de reducción de tiempos
Reducción estimada	99,66 %	Resultado favorable frente a la línea base manual estimada
Elementos consultados acumulados	95 629	Muestra con volumen relevante de información consultada

Los resultados obtenidos muestran que el sistema permitió automatizar actividades que anteriormente requerían consolidar información de manera manual, aplicar filtros en hojas de cálculo, revisar datos disponibles en diferentes archivos y generar reportes de forma no centralizada. En la muestra registrada, el tiempo promedio general de respuesta fue de 10,53 segundos, mientras que el tiempo acumulado de las 80 operaciones fue de 14,04 minutos. Al compararlo con la línea base manual estimada de 68,17 horas, se obtuvo un ahorro estimado de 67,93 horas, equivalente a una reducción aproximada del 99,66 %. Este resultado evidencia una mejora significativa en el tiempo de reportería institucional, aunque debe interpretarse como una estimación frente al proceso manual previo y no como una medición absoluta cronometrada antes del sistema.

Tabla 43. Base de medición para la reducción estimada de tiempos de reportería.

Elemento evaluado	Fuente	Cantidad de mediciones	Resultado
Operaciones reales del sistema	Tabla ReportingPerformanceMetrics	80 operaciones	14,04 minutos acumulados
Operaciones exitosas	Registro interno del sistema	80 operaciones	100 % de éxito
Operaciones fallidas	Registro interno del sistema	80 operaciones	0 fallos
Tiempo promedio real del sistema	Métricas registradas por operación	80 operaciones	10,53 segundos
Línea base manual equivalente	Estimación operativa del responsable de reportería	80 operaciones equivalentes	68,17 horas estimadas
Ahorro estimado acumulado	Comparación sistema vs. línea base manual	80 operaciones equivalentes	67,93 horas
Reducción estimada	Fórmula comparativa	80 operaciones equivalentes	99,66 %

Tabla 44. Tiempos de reportería por tipo de operación.

Operación	Ejecuciones	Promedio	Mediana	P90	Máximo	Interpretación
Dashboard institucional	32	16,37 s	0,44 s	96,01 s	96,40 s	Promedio aceptable, con picos altos en consultas pesadas
Panel autores/coautoría	29	8,94 s	0,38 s	15,17 s	99,16 s	Comportamiento promedio favorable, con un pico aislado
ETL reportería	12	2,63 s	2,71 s	3,41 s	3,43 s	Tiempo estable y bajo en la muestra
Exportación Excel	3	6,71 s	7,07 s	7,12 s	7,14 s	Generación consistente menor a 10 segundos
Exportación PDF	3	2,12 s	0,55 s	4,41 s	5,38 s	Generación rápida en la muestra

PDF compuesto	1	0,83 s	0,82 s	0,82 s	0,83 s	Evidencia puntual, requiere más mediciones
---------------	---	--------	--------	--------	--------	--

La matriz por tipo de operación permite interpretar los resultados de forma equilibrada. Aunque el promedio general fue favorable, se identificaron picos de carga en el dashboard institucional y en el panel de autores/coautoría, con valores máximos cercanos a 96 y 99 segundos. Estos casos corresponden a consultas más pesadas y deben considerarse como oportunidades de optimización para una fase de producción con mayor volumen de datos. No obstante, estos tiempos continúan siendo menores que el tiempo manual estimado para la generación de reportes institucionales.

Como complemento a las métricas técnicas de desempeño, se aplicó una entrevista estructurada post-implantación basada en dimensiones CSUQ, adaptada al contexto del sistema desarrollado. El instrumento fue aplicado en línea mediante Google Forms a cinco participantes, en coherencia con la población definida para la investigación, utilizando una escala tipo Likert de 1 a 5. Esta entrevista permitió evaluar la utilidad del sistema, la calidad de la información y reportería, la facilidad de uso e interfaz, las funciones de Inteligencia Artificial, la satisfacción y aceptación inicial.

Los resultados obtenidos evidenciaron un promedio general de 4,59 sobre 5 y una aceptación global del 100 %, considerando como respuestas favorables aquellas ubicadas en los niveles 4 y 5 de la escala. De las 100 respuestas procesadas, 41 correspondieron al valor 4 y 59 al valor 5, sin registrarse valoraciones bajas o neutras. Además, el Alfa de Cronbach fue de 0,973, lo que refleja una alta consistencia interna del instrumento, interpretada como referencia descriptiva debido al tamaño reducido de la población.

Tabla 45. Resultados de la entrevista estructurada post-implantación basada en dimensiones CSUQ.

Dimensión evaluada	Promedio	Aceptación	Interpretación
Utilidad del sistema	4,60	100 %	Percepción favorable sobre el aporte funcional del sistema.
Calidad de la información y reportería	4,60	100 %	Reportes e indicadores valorados como claros y útiles.
Facilidad de uso e interfaz	4,55	100 %	Navegación e interfaz valoradas positivamente.

Funciones de Inteligencia Artificial	4,55	100 %	Predicción y recomendación percibidas como apoyo útil.
Satisfacción y aceptación inicial	4,65	100 %	Mayor valoración general sobre satisfacción y condiciones favorables para adopción progresiva.

Al analizar los resultados por dimensión, se observa que la valoración más alta corresponde a satisfacción y aceptación inicial, con un promedio de 4,65. Las dimensiones de utilidad del sistema y calidad de la información y reportería obtuvieron un promedio de 4,60, mientras que facilidad de uso e interfaz y funciones de Inteligencia Artificial alcanzaron 4,55. Estos resultados demuestran una percepción favorable hacia el sistema implementado y complementan la evaluación del tercer objetivo específico, al evidenciar no solo reducción de tiempos de reportería, sino también utilidad percibida, satisfacción y aceptación inicial.

3.5.4 Métricas del sistema

Las métricas del sistema permitieron consolidar los resultados técnicos y perceptuales obtenidos durante las pruebas, mediciones de reportería y evaluación post-implantación. Estas métricas fueron agrupadas en indicadores de estabilidad, rendimiento, reportería, carga funcional, exportación, modelo analítico, Inteligencia Artificial y utilidad percibida. Su finalidad fue proporcionar una visión general del comportamiento del sistema frente a los principales procesos institucionales definidos en la propuesta, así como complementar la evaluación técnica con la percepción de los participantes respecto al uso de la plataforma.

Tabla 46. Métricas generales del sistema desarrollado.

Métrica	Resultado	Interpretación
Casos principales de carga masiva	10 casos aprobados	Validación favorable de integridad y control de datos
Operaciones de reportería medidas	80	Muestra inicial de desempeño operativo
Tasa de éxito en reportería	100 %	No se observaron fallos en las operaciones registradas
Tiempo promedio general de reportería	10,53 segundos	Tiempo adecuado para consultas institucionales
Tiempo acumulado del sistema	14,04 minutos	Procesamiento total bajo frente a línea base manual
Reducción estimada de tiempo	99,66 %	Mejora significativa frente al proceso manual estimado
Datasets de carga funcional	Hasta 1000 registros	Validación de carga con volumen alto prudente
Exportación PDF	Validada	Generación de reportes con contenido

Exportación Excel	Validada	Generación de archivos exportables con contenido
Modelo OLAP	Constelación dimensional implementada	Base analítica para BI, KPIs e IA
Módulo IA	Predicción a seis meses y recomendaciones explicables	Soporte para análisis predictivo y prescriptivo inicial
Participantes de la entrevista estructurada post-implantación	5	Aplicación en línea mediante Google Forms, coherente con la población definida
Respuestas procesadas en la entrevista post-implantación	100	20 ítems aplicados a 5 participantes
Promedio general post-implantación	4,59 / 5	Percepción favorable del sistema
Aceptación general post-implantación	100 %	Todas las respuestas se ubicaron en valores 4 y 5
Alfa de Cronbach del instrumento post-implantación	0,973	Alta consistencia interna referencial del instrumento

Los resultados de estas métricas muestran que el sistema alcanzó un comportamiento favorable en las pruebas aplicadas, en la medición inicial de reportería y en la evaluación post-implantación. La tasa de éxito del 100 % en las operaciones medidas evidencia estabilidad funcional en la muestra observada. La reducción estimada del 99,66 % en el tiempo de reportería muestra el impacto de automatizar actividades que anteriormente dependían de procesos manuales, como la consolidación de información, aplicación de filtros, revisión de datos y generación de reportes institucionales. Asimismo, la validación de exportación PDF/Excel, carga funcional, ETL y reportería confirma que el sistema cumple con los componentes principales definidos en el objetivo general de la investigación.

La incorporación de la entrevista estructurada post-implantación basada en dimensiones CSUQ permitió complementar las métricas técnicas con una valoración perceptual de los participantes. Este instrumento fue aplicado en línea mediante Google Forms a cinco participantes, considerando dimensiones relacionadas con utilidad del sistema, calidad de la información y reportería, facilidad de uso, funciones de Inteligencia Artificial, satisfacción y aceptación inicial. Los resultados evidenciaron un promedio general de 4,59 sobre 5, una aceptación global del 100 % y un Alfa de Cronbach de 0,973, lo que refleja una percepción favorable del sistema y una alta consistencia interna referencial del instrumento aplicado.

Finalmente, los resultados obtenidos permiten afirmar que la plataforma desarrollada constituye una solución tecnológica capaz de mejorar la gestión informativa de la

DIDE-UTA, automatizar procesos de reportería, consolidar información histórica, generar indicadores institucionales y aportar funciones iniciales de recomendación y predicción. Estos resultados respaldan la pertinencia del sistema como herramienta de apoyo para la toma de decisiones estratégicas basadas en datos, integrando evidencia técnica de funcionamiento y evidencia perceptual relacionada con utilidad, facilidad de uso, satisfacción y aceptación inicial.

Tabla 47. Trazabilidad entre objetivos específicos, evidencias y resultados alcanzados.

Objetivo específico	Evidencia principal	Tablas/Figuras asociadas	Resultado alcanzado	Conclusión parcial
Analizar el estado actual de la gestión informativa en la DIDE	Entrevista estructurada, entrevista semiestructurada y análisis mediante el uso de una herramienta informática.	Figuras 6-10, Tablas 6-9	Se identificaron 8 problemáticas y 10 necesidades institucionales	El diagnóstico justificó el desarrollo de una solución BI-IA para centralizar información, automatizar reportería y mejorar trazabilidad.
Desarrollar un sistema de reportería y visualización con recomendación y predicción	Arquitectura Clean Architecture, modelo OLTP, modelo OLAP, ETL, BI e IA	Figuras 32-46, Tablas 10-31	Se implementó una plataforma modular con OLTP extensible, Data Warehouse tipo constelación, dashboards, reportes e IA	El sistema desarrollado integra procesos operativos, analíticos y predictivos alineados con los requerimientos institucionales.
Evaluar el desempeño del sistema desarrollado	Pruebas funcionales, integración, smoke, carga, reportería,	Tablas 32-45	Se registraron 80 operaciones exitosas, 10,53 s promedio, reducción estimada del	La evaluación evidenció estabilidad funcional,

	exportación, métricas de tiempo y entrevista post-implantación		99,66 %, CSUQ 4,59/5 y aceptación inicial del 100 %	mejora inicial en tiempos de reportería y percepción favorable de utilidad.
--	--	--	---	---

3.6 Implantación del sistema

La implantación del sistema se realizó con el propósito de dejar disponible una versión funcional de la plataforma en un entorno institucional controlado, similar al contexto operativo donde será utilizada por la DIDE-UTA. Para ello, se preparó una versión específica orientada al despliegue, incorporando configuraciones relacionadas con contenedores, variables de entorno, servidor web y conexión con las bases de datos requeridas por el sistema. Esta estrategia permitió separar la versión de desarrollo de la versión preparada para ejecución en servidor, facilitando el control de cambios y reduciendo el riesgo de afectar otros servicios institucionales existentes.

3.6.1 Estrategia de implantación

El sistema fue implantado sobre un servidor institucional basado en Rocky Linux, utilizando Docker como plataforma de contenerización. Esta decisión permitió desplegar los componentes principales de la solución de forma aislada, reproducible y administrable, manteniendo independencia entre los servicios del sistema y otros aplicativos previamente instalados en el servidor. La arquitectura de despliegue se organizó mediante dos contenedores principales: uno destinado al backend y otro destinado al frontend.

El backend fue desarrollado en ASP.NET Core y concentra la lógica de negocio, los servicios internos, autenticación, gestión de artículos científicos, carga masiva, workflow, reportería, procesos ETL e Inteligencia Artificial. Este componente se ejecuta como un servicio independiente dentro del entorno Docker y se comunica con las bases de datos institucionales restauradas para el sistema. Por su parte, el frontend

fue desarrollado mediante Blazor WebAssembly y publicado como aplicación web estática, servida mediante Nginx, el cual actúa como servidor web y punto de entrada para los usuarios.

La comunicación entre frontend y backend se configuró internamente dentro de la red de contenedores, de manera que las solicitudes realizadas desde la interfaz web hacia los servicios de la API son redirigidas de forma transparente. Esta organización permite que el usuario acceda al sistema desde una única dirección web institucional, mientras que la comunicación entre los componentes internos se mantiene encapsulada dentro de la arquitectura de despliegue.

Para la persistencia de información se utilizaron bases de datos SQL Server, correspondientes al modelo transaccional y al modelo analítico del sistema. La base transaccional almacena información operativa relacionada con usuarios, artículos científicos, autores, revistas, indexaciones, carga masiva, workflow y trazabilidad. La base analítica almacena la información consolidada para reportería, Business Intelligence, KPIs y procesos vinculados con el módulo de Inteligencia Artificial. Antes de iniciar el sistema, estas bases fueron restauradas y verificadas en el servidor, comprobando la existencia de tablas operativas, esquemas analíticos y componentes necesarios para la ejecución de procesos ETL.

El despliegue se realizó de forma independiente respecto a otros sistemas existentes en el servidor institucional. Para ello, se configuró un stack propio de contenedores, un puerto de exposición separado y variables de entorno específicas para el sistema. Esta decisión permitió evitar interferencias con otros servicios previamente instalados y facilitar la administración individual de la plataforma. Además, se configuraron políticas de reinicio automático para favorecer la disponibilidad del sistema después de reinicios del servidor o recreación de contenedores.

También se incorporaron ajustes específicos en el servidor web para soportar el comportamiento de una aplicación tipo SPA y permitir operaciones de mayor duración, como generación de reportes, ejecución de procesos ETL, carga masiva de información y exportación de datos. De igual manera, se configuraron volúmenes persistentes para conservar archivos generados y recursos utilizados por el sistema, aun cuando los contenedores sean actualizados o recreados.

Tabla 48. Información operativa no sensible del entorno de implantación

Aspecto	Información operativa documentada
Entorno de despliegue	Servidor institucional asignado
Sistema operativo	Rocky Linux 9.5
Acceso remoto	Acceso mediante SSH con usuario institucional autorizado
Tecnología de despliegue	Docker y Docker Compose
Arquitectura de despliegue	Contenedores independientes para frontend y backend
Servicio frontend	Aplicación Blazor WebAssembly publicada mediante Nginx
Servicio backend	API desarrollado en ASP.NET Core / .NET
Comunicación entre servicios	Comunicación interna entre contenedores mediante red Docker
Exposición del frontend	Puerto institucional asignado para acceso HTTP controlado
Puerto interno del frontend	Servicio web publicado internamente mediante Nginx
Puerto interno del backend	API expuesto internamente dentro del entorno de contenedores
Base de datos transaccional	Base de datos OLTP institucional restaurada para el sistema
Base de datos analítica	Base de datos OLAP / Data Warehouse restaurada para reportería
Motor de base de datos	Microsoft SQL Server
Modelo analítico	Data Warehouse extensible con tablas de hechos y dimensiones
Esquemas principales del Data Warehouse	Esquemas orientados a consolidación analítica y procesos ETL
Procedimiento ETL principal	Procedimiento institucional para carga completa hacia el Data Warehouse
Seguridad del sistema	Autenticación mediante JWT y control de acceso por roles
Roles principales	Administrador, autor, revisores institucionales, perfiles de reportería e IA
Administración de permisos	Gestión configurable de permisos por usuario y rol desde el módulo administrativo
Variables de entorno	Configuración mediante archivo de entorno dentro del servidor
Manejo de credenciales	Las credenciales, claves JWT, claves SQL y tokens externos no se encuentran documentan por seguridad
APIs externas contempladas	Scopus, Crossref, OpenAlex y Semantic Scholar
Persistencia de datos	Bases de datos restauradas desde respaldos institucionales
Respaldos	Respaldos de las bases OLTP y OLAP gestionados en el entorno institucional
Logs del sistema	Revisión mediante logs de contenedores backend y frontend
Verificación de contenedores	Validación del estado activo de los servicios desplegados
Validación del frontend	Respuesta HTTP correcta desde el servicio web publicado

Validación de API protegida	Respuesta no autorizada esperada al consultar rutas protegidas sin token
Usuarios de prueba	Usuarios definidos según ambiente institucional y roles de validación
Credenciales de prueba	Administradas por el responsable técnico; no se incluyen en el documento final
Estado operativo	Sistema desplegado, accesible en el entorno institucional y conectado a las bases de datos de prueba

En Tabla 48 se muestra la evidencia de que el sistema fue implantado en un entorno institucional controlado, utilizando contenedores independientes para frontend y backend, servidor web Nginx, bases de datos OLTP y OLAP en SQL Server, autenticación mediante JWT y control de acceso por roles. Además, se verificó el acceso al frontend, la protección de rutas API, la conexión con la base de datos, la disponibilidad de esquemas principales, la ejecución de procesos ETL y el estado activo de los servicios desplegados. Con ello, se demuestra que la solución no quedó limitada a un entorno local de desarrollo, sino que fue publicada en un servidor institucional preparado para validación funcional y operación progresiva.

3.6.2 Verificación del despliegue

La implantación fue validada en dos momentos. Primero, se realizaron pruebas locales de contenerización para comprobar la construcción de imágenes, ejecución de contenedores, carga del frontend, conexión del backend con la base de datos y respuesta de endpoints protegidos. Posteriormente, se validó el despliegue en el servidor institucional, verificando la disponibilidad del entorno Docker, la restauración de bases de datos, la ejecución de contenedores, la respuesta del frontend, la comunicación del servidor web con el backend y el comportamiento esperado de rutas protegidas mediante autenticación.

Durante la verificación se comprobó que el frontend respondiera correctamente desde el servidor y que los endpoints protegidos del backend respondieran con comportamiento esperado cuando no se presentaba un token de autenticación, se puede evidenciar en Tabla 49. Este resultado fue considerado correcto, debido a que dichos servicios requieren autorización mediante credenciales válidas. Asimismo, se verificó

que el backend pudiera conectarse a la base transaccional restaurada y reconocer entidades operativas como artículos, participantes, lotes de importación, workflow, métricas de reportería e historial de entrenamiento de Inteligencia Artificial.

Tabla 49. Verificaciones realizadas durante la implantación del sistema.

Verificación	Resultado obtenido	Interpretación
Construcción de imágenes Docker	Correcta	Los componentes backend y frontend pudieron empaquetarse para despliegue
Ejecución de contenedores	Correcta	Los servicios principales quedaron activos en el servidor
Respuesta del frontend	Correcta	La interfaz web pudo ser servida desde el entorno institucional
Comunicación frontend-backend	Correcta	Las solicitudes hacia la API fueron redirigidas internamente
Conexión con SQL Server	Correcta	El backend accedió a las bases restauradas del sistema
Restauración de base OLTP	Correcta	La base transaccional quedó disponible para operación
Restauración de base OLAP	Correcta	La base analítica quedó disponible para reportería y BI
Verificación de ETL	Correcta	Se confirmó la existencia de componentes requeridos para actualización analítica
Rutas protegidas	Correcta	Los endpoints restringidos respondieron conforme al control de autenticación
Aislamiento respecto a otros sistemas	Correcto	El sistema fue desplegado sin interferir con servicios existentes

Los resultados de estas verificaciones permitieron confirmar que el sistema se encontraba operativo dentro del entorno institucional. Además, evidenciaron que la solución puede ejecutarse de manera contenerizada, conectarse con sus bases de datos, servir la interfaz web, exponer servicios protegidos y mantener independencia respecto a otros sistemas del servidor.

3.6.3 Resultado de la implantación

Como resultado del proceso de implantación, el sistema quedó disponible en el servidor institucional para revisión, validación y uso controlado por parte de los

usuarios involucrados. La plataforma fue desplegada mediante una arquitectura basada en contenedores, compuesta por un servicio backend, un servicio frontend, servidor web Nginx y conexión con bases de datos SQL Server. Esta estructura permitió disponer de una solución funcional, modular y preparada para futuras actualizaciones.

La implantación evidencia la portabilidad del sistema, ya que la solución pudo ser trasladada desde el entorno de desarrollo hacia un servidor institucional mediante una configuración reproducible basada en Docker. También demuestra su capacidad de integración con infraestructura existente, al operar sin afectar otros servicios instalados. De igual manera, la separación entre frontend, backend y bases de datos facilita el mantenimiento, la actualización progresiva y el monitoreo posterior de la plataforma.

En conjunto, la implantación permitió validar que el sistema desarrollado no permanece únicamente como prototipo local, sino que puede ejecutarse en un entorno institucional real. Esto fortalece el cumplimiento del objetivo general de la investigación, al disponer de una plataforma operativa para la gestión, reportería, visualización, recomendación y predicción relacionada con la producción científica de la DIDE-UTA.

3.6.4 Escalabilidad y mantenimiento

La arquitectura de implantación utilizada favorece la escalabilidad y el mantenimiento del sistema debido a la separación de responsabilidades entre sus componentes. El uso de contenedores permite actualizar el backend o frontend de forma independiente, sin modificar directamente el servidor anfitrión. Asimismo, la configuración mediante variables de entorno facilita adaptar el sistema a nuevos entornos, credenciales, puertos o bases de datos sin alterar el código fuente.

Desde el punto de vista operativo, el sistema puede mantenerse mediante actualizaciones controladas del código fuente, reconstrucción de imágenes y reinicio de contenedores. Además, la existencia de bases separadas para el entorno transaccional y analítico permite conservar una organización adecuada entre los procesos diarios de gestión y los procesos de reportería institucional. Esta separación también facilita futuras mejoras relacionadas con optimización de consultas,

ampliación de módulos, incorporación de nuevas vistas analíticas y evolución de las funciones de Inteligencia Artificial.

Finalmente, la implantación sobre infraestructura institucional permite proyectar una fase posterior de monitoreo continuo, ampliación de validación con nuevos usuarios, optimización de consultas pesadas y fortalecimiento de mecanismos de seguridad. Estas acciones contribuirán a consolidar progresivamente el sistema como herramienta de apoyo para la toma de decisiones estratégicas dentro de la DIDE-UTA.

CAPÍTULO IV. CONCLUSIONES Y RECOMENDACIONES

4.1 Conclusiones

- El análisis realizado en la DIDE-UTA identificó 8 problemáticas principales asociadas a la gestión de información científica, entre ellas la gestión manual de registros, dispersión de datos, ausencia de reportería centralizada, limitada trazabilidad, procesos manuales de revisión, falta de integración externa y ausencia de dashboards analíticos. A partir de este diagnóstico se establecieron 10 necesidades institucionales, registro de información, visualización y gestión de datos, automatización de validaciones, carga masiva, workflow, reportería BI, integración con fuentes externas y funciones de IA. Por tanto, se concluye que el análisis inicial permitió sustentar técnica y funcionalmente el desarrollo de una solución tecnológica alineada con las necesidades reales de la institución.
- El sistema desarrollado implementa una plataforma modular compuesta por: componentes de seguridad, gestión de artículos científicos, carga masiva, workflow, integración externa, modelo OLTP, procesos ETL, modelo OLAP, reportería BI, auditoría e Inteligencia Artificial. En el modelo transaccional se organizaron entidades relacionadas con artículos, autores, revistas, indexaciones, catálogos, formularios dinámicos, staging y trazabilidad. A nivel analítico, el Data Warehouse fue estructurado como una constelación dimensional, compuesta por 4 tablas de hechos principales: FactArticlePublication, FactArticleIndexing, FactVenueMetricYear y FactArticleAuthor, junto con dimensiones compartidas como fecha, artículo, autor, revista, campo académico, línea de investigación, estado e indexación. Esta arquitectura permitió consolidar información histórica para dashboards, KPIs como total de artículos científicos registrados, producción por año y mes, producción por facultad, línea de investigación, autores, revistas, fuentes de indexación, cuartiles, acceso abierto, trazabilidad institucional, registros procesados y tiempos de revisión, además de reportes orientados a la toma de decisiones sobre productividad científica institucional, desempeño por unidades académicas, calidad editorial, colaboración autoral, seguimiento del

workflow, integración de datos externos y planificación de estrategias para fortalecer la producción científica mediante BI.

- El módulo de Inteligencia Artificial incorporó funciones de predicción y recomendación sobre la información histórica consolidada en el Data Warehouse. Para la predicción de producción científica se consideraron modelos como persistencia del último valor, promedio móvil de tres meses, regresión lineal temporal y ForecastBySsa de ML.NET. La selección del modelo se sustentó mediante métricas como MAE, RMSE y MAPE, donde el error absoluto medio se expresa como $MAE = \sum |Real - Predicho| / n$. Bajo este criterio, el sistema selecciona como mejor modelo aquel que presenta el menor MAE y, en caso de empate, el menor RMSE, permitiendo priorizar el algoritmo con mejor comportamiento sobre la serie histórica mensual. En el estado actual, el módulo opera con un enfoque híbrido, utilizando modelos base explicables cuando el volumen histórico es limitado y activando ForecastBySsa cuando existen suficientes meses de datos para realizar predicción temporal con mayor estabilidad. Esto permitió generar estimaciones de producción científica para periodos posteriores, junto con niveles de confianza e intervalos esperados que facilitan la interpretación de los resultados.
- Además, el módulo incorporó un indicador de preparación IA mediante la fórmula ponderada $OverallScore = (Volumen \times 0,30) + (Trazabilidad\ de\ autores \times 0,22) + (Cobertura\ de\ indexación \times 0,20) + (Calidad\ de\ carga \times 0,14) + (Señales\ de\ workflow \times 0,14)$. A partir de este enfoque, se concluye que la IA implementada aporta una base funcional para el análisis predictivo, generación de recomendaciones explicables y apoyo inicial a la toma de decisiones institucionales. Las recomendaciones generadas se orientan a campos de interés para la gestión científica, como trazabilidad autoral, normalización de filiación, ORCID, cobertura de indexación, cuartiles, producción por facultad, líneas de investigación, revistas, acceso abierto, tendencias descendentes y registros con baja calidad de datos. De esta forma, el módulo no solo predice posibles comportamientos futuros, sino que también entrega acciones interpretables para mejorar la calidad de la información, la planificación editorial y el monitoreo de la producción científica institucional.

- Las pruebas aplicadas permitieron validar los flujos críticos del sistema durante y después del desarrollo. Se ejecutaron pruebas de carga masiva, integración, reportería, exportación, smoke, ETL y carga funcional, incluyendo 10 casos principales de carga masiva e integridad, todos aprobados. También se utilizaron datasets de prueba con diferentes volúmenes, incluyendo cargas de 20, 25, 18, 15, 16 y hasta 1000 registros, lo que permitió evaluar el comportamiento del sistema frente a escenarios mínimos, completos, con errores controlados y de mayor volumen. Los resultados demostraron que el sistema puede registrar información válida, detectar inconsistencias como DOI duplicado, claves foráneas inválidas y campos obligatorios vacíos, generar reportes PDF/Excel y actualizar el entorno analítico mediante ETL. Por ello, se concluye que la plataforma alcanzó un comportamiento funcional adecuado para una fase inicial de uso institucional.
- La medición de tiempos de reportería evidenció una mejora significativa frente al proceso manual previo. En la matriz de desempeño se registraron 80 operaciones de reportería, con 80 operaciones exitosas, 0 fallidas y una tasa de éxito del 100 %. El tiempo promedio general del sistema fue de 10,53 segundos, con un tiempo acumulado de 14,04 minutos para todas las operaciones medidas. Al compararlo con una línea base manual estimada de 68,17 horas, se obtuvo un ahorro aproximado de 67,93 horas y una reducción estimada del 99,66 %, calculada mediante la relación entre el tiempo ahorrado y el tiempo manual estimado. Este resultado permite concluir que la automatización de reportería reduce considerablemente el esfuerzo operativo requerido para generar información institucional, aunque debe interpretarse como una estimación comparativa frente al proceso manual descrito por los usuarios.
- La evaluación post-implantación permitió complementar los resultados técnicos con la percepción de los participantes sobre el uso del sistema. La entrevista estructurada post-implantación basada en dimensiones CSUQ, aplicada en línea mediante Google Forms a 5 participantes, obtuvo un promedio general de 4,59 sobre 5, una aceptación global del 100 % y un Alfa de Cronbach de 0,973. Estos resultados evidencian una valoración favorable en dimensiones como utilidad del sistema, calidad de la información y

reportería, facilidad de uso, funciones de Inteligencia Artificial, satisfacción y aceptación inicial. Por tanto, se concluye que el sistema no solo presentó mejoras técnicas en tiempos y estabilidad funcional, sino también una percepción positiva respecto a su utilidad práctica y condiciones favorables para la adopción progresiva del mismo.

- La implantación del sistema en un servidor institucional permitió comprobar que la solución puede operar fuera del entorno local de desarrollo. El despliegue se realizó mediante contenedores independientes para backend y frontend, integrando ASP.NET Core, Blazor WebAssembly, Nginx, Docker y SQL Server. Además, se verificó la restauración y disponibilidad de las bases de datos transaccional y analítica, con 48 tablas en la base OLTP y 34 tablas en la base OLAP, junto con los esquemas requeridos para reportería y procesos ETL. Este resultado demuestra que el sistema posee condiciones de portabilidad, mantenibilidad y despliegue institucional, por lo que no se limita a un prototipo académico, sino que queda preparado para revisión, validación y operación progresiva dentro de la infraestructura universitaria.
- En conjunto, la implementación del sistema de reportería y visualización de datos con funciones de recomendación y predicción permitió fortalecer la gestión informativa de la DIDE-UTA mediante una solución basada en BI e IA. La plataforma centraliza información científica, automatiza procesos, incorpora carga masiva, mejora la trazabilidad, consolida datos históricos en una constelación dimensional, genera dashboards y reportes, y aporta funciones iniciales de análisis predictivo. Los resultados obtenidos, especialmente la validación de pruebas, la tasa de éxito del 100 %, la reducción estimada del 99,66 % en tiempos de reportería y la evaluación post-implantación con promedio de 4,59/5, permiten concluir que el sistema desarrollado constituye una herramienta tecnológica pertinente para apoyar la toma de decisiones estratégicas basadas en datos dentro del contexto institucional analizado.

4.2 Recomendaciones

- Resulta pertinente continuar la evaluación del sistema durante un periodo prolongado de uso institucional, con el fin de medir adopción real, frecuencia de uso, utilidad percibida y satisfacción después de interactuar con la plataforma en actividades cotidianas. Esta evaluación permitiría complementar las métricas técnicas obtenidas en la investigación con evidencia relacionada con experiencia de usuario, aceptación institucional y mejora de procesos. Además, sería conveniente aplicar instrumentos post-implantación, como encuestas de satisfacción o cuestionarios basados en dimensiones de utilidad, facilidad de uso, calidad de información y adopción, para fortalecer futuras investigaciones sobre el impacto del sistema.
- Es necesario mantener un proceso permanente de actualización, depuración y validación de la información registrada en la plataforma, debido a que la calidad de los dashboards, reportes, indicadores, predicciones y recomendaciones depende directamente de la consistencia de los datos. Para ello, deben establecerse controles periódicos sobre artículos, autores, revistas, indexaciones, DOI, cuartiles, líneas de investigación y estados de publicación. De igual manera, conviene definir una política regular de ejecución y supervisión de procesos ETL, garantizando que el Data Warehouse se mantenga actualizado y que la reportería institucional refleje información confiable y reciente.
- Como línea de evolución tecnológica, se sugiere ampliar progresivamente las capacidades del módulo de Inteligencia Artificial conforme aumente el volumen de datos históricos disponibles. La predicción implementada constituye una base funcional para anticipar tendencias de producción científica, sin embargo, futuras versiones podrían comparar nuevos algoritmos como ARIMA, Prophet, Random Forest, XGBoost, redes neuronales o modelos híbridos, utilizando métricas como MAE, RMSE, MAPE y R^2 . Asimismo, sería conveniente extender las recomendaciones hacia escenarios como riesgo de demora en workflow, recomendación de revistas, análisis de colaboración científica, detección de registros inconsistentes y explicabilidad de resultados para facilitar su interpretación.

- Desde el componente analítico, conviene continuar optimizando y extendiendo la arquitectura BI implementada mediante nuevas vistas, dimensiones, tablas de hechos e indicadores institucionales. La constelación dimensional desarrollada permite incorporar nuevos procesos de análisis sin rediseñar completamente el Data Warehouse. Por ello, futuras mejoras podrían incluir indicadores de impacto científico, colaboración interinstitucional, acceso abierto, productividad por unidad académica, errores de carga masiva, tiempos de revisión y cumplimiento de líneas estratégicas. Además, se podrían fortalecer consultas, índices, filtros, paginación y almacenamiento temporal de resultados frecuentes, especialmente en los componentes que presentaron picos de tiempo durante la evaluación.
- Para fortalecer la sostenibilidad técnica a largo plazo, es importante mantener documentación, capacitación y pruebas complementarias. La implantación mediante contenedores facilita la portabilidad y administración del sistema, no obstante, deben conservarse procedimientos documentados sobre despliegue, actualización, respaldo, recuperación, monitoreo y revisión de logs. También resulta conveniente incorporar pruebas automatizadas complementarias para validar reglas de negocio, servicios ETL, carga masiva, generación de reportes y componentes de IA en futuras versiones. Finalmente, el uso del sistema debería acompañarse con capacitaciones breves sobre registro de información, carga masiva, workflow, dashboards, exportación de reportes e interpretación de predicciones y recomendaciones.

REFERENCIAS BIBLIOGRÁFICAS

- [1] R. R. Tirno, "Effect of Business Intelligence on the Organizational Competitiveness: Exploring the Mediation of the Technology Anxiety," *Computers in Human Behavior Reports*, nov. 2024, art. no. 100536. doi: 10.1016/j.chbr.2024.100536. Disponible en: <https://doi.org/10.1016/j.chbr.2024.100536>
- [2] Y. Niu et al., "Organizational business intelligence and decision making using big data analytics," *Information Processing & Management*, vol. 58, no. 6, nov. 2021, art. no. 102725. doi: 10.1016/j.ipm.2021.102725. Disponible en: <https://doi.org/10.1016/j.ipm.2021.102725>
- [3] C. Nyanga, J. Pansiri y D. Chatibura, "Enhancing competitiveness in the tourism industry through the use of business intelligence: A literature review," *Journal of Tourism Futures*, vol. 6, no. 2, nov. 2019, pp. 139–151. doi: 10.1108/JTF-11-2018-0069. Disponible en: <https://doi.org/10.1108/JTF-11-2018-0069>
- [4] Y.-S. Wang et al., "Developing and validating a mobile catering app success model," *International Journal of Hospitality Management*, vol. 77, ene. 2019, pp. 19–30. doi: 10.1016/j.ijhm.2018.06.002. Disponible en: <https://doi.org/10.1016/j.ijhm.2018.06.002>
- [5] Sudaryono, U. Rahardja y E. P. Harahap, "Implementation of Information Planning and Strategies Industrial Technology 4.0 to Improve Business Intelligence Performance on Official Site APTISI," *Journal of Physics: Conference Series*, vol. 1179, jul. 2019, art. no. 012111. doi: 10.1088/1742-6596/1179/1/012111. Disponible en: <https://doi.org/10.1088/1742-6596/1179/1/012111>
- [6] V. Kalina, J. Lhota y Z. T. Kalender, "Effective project portfolio management for SMEs: A conceptual framework using business intelligence tools," *Procedia Computer Science*, vol. 253, 2025, pp. 745–756. doi: 10.1016/j.procs.2025.01.136. Disponible en: <https://doi.org/10.1016/j.procs.2025.01.136>
- [7] M. Correa-Peralta, J. Vinueza-Martínez y L. Castillo-Heredia, "Evolución, temas y metodologías de investigación relevantes en inteligencia de negocios y análisis de

datos en la gestión académica de instituciones de educación superior,” *Scopus-indexed bibliometric review*, 2023.

[8] H. A. Mengash, “Using Data Mining Techniques to Predict Student Performance to Support Decision Making in University Admission Systems,” *IEEE Access*, vol. 8, 2020, pp. 55462–55470. doi: 10.1109/ACCESS.2020.2981905

[9] L. A. Huamán y M. R. Arangüena Yllanes, “Business intelligence en la gestión académica de la educación superior universitaria,” *Revista de Investigaciones de la Escuela de Posgrado*, vol. 2, no. 2, 2018, pp. 77–88. doi: 10.26788/riepg.2018.2.77

[10] B. Zhang et al., “Sustainable Digital Marketing: Factors of Adoption of M-Technologies by Older Adults in the Chinese Market,” *Sustainability*, vol. 15, no. 3, ene. 2023, art. no. 1972. doi: 10.3390/su15031972. Disponible en: <https://doi.org/10.3390/su15031972>

[11] Hofstede Insights, “Country Comparison: What about Bangladesh?” 2023. Disponible en: <https://www.hofstede-insights.com/country-comparison/bangladesh/> (consultado el 09-jul-2025).

[12] S. Prayogo, M. B. Hidayanto y M. Lubis, “Business Intelligence in e-learning for Higher Education,” en *Proceedings of the 11th International Conference on Computer and Communications Management (ICCCM 2023)*, Nagoya, Japan, ago. 2023, pp. 1–6. doi: 10.1145/3617733.3617768

[13] G. Li et al., “Evaluación de los desafíos de la transformación digital para la innovación del modelo de negocio en el contexto de las instituciones de educación superior utilizando un enfoque de apoyo a la toma de decisiones,” *Journal of Innovation & Knowledge*, vol. 9, no. 1, 2024, art. no. 100527. doi: 10.1016/j.jik.2024.100527

[14] J. Wolf, T. Hall y L. Robersshaw, “Business Intelligence in Research Administration: An Emerging Frontier,” *Journal of Higher Education Management*, vol. 29, no. 1, 2021, pp. 101–120.

[15] J. Macías y F. Borges, “Paneles interactivos de BI para evaluar la experiencia de usuario,” *Revista de Tecnología y Sociedad*, vol. 12, no. 2, 2023, pp. 45–60.

- [16] Number Analytics, *Global Educational Analytics Market Size 2020–2025*, 2024. Disponible en: <https://numberanalytics.com/educationBI-growth>
- [17] MoldStud, *How BI Improved Admission Processing in Higher Education*, 2023. Disponible en: <https://moldstud.com/articles/bi-admissions-efficiency>
- [18] B. Anđić, “Artificial Intelligence and Deepfake Learning in Higher Education,” *Journal of Baltic Science Education*, vol. 24, no. 1, feb. 2025, pp. 4–6. doi: 10.33225/jbse/25.24.04. Disponible en: <https://doi.org/10.33225/jbse/25.24.04>
- [19] H. Xueliang, T. X. Lin y W. Xiao, “Mitigate Cross-Market Competition Caused by the Risk of Uncertainty and Improve Firm Performance through Business Intelligence,” *Heliyon*, jul. 2024. doi: 10.1016/j.heliyon.2024.e34542. Disponible en: <https://doi.org/10.1016/j.heliyon.2024.e34542>
- [20] T. N. Mudau, J. F. Cohen y E. Papageorgiou, “Determinants and consequences of routine and advanced use of business intelligence (BI) systems by management accountants,” *Information & Management*, vol. 61, 2024, art. no. 103888. doi: 10.1016/j.im.2023.103888. Disponible en: <https://doi.org/10.1016/j.im.2023.103888>
- [21] M. Gonçalves et al., “Decision support system for port terminals: Design and development of a business intelligence tool,” *Procedia Computer Science*, vol. 243, 2025, pp. 2438–2447. doi: 10.1016/j.procs.2025.01.304
- [22] C. M. Olszak, “Business Intelligence Systems for Innovative Development of Organizations,” *Procedia Computer Science*, vol. 207, 2022, pp. 1754–1762. doi: 10.1016/j.procs.2022.09.233. Disponible en: <https://doi.org/10.1016/j.procs.2022.09.233>
- [23] A. Sorour y S. A. Atkins, “Big data challenge for monitoring quality in higher education institutions using business intelligence dashboards,” *Journal of Electronic Science and Technology*, vol. 22, no. 1, 2024, art. no. 100233. doi: 10.1016/j.jnlest.2024.100233
- [24] P. P. Singh et al., “Investigating Customer Churn in Banking: A Machine Learning Approach and Visualization App for Data Science and Management,” *Data Science and Management*, vol. 7, no. 1, mar. 2024, pp. 7–16. doi: 10.1016/j.dsm.2023.09.002

- [25] S. L. van der Meijden et al., “Development and validation of artificial intelligence models for early detection of postoperative infections (PERISCOPE): A multicentre study using electronic health record data,” *The Lancet Regional Health – Europe*, vol. 49, 2024, art. no. 101163. doi: 10.1016/j.lanepe.2024.101163
- [26] L. G. M. Siqueira et al., “Implementation of a Business Intelligence System in the Brazilian Nuclear Industry: An Action Research,” *Procedia Computer Science*, vol. 232, 2024, pp. 956–965. doi: 10.1016/j.procs.2023.01.304. Disponible en: <https://doi.org/10.1016/j.procs.2023.01.304>
- [27] J. Duque, A. Godinho y J. Vasconcelos, “Knowledge data extraction for business intelligence: A design science research approach,” *Procedia Computer Science*, vol. 204, 2022, pp. 131–139. doi: 10.1016/j.procs.2022.08.016. Disponible en: <https://doi.org/10.1016/j.procs.2022.08.016>
- [28] S. Haymond, “Create laboratory business intelligence dashboards for free using R: A tutorial using the flexdashboard package,” *Journal of Mass Spectrometry and Advances in the Clinical Lab*, vol. 23, 2022, pp. 39–43. doi: 10.1016/j.jmsacl.2021.12.002. Disponible en: <https://doi.org/10.1016/j.jmsacl.2021.12.002>
- [29] W. Strielkowski et al., “Predictive demand analytics and machine learning in electric power systems for enhancing resilience and efficiency,” *Sustainable Energy, Grids and Networks*, vol. 42, 2025, art. no. 101722. doi: 10.1016/j.segan.2025.101722. Disponible en: <https://doi.org/10.1016/j.segan.2025.101722>
- [30] J. Yoo et al., “Development of an Interoperable and Easily Transferable Clinical Decision Support System Deployment Platform: System Design and Development Study,” *Journal of Medical Internet Research*, vol. 24, no. 7, jul. 2022, e37928. doi: 10.2196/37928. Disponible en: <https://doi.org/10.2196/37928>
- [31] F. E. Mouas et al., “Machine learning prediction of histological types of breast cancer: A case study in Morocco,” *Intelligence-Based Medicine*, vol. 12, jul. 2025, art. no. 100275. doi: 10.1016/j.ibmed.2025.100275. Disponible en: <https://doi.org/10.1016/j.ibmed.2025.100275>

- [32] T. Cerratto-Pargman, C. McGrath y M. Milrad, “Towards Responsible AI in Education: Challenges and Implications for Research and Practice,” *Computers and Education: Artificial Intelligence*, dic. 2024, art. no. 100345. doi: 10.1016/j.caeai.2024.100345. Disponible en: <https://doi.org/10.1016/j.caeai.2024.100345>
- [33] S. Haque et al., “Towards development of functional climate-driven early warning systems for climate-sensitive infectious diseases: Statistical models and recommendations,” *Environmental Research*, vol. 249, mayo 2024, art. no. 118568. doi: 10.1016/j.envres.2024.118568. Disponible en: <https://doi.org/10.1016/j.envres.2024.118568>
- [34] D. Sjödin et al., “How AI capabilities enable business model innovation: Scaling AI through co-evolutionary processes and feedback loops,” *Journal of Business Research*, vol. 134, jun. 2021, pp. 574–587. doi: 10.1016/j.jbusres.2021.05.009. Disponible en: <https://doi.org/10.1016/j.jbusres.2021.05.009>
- [35] A. Berges et al., “A Framework for Strategic Intelligence Systems Applied to Education Management: A Pilot Study in the Community of Madrid,” *IEEE Access*, vol. 9, 2021, pp. 75313–75323. doi: 10.1109/ACCESS.2021.3081734. Disponible en: <https://doi.org/10.1109/ACCESS.2021.3081734>
- [36] W. D. Anjaningrum, N. Azizah y N. Suryadi, “Spurring SMEs’ Performance Through Business Intelligence, Organizational and Network Learning, Customer Value Anticipation, and Innovation: Empirical Evidence of the Creative Economy Sector in East Java, Indonesia,” *Heliyon*, vol. 10, no. 7, 2024, e27998. doi: 10.1016/j.heliyon.2024.e27998. Disponible en: <https://doi.org/10.1016/j.heliyon.2024.e27998>
- [37] P. Chinnasamy et al., “Health Recommendation System Using Deep Learning-Based Collaborative Filtering,” *Heliyon*, vol. 9, no. 12, 2023, e22844. doi: 10.1016/j.heliyon.2023.e22844. Disponible en: <https://doi.org/10.1016/j.heliyon.2023.e22844>
- [38] R. Chugh y S. Grandhi, “Business intelligence: A review,” *Journal of Technology Research*, vol. 5, 2014.

- [39] H. Chen, R. H. L. Chiang y V. C. Storey, “Business Intelligence and Analytics: From Big Data to Big Impact,” *MIS Quarterly*, vol. 36, no. 4, 2012, pp. 1165–1188. doi: 10.2307/41703503
- [40] R. Sharda, D. Delen y E. Turban, *Analytics, Data Science, and Artificial Intelligence: Systems for Decision Support*. Pearson, 2019.
- [41] Gartner, “Augmented Analytics is the Future of Data and Analytics,” 2019. Disponible en: <https://www.gartner.com/smarterwithgartner/augmented-analytics-is-the-future-of-data-and-analytics>
- [42] J. Kelleher, B. Mac Namee y A. D’Arcy, *Fundamentals of Machine Learning for Predictive Data Analytics*. MIT Press, 2015.
- [43] S. Few, *Information Dashboard Design: The Effective Visual Communication of Data*. O’Reilly Media, 2006.
- [44] F. Ricci, L. Rokach y B. Shapira, *Introduction to Recommender Systems*. Springer, 2011. doi: 10.1007/978-0-387-85820-3
- [45] P. Roth, *Blazor Revealed: Building Web Applications in .NET 6 and Beyond*. Apress, 2021. doi: 10.1007/978-1-4842-7136-4
- [46] R. Elmasri y S. Navathe, *Fundamentals of Database Systems*, 7th ed. Pearson, 2016.
- [47] C. Martínez-Salgado, “El muestreo en la investigación cualitativa: Principios básicos y algunas controversias,” *Ciencia & Saúde Coletiva*, vol. 17, no. 3, 2012, pp. 613–619. doi: 10.1590/S1413-81232012000300006
- [48] “Building Blazor WebAssembly Apps with Clean Architecture”. EzzyLearning.net. Accedido el 4 de junio de 2026. [En línea]. Disponible: <https://www.ezzylearning.net/tutorial/building-blazor-webassembly-apps-with-clean-architecture>

ANEXOS

Anexo 1. Entrevista estructurada inicial con escala tipo Likert

Objetivo: obtener información cuantificable sobre la percepción de los participantes respecto al estado actual de los procesos de gestión informativa, reportería institucional y pertinencia de implementar una solución basada en Business Intelligence e Inteligencia Artificial.

Dirigido a: funcionarios vinculados con procesos de gestión, registro, seguimiento o reportería de información científica institucional.

Instrucción: marque el valor que mejor represente su nivel de acuerdo con cada afirmación. La información será utilizada únicamente con fines académicos.

Valor	Interpretación
1	Totalmente en desacuerdo
2	En desacuerdo
3	Neutral
4	De acuerdo
5	Totalmente de acuerdo

Código	Ítem	Valoración
I1	Los procesos actuales de registro de artículos son ágiles y eficientes.	1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐
I2	El tiempo de generación de reportes es adecuado para las necesidades institucionales.	1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐
I3	Los datos registrados se mantienen organizados y actualizados.	1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐
I4	Me siento preparado/a para utilizar herramientas digitales más avanzadas.	1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐
I5	La implementación de un sistema BI-IA sería beneficiosa para la gestión institucional.	1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐
I6	Un sistema de reportería inteligente mejoraría la calidad de los informes institucionales.	1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐

I7	Las funciones de predicción y recomendación ayudarían a tomar mejores decisiones.	1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐
I8	Confío en la utilidad práctica del sistema propuesto.	1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐
I9	Considero que la interfaz debe ser simple y de fácil uso.	1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐
I10	Me gustaría participar en la validación o pruebas del nuevo sistema.	1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐

Relación de pertinencia: los ítems I1 a I3 permiten establecer la línea base sobre el proceso actual, los ítems I4 a I10 permiten valorar disposición tecnológica, utilidad esperada, aceptación de funciones BI-IA, usabilidad y participación en la validación del sistema.

Anexo 2. Entrevista semiestructurada con preguntas abiertas.

Objetivo: profundizar en el contexto institucional relacionado con la gestión de información científica, reportería, validación de registros, trazabilidad, integración de información externa y necesidades tecnológicas.

Instrucción: responda de forma abierta de acuerdo con su experiencia y conocimiento del proceso institucional. Las respuestas serán utilizadas únicamente con fines académicos.

Código	Pregunta abierta	Respuesta
P1	¿Cómo se realiza actualmente el registro y seguimiento de artículos científicos dentro de la institución?	
P2	¿Qué dificultades se presentan durante la consolidación, revisión o actualización de información científica?	
P3	¿Qué herramientas se utilizan actualmente para generar reportes institucionales y cuáles son sus principales limitaciones?	
P4	¿Qué tipo de información o indicadores se requieren con mayor frecuencia para procesos de seguimiento y toma de decisiones?	
P5	¿Qué problemas se presentan al trabajar con archivos, hojas de cálculo o fuentes de información separadas?	
P6	¿Qué mecanismos de validación o trazabilidad considera necesarios para mejorar el control de la información registrada?	
P7	¿Qué beneficios tendría contar con carga masiva de información y validación automática de registros?	
P8	¿Qué fuentes externas de información científica deberían integrarse o consultarse para enriquecer los registros institucionales?	
P9	¿Qué características debería tener un sistema de reportería y visualización para ser útil en el análisis institucional?	

P10	¿Qué tipo de recomendaciones o predicciones serían útiles para apoyar la planificación y seguimiento de la producción científica?	
P11	¿Qué aspectos considera importantes para facilitar la adopción del sistema?	
P12	¿Qué observaciones o sugerencias adicionales considera relevantes para el desarrollo de la propuesta tecnológica?	

Anexo 3. Entrevista estructurada post-implantación basada en dimensiones CSUQ.

Objetivo: valorar la percepción posterior a la revisión o uso del sistema, considerando utilidad, calidad de la información, facilidad de uso, funciones de Inteligencia Artificial, satisfacción general y adopción institucional inicial.

Instrucción: marque el valor que mejor represente su nivel de acuerdo con cada afirmación. La escala utilizada es de 1 a 5.

Valor	Interpretación
1	Totalmente en desacuerdo
2	En desacuerdo
3	Neutral
4	De acuerdo
5	Totalmente de acuerdo

Código	Dimensión	Ítem	Valoración
I1	Utilidad del sistema	El sistema facilita la gestión de información científica institucional.	1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐
I2	Utilidad del sistema	El sistema permite reducir el esfuerzo requerido para generar reportes institucionales.	1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐
I3	Utilidad del sistema	El sistema mejora el seguimiento de artículos, autores, revistas e indexaciones.	1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐
I4	Utilidad del sistema	Las funcionalidades del sistema responden a necesidades reales de gestión institucional.	1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐
I5	Calidad de información y reportería	La información presentada por el sistema es clara y comprensible.	1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐
I6	Calidad de información y reportería	Los reportes generados contienen información útil para el análisis institucional.	1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐
I7	Calidad de información y reportería	Los indicadores, gráficos y filtros permiten interpretar mejor la producción científica.	1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐

I8	Calidad de información y reportería	La exportación de reportes en PDF o Excel aporta valor para los procesos institucionales.	1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐
I9	Facilidad de uso e interfaz	La interfaz del sistema es fácil de utilizar.	1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐
I10	Facilidad de uso e interfaz	La navegación entre módulos resulta clara y ordenada.	1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐
I11	Facilidad de uso e interfaz	Los dashboards, tablas y reportes se presentan de forma visualmente adecuada.	1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐
I12	Facilidad de uso e interfaz	El sistema permite encontrar la información requerida de manera más rápida que el proceso manual.	1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐
I13	Funciones de IA	Las funciones de predicción ayudan a anticipar tendencias de producción científica.	1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐
I14	Funciones de IA	Las recomendaciones generadas por el sistema son comprensibles para el usuario.	1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐
I15	Funciones de IA	Las funciones de IA aportan valor como apoyo inicial para la planificación institucional.	1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐
I16	Funciones de IA	El sistema comunica de forma adecuada que las predicciones deben interpretarse como apoyo y no como decisiones absolutas.	1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐
I17	Satisfacción y adopción	En general, me siento satisfecho/a con el sistema desarrollado.	1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐
I18	Satisfacción y adopción	Considero que el sistema puede ser adoptado como apoyo para los procesos institucionales.	1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐
I19	Satisfacción y adopción	Recomendaría el uso del sistema para mejorar la reportería y análisis institucional.	1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐
I20	Satisfacción y adopción	El sistema contribuye a mejorar la toma de decisiones basada en datos.	1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐

Preguntas abiertas complementarias

Código	Pregunta	Respuesta
P1	¿Qué funcionalidad considera más útil dentro del sistema?	

P2	¿Qué aspecto del sistema considera que debería mejorarse?	
P3	¿Considera que el sistema puede apoyar los procesos de reportería y seguimiento institucional? Explique brevemente.	

Anexo 4. Product Backlog.

El Product Backlog organiza las historias de usuario priorizadas para el desarrollo del sistema. Cada historia se relaciona con un módulo funcional, una prioridad, un nivel de riesgo y el sprint asignado. Este anexo respalda la planificación metodológica aplicada durante el desarrollo de la propuesta.

ID	Historia	Usuario	Prioridad	Riesgo	Puntos	Sprint	Criterio principal
HU-01	Autenticación de usuario	Administrador, Analista DIDE, Autoridad	Alta	Baja	3	Sprint 1	Validar credenciales y permitir acceso según rol.
HU-02	Recuperación de contraseña	Usuario institucional	Media	Baja	3	Sprint 1	Solicitar recuperación, establecer nueva contraseña y validar seguridad básica.
HU-03	Gestión de usuarios	Administrador	Alta	Media	5	Sprint 1	Crear, editar, eliminar usuarios y asignar roles.
HU-04	Registro de artículos científicos	Analista DIDE / Investigador	Alta	Baja	5	Sprint 1	Registrar artículos científicos y almacenar información en base de datos.
HU-05	Carga masiva de artículos	Analista DIDE	Alta	Media	5	Sprint 1	Importar archivos Excel o CSV, validar estructura

							insertar registros.
HU-06	Visualización de dashboard	Autoridad DIDE / Analista	Alta	Media	5	Sprint 2	Visualizar indicadores y gráficos de producción científica.
HU-07	Generación de reportes	Analista DIDE	Alta	Media	5	Sprint 2	Aplicar filtros y generar reportes institucionales.
HU-08	Exportación de reportes	Analista DIDE	Alta	Baja	3	Sprint 2	Exportar reportes en PDF o Excel.
HU-09	Gestión de fuentes de datos	Administrador / Analista DIDE	Media	Media	5	Sprint 2	Registrar, editar o eliminar fuentes de datos.
HU-10	Limpieza y transformación de datos	Sistema / Analista DIDE	Media	Alta	5	Sprint 3	Eliminar duplicados y normalizar información.
HU-11	Entrenamiento del modelo predictivo	Administrador / Analista DIDE	Media	Alta	8	Sprint 3	Entrenar modelos con datos históricos y guardar resultados.
HU-12	Generación de predicciones	Autoridad DIDE / Analista	Alta	Media	5	Sprint 3	Generar predicciones basadas en datos disponibles.
HU-13	Generación de recomendaciones	Autoridad DIDE / Analista	Alta	Media	5	Sprint 3	Generar recomendaciones explicables basadas en

							análisis de datos.
HU-14	Gestión de modelos de IA	Administrador / Analista DIDE	Media	Media	5	Sprint 3	Activar, desactivar o revisar versiones de modelos de IA.
HU-15	Clasificación de artículos según OCDE	Administrador / Analista DIDE	Alta	Baja	3	Sprint 4	Clasificar artículos según estructura OCDE.
HU-16	Registro de auditoría del sistema	Administrador / Sistema	Media	Baja	3	Sprint 4	Registrar eventos y consultar historial de acciones.

Sprint	Historias incluidas	Resultado esperado
Sprint 1	HU-01 a HU-05	Sistema base con autenticación, usuarios, registro y carga masiva inicial.
Sprint 2	HU-06 a HU-09	Módulo de reportería, dashboards, exportación y fuentes de datos.
Sprint 3	HU-10 a HU-14	Procesamiento de datos, predicción, recomendaciones y gestión de modelos IA.
Sprint 4	HU-15 a HU-16	Clasificación OCDE, auditoría, integración final y ajustes generales.

Anexo 5. Plan de sprints y cronograma de desarrollo.

La planificación del desarrollo se basó en Scrum + XP, mediante sprints semanales orientados a entregar incrementos funcionales progresivos. Scrum permitió organizar backlog, planificación, revisión y seguimiento; XP complementó el proceso mediante prácticas técnicas de calidad, validación frecuente, refactorización, integración continua y control de versiones.

Fase	Actividad
Inicio del sprint	Sprint Planning y priorización de historias.
Desarrollo	Implementación de historias de usuario y componentes del sistema.
Seguimiento	Revisiones técnicas con el tutor y control de avances.
Validación	Revisión funcional con el responsable institucional o usuario relacionado.
Mejora continua	Ajustes derivados de la retroalimentación y revisión técnica.

Sprint	Objetivo	Funcionalidades principales	Entregable
Sprint 1	Establecer la infraestructura base del sistema.	Inicio de sesión, gestión de usuarios, registro de artículos y carga masiva inicial.	Sistema base funcional con usuarios, registro y almacenamiento operativo.
Sprint 2	Implementar reportería y visualización institucional.	Dashboards, generación de reportes, filtros, exportación y fuentes de datos.	Sistema con indicadores, reportes y visualización analítica.
Sprint 3	Implementar análisis avanzado e IA.	Procesamiento de datos, entrenamiento, predicciones, recomendaciones y gestión de modelos IA.	Sistema con funciones de predicción y recomendación basadas en datos.
Sprint 4	Integrar módulos y cerrar funcionalmente el sistema.	Clasificación OCDE, auditoría, pruebas finales, ajustes de integración y validación general.	Versión final lista para evaluación e implantación.

Semana	Actividades principales
Semana 1	Desarrollo de infraestructura base, autenticación, usuarios, registro y carga masiva inicial.
Semana 2	Desarrollo de reportería, dashboards, filtros y exportación.
Semana 3	Implementación de procesamiento de datos, predicción, recomendaciones y gestión IA.
Semana 4	Integración final, auditoría, clasificación OCDE, pruebas y ajustes.

Anexo 6. Definition of Done (DoD).

La Definition of Done establece los criterios mínimos para considerar que una historia de usuario o incremento del sistema fue completado satisfactoriamente. En esta versión se ajusta la redacción para mantener coherencia con la tesis: se documenta validación técnica frecuente y pruebas funcionales/integración, sin afirmar la existencia de un proyecto formal de pruebas unitarias automatizadas.

Criterio DoD	Aplicación dentro del proyecto
Implementación funcional	La funcionalidad debe estar desarrollada conforme a la historia de usuario y a sus criterios de aceptación.
Integración al sistema	El código debe estar integrado al repositorio y no afectar negativamente otros módulos.
Compilación sin errores	El sistema debe ejecutarse sin errores críticos de compilación.
Validación técnica frecuente	La funcionalidad debe ser validada mediante pruebas funcionales, integración, revisión técnica o evidencia operativa, según corresponda.
Ausencia de errores críticos	No deben existir errores bloqueantes durante la ejecución del flujo principal.
Revisión técnica	La funcionalidad debe ser revisada técnicamente con el tutor del proyecto.

Validación funcional	Los módulos orientados al usuario deben ser demostrados y validados funcionalmente.
Trazabilidad	La historia debe conservar relación con el backlog, sprint y evidencia de revisión.

Lista de control	Cumplimiento
Funcionalidad completamente implementada	☐
Criterios de aceptación cumplidos	☐
Compilación correcta	☐
Código integrado al repositorio	☐
Validación técnica o funcional ejecutada	☐
Sin errores críticos	☐
Revisión técnica realizada	☐
Validación funcional realizada cuando aplica	☐

Relación con XP: La DoD se complementa con integración continua, refactorización, entregas frecuentes y validación técnica frecuente. Las pruebas unitarias automatizadas se consideran una mejora técnica futura para fortalecer el aseguramiento de calidad en versiones posteriores.

Anexo 7. Plan de aseguramiento de calidad del software.

El plan de aseguramiento de calidad se fundamenta en prácticas de Extreme Programming (XP) adaptadas al contexto del proyecto. Su finalidad es mantener estabilidad, confiabilidad y mantenibilidad durante el desarrollo del sistema de reportería, visualización, recomendación y predicción.

Práctica de calidad	Aplicación corregida dentro del proyecto
---------------------	--

Validación técnica frecuente	Revisión de componentes críticos mediante pruebas funcionales, integración, revisión de resultados y evidencia operativa.
Integración continua	Uso de GitHub para registrar cambios, integrar avances y mantener trazabilidad del desarrollo.
Refactorización continua	Mejora progresiva de estructura interna del código sin alterar comportamiento funcional.
Entregas frecuentes	Presentación de incrementos funcionales durante cada sprint para recibir retroalimentación.
Control de versiones	Registro histórico de cambios y recuperación de versiones previas cuando sea necesario.

Tipo de prueba	Descripción
Pruebas funcionales	Verifican que las funcionalidades respondan al comportamiento esperado.
Pruebas de integración	Validan la interacción entre API, base de datos, staging, reportería, exportación y ETL.
Pruebas smoke	Comprueban rápidamente que los servicios principales estén operativos.
Pruebas de carga funcional	Evalúan el comportamiento con datasets pequeños, medianos y extensos.
Pruebas de usuario / validación funcional	Permiten revisar funcionalidades con actores relacionados al sistema.
Pruebas unitarias automatizadas	Se consideran mejora técnica futura para validadores, reglas de negocio y servicios críticos.

Métrica de calidad	Descripción
Número de defectos detectados	Errores encontrados durante pruebas funcionales o revisión técnica.
Velocidad del sprint	Historias completadas por sprint.
Frecuencia de integración	Número de integraciones o commits registrados.

Resultados de pruebas	Casos aprobados, rechazados o con observaciones.
Tiempos de reportería	Indicadores de rendimiento registrados por el sistema.

Anexo H. Reporte de pruebas prioritarias del sistema

Este reporte consolida las pruebas más relevantes aplicadas o preparadas para validar el sistema, priorizando los flujos críticos de registro individual, carga masiva mediante staging, validaciones de datos, workflow institucional, reportería, exportación, ETL y rendimiento funcional con volumen de datos.

Estado de automatización

En la versión documentada no existe un proyecto formal de pruebas unitarias integrado con xUnit, NUnit o MSTest. Sin embargo, existen artefactos de prueba operativa, integración, smoke, carga funcional y datasets de validación. Por ello, la tesis documenta pruebas unitarias como mejora técnica futura y evidencia principalmente pruebas funcionales, integración, smoke, reportería, exportación y carga funcional.

Tipo de prueba	Cobertura principal	Importancia
Smoke / end-to-end parcial	Backend, autenticación, usuario actual, salud de reportería, dashboard, analítica de autores, filtros, PDF, Excel y ETL opcional.	Demuestra que autenticación, reportería, DW, exportación y consultas principales trabajan en conjunto.
Integración de registro individual	Registro mínimo válido, DOI único, persistencia de revista, métricas editoriales, tiempos y control de fallos.	Evidencia estabilidad del flujo individual bajo repetición controlada.
Carga masiva con staging	Upload, resumen de lote, validación, procesamiento, tiempos y persistencia.	Demuestra recepción, validación y procesamiento controlado de información en masa.
Errores controlados	Campos obligatorios vacíos, fechas inválidas, autores incompletos, errores por fila y no persistencia parcial.	Evidencia control de calidad antes de alimentar OLTP, reportería y ETL.

Carga funcional	Datasets de 15, 16, 18, 20, 25 y 1000 registros.	Permite evaluar comportamiento con volúmenes pequeños, medianos y altos.
Reportería, ETL y exportación	Dashboard, filtros, analítica de autores, PDF, Excel y actualización posterior al ETL.	Demuestra el funcionamiento del componente BI y generación de reportes institucionales.

ID	Componente	Prueba unitaria prioritaria
UT-01	Validador de campos dinámicos	Campo requerido vacío devuelve error.
UT-02	Validador de campos dinámicos	Campo numérico rechaza texto.
UT-03	Validador de campos dinámicos	Longitud máxima se respeta.
UT-04	Validación DOI	DOI duplicado o inválido se rechaza.
UT-05	Participantes	Debe existir al menos un autor o participante.
UT-06	Venue / revista	La revista se resuelve por nombre o ISSN sin duplicar innecesariamente.
UT-07	Carga masiva	Una fila inválida genera error de staging y no excepción general.
UT-08	Workflow	El flujo respeta reglas de devolución, revisión y procesamiento.
UT-09	Reportería	Generación de indicadores mantiene conteos esperados.
UT-10	IA	Reentrenamiento con datos suficientes registra historial y métricas.

Conclusión del anexo: La mayor cobertura existente corresponde a pruebas de integración, smoke, carga masiva, errores controlados, reportería, exportación y carga funcional. La mejora técnica recomendada consiste en implementar un proyecto formal de pruebas unitarias para validadores, reglas de negocio y servicios críticos.

Anexo 8. Casos de prueba de carga masiva e integridad de datos.

Los siguientes casos consolidan las pruebas de carga masiva e integridad de datos. Se corrigió la numeración para evitar duplicidad del código CI-LOAD-08 en los escenarios de error.

ID del caso	Nombre del caso	Módulo	Tipo de prueba	Prioridad	Estado	Observación
CI-LOAD-01	Artículo con solo campos básicos	Importación / carga masiva	Carga de datos e integridad	Alta	Aprobado	Sin observaciones.
CI-LOAD-02	Artículo con catálogos completos	Importación / carga masiva	Carga de datos e integridad	Alta	Aprobado	Sin observaciones.
CI-LOAD-03	Artículo con campos dinámicos	Importación / carga masiva	Carga de datos e integridad	Alta	Aprobado	Sin observaciones.
CI-LOAD-04	Artículo con un autor	Importación / carga masiva	Carga de datos e integridad	Alta	Aprobado	Sin observaciones.
CI-LOAD-05	Artículo con varios autores/coautores ordenados	Importación / carga masiva	Carga de datos e integridad	Alta	Aprobado	Sin observaciones.
CI-LOAD-06	Artículo con datos provenientes de integración externa	Importación / carga masiva	Carga de datos e integridad	Alta	Aprobado	Sin observaciones.
CI-LOAD-07	Carga masiva con registros válidos	Importación / carga masiva	Carga de datos e integridad	Alta	Aprobado	Sin observaciones.
CI-LOAD-08	Carga masiva con errores: DOI duplicado	Importación / carga masiva	Validación negativa	Alta	Aprobado	El sistema detecta duplicidad y

						evita consolidación inválida.
CI-LOAD-09	Carga masiva con errores: FK inválida	Importación / carga masiva	Validación negativa	Alta	Aprobado	El sistema identifica referencia inválida antes de persistir.
CI-LOAD-10	Carga masiva con errores: campo requerido vacío	Importación / carga masiva	Validación negativa	Alta	Aprobado	El sistema marca error por obligatoriedad y bloquea la fila inválida.

Detalle resumido de los casos principales

CI-LOAD-01. Artículo con solo campos básicos

Objetivo: Validar que el modelo permita registrar un artículo con campos mínimos, sin catálogos, participantes ni campos dinámicos.

Criterio de aceptación: El artículo se inserta correctamente, los campos de catálogo quedan nulos y no se generan registros asociados innecesarios.

CI-LOAD-02. Artículo con catálogos completos

Objetivo: Validar que el artículo se registre con revistas, periodos, estados, líneas y clasificación académica jerárquica.

Criterio de aceptación: Todas las claves foráneas apuntan a registros válidos y las consultas con joins devuelven la clasificación completa.

CI-LOAD-03. Artículo con campos dinámicos

Objetivo: Validar la extensibilidad del modelo mediante atributos adicionales sin modificar la tabla principal.

Criterio de aceptación: Los valores dinámicos se almacenan correctamente y se asocian al artículo mediante el catálogo de campos.

CI-LOAD-04. Artículo con un autor

Objetivo: Validar que un artículo pueda registrar un participante principal manteniendo orden de autoría.

Criterio de aceptación: El autor se registra y se asocia correctamente al artículo.

CI-LOAD-05. Artículo con varios autores/coautores ordenados

Objetivo: Validar la relación de múltiples participantes y el orden de autoría.

Criterio de aceptación: Los autores y coautores quedan asociados en el orden esperado.

CI-LOAD-06. Artículo con datos de integración externa

Objetivo: Validar la incorporación de información proveniente de servicios externos.

Criterio de aceptación: Los datos externos se normalizan y almacenan sin afectar la integridad del modelo.

CI-LOAD-07. Carga masiva con registros válidos

Objetivo: Validar el procesamiento de lotes con filas correctas.

Criterio de aceptación: El lote se valida y procesa, consolidando registros válidos.

CI-LOAD-08. Error por DOI duplicado

Objetivo: Comprobar que el sistema detecte artículos duplicados.

Criterio de aceptación: La fila queda marcada con error y no se consolida como registro definitivo.

CI-LOAD-09. Error por FK inválida

Objetivo: Comprobar que el sistema detecte referencias inexistentes o inválidas.

Criterio de aceptación: La fila inválida se rechaza o queda pendiente de corrección.

CI-LOAD-10. Error por campo requerido vacío

Objetivo: Comprobar que el sistema valide obligatoriedad de campos.

Criterio de aceptación: La fila con campo requerido vacío se marca con error y no se procesa.

Anexo 9. Matriz de tiempos de reportería.

El sistema reduce significativamente el esfuerzo manual de reportería y presenta tiempos promedio adecuados, aunque existen picos de carga que deben considerarse oportunidades de optimización en producción.

Fuente de datos	Valor
Tabla consultada	dbo.ReportingPerformanceMetrics
Base de datos	TesisDB_Extensible
Periodo medido	2026-05-22 05:28:11 UTC a 2026-05-24 23:36:59 UTC
Cantidad de mediciones	80 operaciones de reportería
Tasa de éxito	80 exitosas / 80 totales = 100 %

Indicador	Valor observado	Lectura
Operaciones medidas	80	Muestra inicial suficiente para describir comportamiento operativo preliminar.
Operaciones exitosas	80	No se observaron fallos en las operaciones registradas.
Operaciones fallidas	0	La reportería se mantuvo estable durante la muestra.
Tasa de éxito	100 %	Indicador favorable de estabilidad funcional.
Tiempo promedio real	10,53 segundos	Tiempo promedio adecuado para consultas institucionales.
Tiempo real acumulado	14,04 minutos	Las 80 operaciones se resolvieron en menos de 15 minutos acumulados.
Tiempo manual estimado acumulado	68,17 horas	Línea base estimada para tareas manuales equivalentes.
Ahorro estimado acumulado	67,93 horas	Evidencia del potencial de reducción de tiempos.
Reducción estimada	99,66 %	Resultado favorable, presentado como estimación frente a la línea base manual.

Elementos consultados acumulados	95 629	La muestra incluye operaciones con volumen relevante de datos.
----------------------------------	--------	--

Operación	Ejecuciones	Promedio	Mediana	P90	Máximo	Lectura
Dashboard institucional	32	16,37 s	0,44 s	96,01 s	96,40 s	La mayoría de consultas fueron rápidas, pero existen picos altos cercanos a 1,6 minutos.
Panel autores/coautoría	29	8,94 s	0,38 s	15,17 s	99,16 s	Comportamiento promedio favorable; un valor máximo aislado evidencia necesidad de monitoreo.
ETL reportería	12	2,63 s	2,71 s	3,41 s	3,43 s	Tiempo estable y bajo en la muestra medida.
Exportación Excel	3	6,71 s	7,07 s	7,12 s	7,14 s	Generación de Excel consistente y menor a 10 segundos.
Exportación PDF	3	2,12 s	0,55 s	4,41 s	5,38 s	Generación de PDF rápida en la muestra.
PDF compuesto	1	0,83 s	0,82 s	0,82 s	0,83 s	Evidencia puntual; requiere más muestras para generalizar.

Rango	Clasificación	Interpretación
0 a 5 segundos	Rápido	Respuesta muy favorable para uso operativo.
5 a 15 segundos	Adecuado	Aceptable para reportería con datos consolidados.
15 a 60 segundos	Moderado	Puede ser aceptable si el reporte es pesado o tiene filtros amplios.
60 a 120 segundos	Alto	Debe monitorearse y optimizarse antes de producción intensiva.
Mayor a 120 segundos	Crítico	Requiere optimización prioritaria.

Dimensión evaluada	Resultado	Conclusión
Estabilidad	80/80 operaciones exitosas	Favorable.
Velocidad promedio	10,53 segundos	Favorable.
Exportes	PDF y Excel menores a 10 segundos promedio	Favorable.
ETL	2,63 segundos promedio	Favorable en la muestra.
Dashboard general	Promedio aceptable, con picos altos	Requiere monitoreo.
Autores/coautoría	Promedio aceptable, con un pico alto	Requiere monitoreo.
Reducción de tiempo	99,66 % estimado	Muy favorable, con aclaración metodológica.

Anexo 10. Selección de metodología Scrum + XP.

Campo	Descripción
Macroproceso	Investigación, innovación y emprendimiento
Proceso	Gestión de producción científica
Documento	Metodología ágil aplicada
Fecha de aprobación	Aprobado con fecha 02 de septiembre de 2025. Resolución N.º 0920-P-CD-FISEI-UTA-2025
Propuesta del trabajo de titulación	Implementación de un sistema de reportería y visualización de datos con funciones de recomendación y predicción, utilizando técnicas de Business Intelligence (BI) e Inteligencia Artificial (IA), en la Dirección de Investigación y Desarrollo (DIDE) de la Universidad Técnica de Ambato (UTA).
Autor	Christopher Paul Santamaría Márquez

1. Fundamentos de las metodologías ágiles

Las metodologías ágiles de desarrollo de software surgen como una alternativa a los enfoques tradicionales de tipo secuencial, proponiendo ciclos de desarrollo iterativos e incrementales que favorecen la adaptación al cambio, la entrega temprana de valor y la participación de los usuarios finales. Estos enfoques priorizan la colaboración continua, la retroalimentación frecuente y la mejora progresiva del producto, aspectos especialmente relevantes en proyectos donde los requerimientos evolucionan conforme se obtiene un mayor entendimiento del dominio del problema.

En el contexto institucional de la Dirección de Investigación y Desarrollo (DIDE) de la Universidad Técnica de Ambato, el desarrollo del sistema de reportería y visualización de datos con funciones de recomendación y predicción requiere un enfoque metodológico que permita equilibrar la flexibilidad propia de los entornos ágiles con las necesidades de control, trazabilidad y aprobación formal propias de una

institución académica. En este sentido, no todas las metodologías ágiles presentan el mismo grado de adecuación al contexto del proyecto, lo que hace necesaria una evaluación comparativa para seleccionar el enfoque más apropiado.

2. Comparación cualitativa de metodologías ágiles

La comparación cualitativa permitió revisar las principales metodologías ágiles consideradas para el desarrollo del sistema, evaluando su grado de ajuste en función de las características operativas del proyecto: cadencia semanal de trabajo, gestión de cambios, documentación institucional, elaboración de artefactos de diseño, planificación por módulos y control del flujo de actividades.

Metodología	Enfoque y flujo	Fortalezas	Limitaciones	Ajuste global
Scrum	Iterativo por sprints, roles y ceremonias.	Excelente para cadencia fija, control por incremento y gestión por backlog.	Puede resultar rígido si ingresan urgencias durante el sprint.	Alto
Kanban	Flujo continuo, límites WIP, tableros y mejora continua.	Adecuado para cambios constantes, urgencias y visibilidad del flujo.	Menos natural para metas por fecha si no se definen cadencias.	Medio-Alto
Scrumban	Híbrido entre Scrum y Kanban.	Combina cadencia iterativa con control de flujo y manejo de cambios.	Requiere reglas claras para evitar pérdida de estructura metodológica.	Muy alto
Extreme Programming (XP)	Metodología enfocada en prácticas técnicas de ingeniería.	Fortalece calidad del software mediante integración continua, refactorización y validación frecuente.	No resuelve por sí sola la gestión estructural del proyecto.	Alto

Lean Software Development	Optimización del valor y eliminación de desperdicio.	Útil para mejorar eficiencia y reducir actividades innecesarias.	Funciona más como conjunto de principios que como método operativo completo.	Medio
DSDM / AgilePM	Enfoque ágil con gobernanza, fases, roles y timeboxing.	Fuerte para control, documentación, aprobaciones y gestión formal.	Puede resultar más pesado para proyectos de desarrollo individual.	Muy alto
FDD	Desarrollo guiado por funcionalidades pequeñas.	Adecuado cuando las funcionalidades son claras y medibles.	Menor flexibilidad frente a cambios constantes.	Medio-Alto
Scrum + XP	Gestión iterativa por sprints combinada con prácticas técnicas XP.	Integra planificación, control, revisión técnica, refactorización, integración continua y validación frecuente.	Requiere disciplina técnica y evidencias de validación durante cada incremento.	Muy alto

3. Matriz cuantitativa ponderada para la selección metodológica

3.1 Definición de criterios de ponderación

Código	Criterio	Peso
C1	Entregas semanales y cadencia de reuniones	20 %
C2	Gestión de cambios constantes	20 %
C3	Documentación y aprobaciones institucionales	15 %

C4	Diseño y artefactos: mockups, casos de uso y diagramas	10 %
C5	Metas mensuales por módulos o releases	15 %
C6	Planificación de pruebas y despliegue	10 %
C7	Visibilidad y control del flujo mediante tablero y límites WIP	10 %
	Total	100 %

3.2 Asignación de puntajes por metodología

Para la evaluación de las metodologías ágiles se empleó una escala ordinal de cinco niveles. No obstante, en el presente estudio se utilizaron de forma efectiva los valores 3, 4 y 5, correspondientes a los niveles medio, alto y muy alto, respectivamente. Esta decisión se fundamenta en que todas las metodologías analizadas pertenecen al paradigma ágil, por lo que ninguna presenta un desempeño muy bajo o bajo respecto a los criterios considerados. El uso de los valores 1 y 2 resultaría inapropiado y artificial, al no representar adecuadamente la naturaleza de las metodologías evaluadas ni el contexto del proyecto.

La asignación de los valores numéricos se realizó conforme a la siguiente regla: 3 significa que la metodología satisface el criterio de forma parcial o indirecta; 4 significa que lo satisface de manera adecuada con ajustes menores; y 5 significa que lo satisface de forma directa y natural, sin requerir modificaciones sustanciales para ajustarse al contexto del proyecto.

Metodología	C1	C2	C3	C4	C5	C6	C7
Scrumban	5	5	4	5	5	4	5
Scrum	5	4	4	5	5	4	3
Kanban	3	5	3	3	3	3	5

XP	5	5	3	3	4	4	3
Lean	3	5	3	3	3	3	4
DSDM/AgilePM	5	4	5	5	5	5	3
FDD	4	3	3	5	5	4	3
Scrum + XP	5	5	4	5	5	5	4

3.3 Cálculo del puntaje ponderado

Metodología	Puntaje (0-5)	Porcentaje
Scrumban	4,75	95 %
Scrum + XP	4,75	95 %
DSDM/AgilePM	4,60	92 %
Scrum	4,35	87 %
XP	4,05	81 %
FDD	3,80	76 %
Kanban	3,60	72 %
Lean	3,50	70 %

4. Selección de metodología ágil

Con base en la evaluación cualitativa y en los resultados obtenidos a partir de la matriz cuantitativa ponderada, se observa que Scrumban y el modelo híbrido Scrum + XP

alcanzan el mayor nivel de adecuación global al contexto del proyecto, obteniendo ambos el puntaje más alto en la comparación. No obstante, se seleccionó Scrum + XP como metodología a aplicar, debido a que, manteniendo el mismo nivel de ajuste global, presentó mejor desempeño en el criterio C6, relacionado con la planificación de pruebas y despliegue. Este criterio fue considerado crítico para asegurar la calidad, estabilidad y confiabilidad del software, especialmente porque el sistema contempla procesos de reportería, ETL, análisis de datos y funciones de recomendación y predicción.

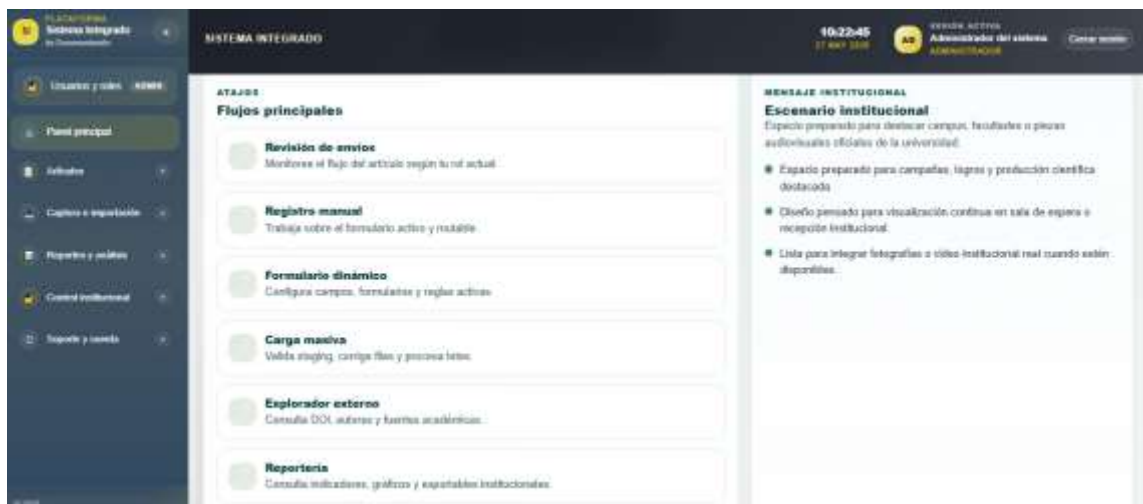
En consecuencia, la combinación de Scrum y XP permitió integrar la gestión iterativa del desarrollo con prácticas técnicas orientadas a la calidad del software, tales como integración continua, refactorización, control de versiones, revisiones técnicas y validación frecuente de funcionalidades. Scrum aportó organización mediante sprints, backlog, planificación y revisión de incrementos, mientras que XP complementó el proceso con prácticas de mejora técnica y control de calidad. Aunque las pruebas unitarias automatizadas se reconocen como una mejora técnica futura, durante el desarrollo se aplicaron pruebas funcionales, pruebas de integración, pruebas smoke y validaciones operativas sobre los módulos críticos del sistema. Por ello, Scrum + XP fue considerado el enfoque más pertinente para el desarrollo de la propuesta tecnológica dentro del contexto institucional analizado.

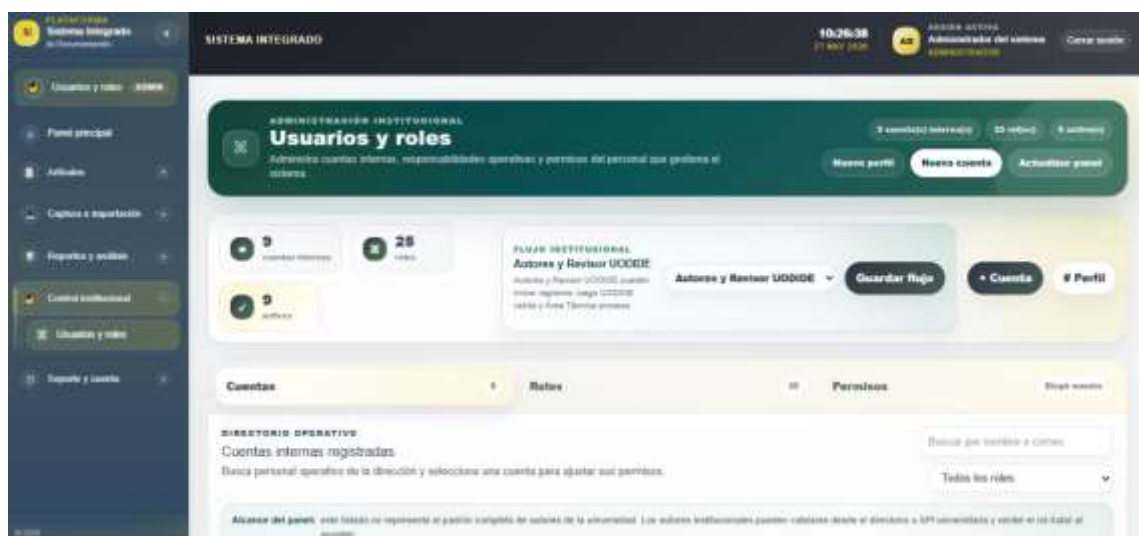
Anexo 11. Diseño de sistema.

Login



Panel Administrativo - Rol administrador

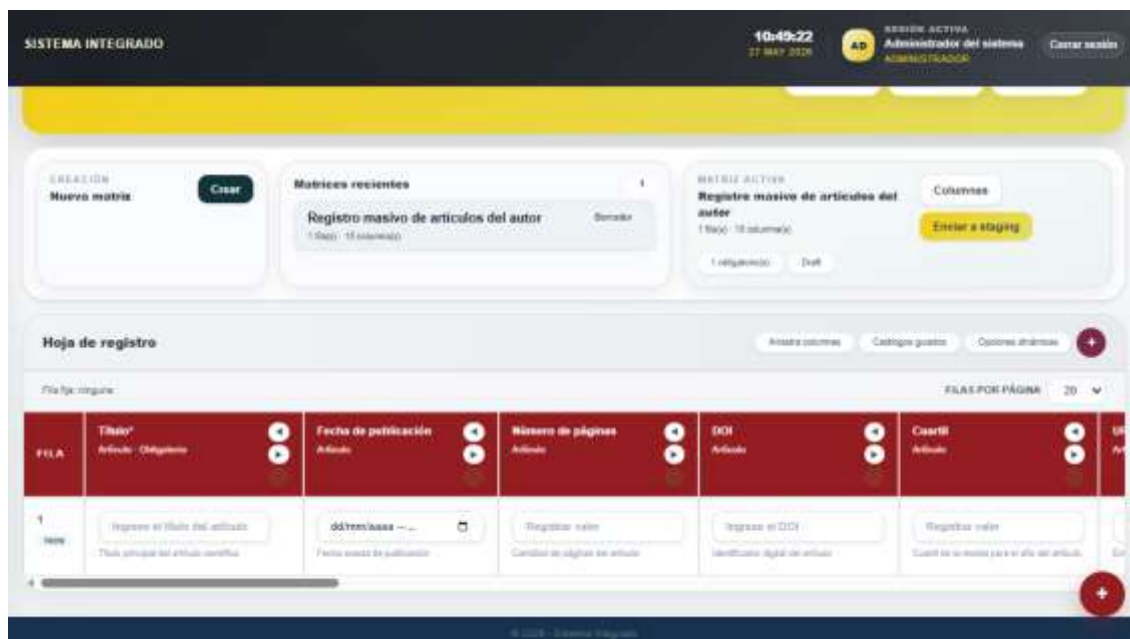




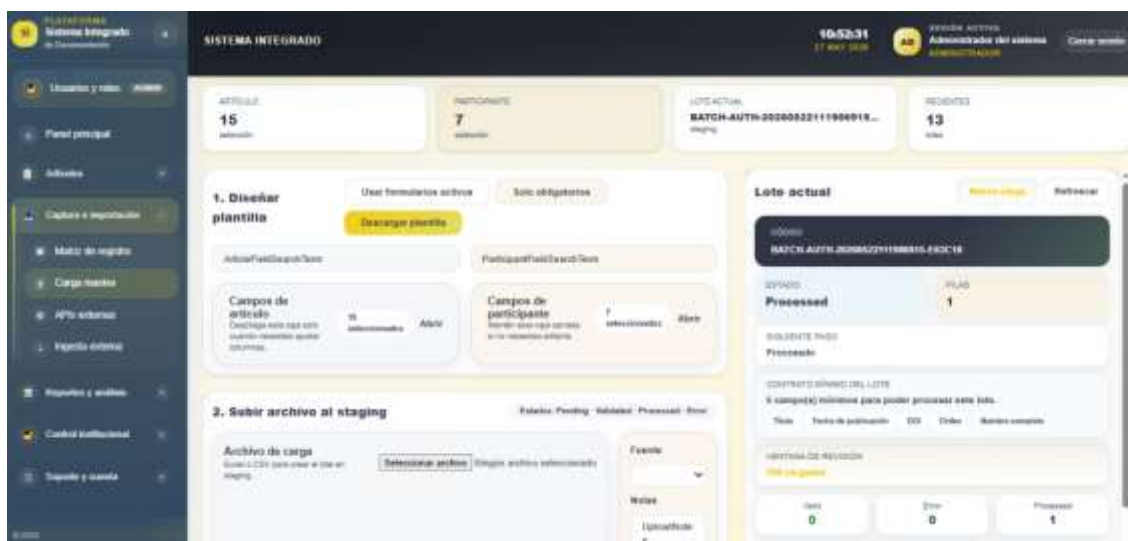
Registro de Artículos - Registro individual y en masa

The screenshot shows the 'Registro individual de artículos' form in the 'SISTEMA INTEGRADO'. The interface includes a sidebar with navigation options like 'Usuarios y roles', 'Panel principal', 'Artículos', 'Registrar artículo', 'Revisión de envío', 'Configuración', 'Lista de artículos', 'Captura e importación', 'Reportes y análisis', 'Control institucional', and 'Soporte y ayuda'. The main content area has a header with 'Registro individual' and 'Métrico de registro'. Below this, the 'FORMULARIO INSTITUCIONAL' section is titled 'Registro individual de artículos' and includes a description: 'Completa un artículo a la vez y luego envíalo a revisión. El formulario responde automáticamente a la configuración activa del sistema.' The form is divided into four main sections: 'ESTRUCTURA ACTIVA' (4 secciones disponibles), 'CAMPOS VISIBLES' (22 incluye participantes), 'COMPLETADOS' (1 con valor registrado), and 'PENDIENTES' (4 obligatorios por completar). A 'EXPLORADOR DEL FORMULARIO' on the right allows searching, expanding, or collapsing sections, and provides a 'Registrar manual de artículo' button. At the bottom, there are tabs for 'Identificación', 'General', 'Publicación', 'Clasificación', and 'Participantes', and a 'Resumen del artículo' section with a 'Publicar' button and a '00 obligatorios' indicator.

The screenshot shows the 'Resumen del registro' sidebar in the 'SISTEMA INTEGRADO'. The sidebar is titled 'Resumen del registro' and includes a 'Edición activa' button. It contains several action buttons: 'Enviar a revisión', 'Imprimir formulario', and 'Cancelar'. Below these, there are two columns of statistics: 'CAMPOS' (22) and 'REGISTRADOS' (1), and 'OBLIGATORIOS PENDIENTES' (4) and 'PARTICIPANTES' (1). The 'FORMULARIO' section indicates 'Registro manual de artículo' and '6 secciones activas'. At the bottom, there is a 'VISTA PREVIA' section with the text 'Sin título todavía'.



Carga Masiva de informacion





PLANTILLA

Sistema Integrado de Investigación

12:24:13

07 MAY 2024

AR

USUARIO ACTIVO

Administrador del sistema

ADMINISTRACION

Cerrar sesión

Sistema Integrado

Usuarios y roles

Panel principal

Activados

Cargas e importación

Módulo de registro

Carga masiva

API's externas

Registros actuales

Reportes y estadísticas

Control institucional

Reporte y consulta

INVESTA INSTITUCIONAL

Carga masiva Scopus

Recupera productos científicos por filitacion, analitica registral y normaliza datos externos en el formato de registro

PASO 1 DE 4

Scopus por filitacion

Configura la filitacion y prepara el dataset recuperado desde Scopus

FILITACION INSTITUCIONAL

Universidad Tecnica de Colombia

TIPO DE REGISTRO

250 negativos

☐

Validar el error

Opcional para validar datos. Si guardas automaticamente se validan antes de registrar

Preparar dataset

Preparar registro

Normalizar registro

RECONSIDERAR

0

Activar en vista previa

LOTES ESTIMADOS

0

Seguir proceso actual

CON DATOS

0

Algunas estadísticas

SIN DATOS

0

Resumen de estadísticas finales

VISTA PREVIA

Dataset Scopus recuperado

Se muestra el resumen

Dataset pendiente

Prepara la carga institucional para ver la muestra de datos recuperados desde Scopus.

0:00

0:00

Validaciones y seguimiento

The screenshot shows the 'Validaciones y seguimiento' section of the 'SISTEMA INTEGRADO' dashboard. The left sidebar contains navigation links: 'Usuarios y roles', 'Panel principal', 'Artículos', 'Registrar artículo', 'Estado de artículos', 'Configuración', 'Estado de artículos', 'Captura e importación', 'Reportes y análisis', 'Control institucional', and 'Soporte y ayuda'. The main content area is divided into three panels:

- PANEL DE VALIDACIÓN:** 'Validación UODIDE'. It includes a message: 'No hay workflows para UODIDE. Cuando se envíe llegar a la etapa de validación UODIDE, aparecerá aquí para revisar y aprobar.' Below this is a 'VALIDACIÓN TÉCNICA EN CURSO' section for 'Docente Autor Demo - 1 artículo', showing a 'Registrado por' field and a 'Registrado por' field.
- PANEL DE VISIÓN:** 'Validación Área Técnica'. It includes a message: 'Permisos flex, transacción flex, aprobables y procesamiento de datos.' Below this is a 'VALIDACIÓN TÉCNICA EN CURSO' section for 'Docente Autor Demo - 1 artículo', showing a 'Registrado por' field and a 'Registrado por' field.
- ACTA:** 'BATCH-AUTH-20200522111906015-E63C10'. It includes a 'Procesado' button and a 'Procesado' button.

Formularios y metadatos

The screenshot shows the 'Formularios y metadatos' section of the 'SISTEMA INTEGRADO' dashboard. The left sidebar contains navigation links: 'Usuarios y roles', 'Panel principal', 'Artículos', 'Registrar artículo', 'Estado de artículos', 'Configuración', 'Estado de artículos', 'Captura e importación', 'Reportes y análisis', 'Control institucional', and 'Soporte y ayuda'. The main content area is divided into two panels:

- ADMINISTRACIÓN SIMBÓLICA:** 'Formularios y metadatos'. It includes a message: 'Definir formularios, asignar campos y controlar qué se elige al registrar datos.' Below this are four statistics: 'FORMULARIOS' (2), 'CAMPOS' (40), 'ASIGNACIONES' (15), and 'CATÁLOGO ACTIVO' (Campos simples).
- NAVIGACIÓN DEL MODELO:** It includes a 'Formulario' button and a 'Formulario' button.

Vista de artículos

PLATONIAS
 Sistema Integrado de Investigaciones

12:42:00

27 MAR 2020

ADMINISTRADOR

Administrador del sistema

Cerrar sesión

Inicio y roles

Panel principal

Artículos

Registrar artículo

Revisión de envío

Configuración

Listado de artículos

Exportar e imprimir

Reportes y estadísticas

Control institucional

Reporte y mensajería

SISTEMA INTEGRADO

REPOSITARIO INSTITUCIONAL

Artículos registrados

TOTAL FOLGADOS

1000

De 1000 registros.

PUBLICACIONES

0

Subidos hoy

DEPTEROS

0

Recibido o publicados

SIN ESTADO

1000

No revisados

EXTRAMETADATOS BÁSICOS

Estados de publicación

ESTADO

Todo

sin estado

Búsqueda por Folio, nombre, DOI...

Crear

Todo

Refrescar

Revisar Registro

A la página siguiente

Mostrando 5-6 de 1000 artículos

Por página

10

Anterior

Página 1 de 167

Siguiente


PLATAFORMA
 Sistema Integrado de Investigación

SISTEMA INTEGRADO


 Usuario y rol: **ADMIN**

Panel principal

Aktivitas

Registrar artículo

Revisión de artículos

Configuración

Listado de artículos

Captura e importación

Reportes y análisis

Control institucional

Soporte y cuenta

Tienda

12:42 PM

21 MAY 2023

ADMINISTRADOR DEL SISTEMA

ADMINISTRACION

Cerrar sesión

Mostrando 5-8 de 1000 artículos

Por página: 5 ▼

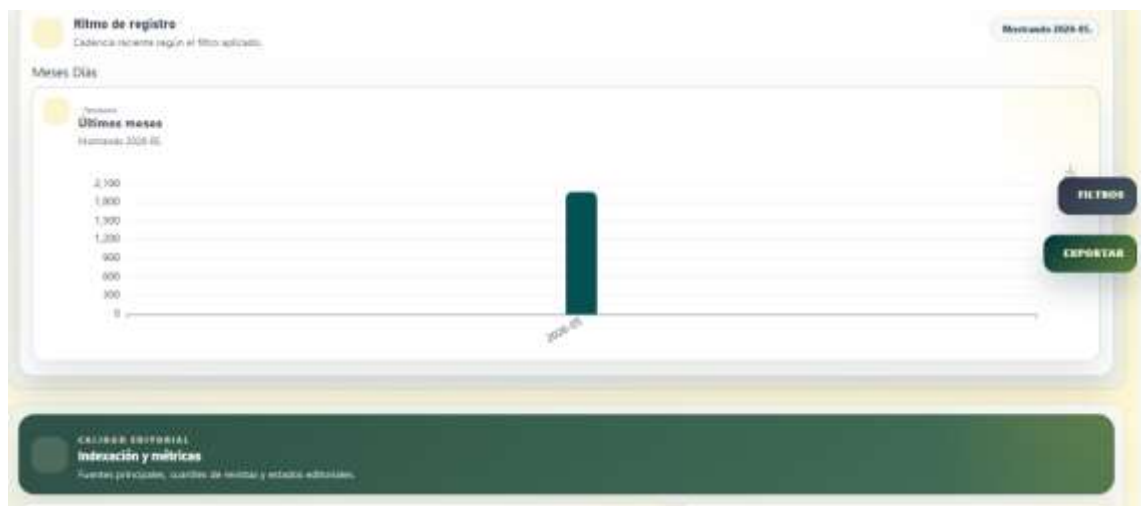
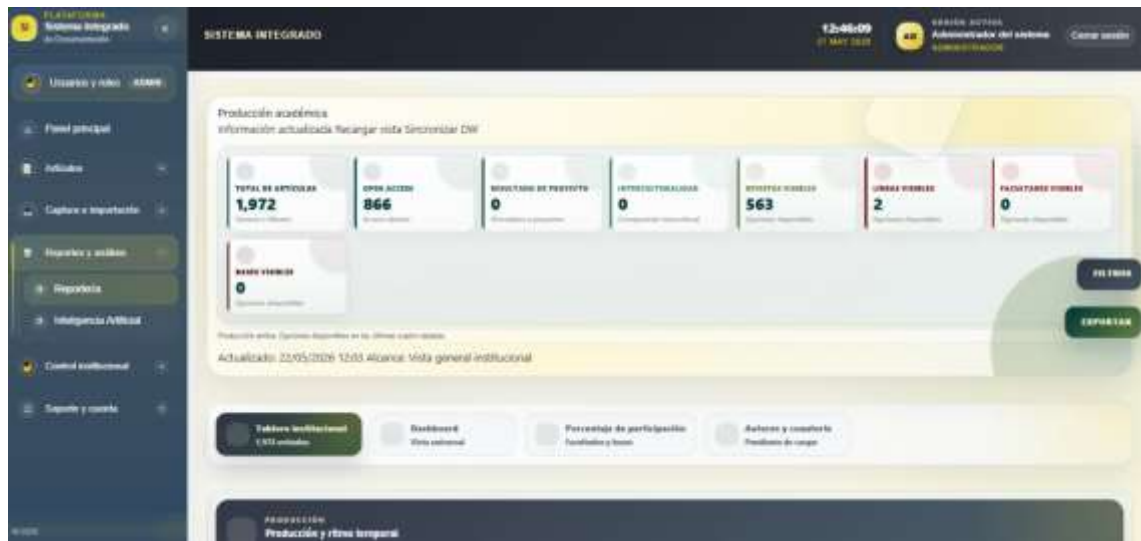
Continuar

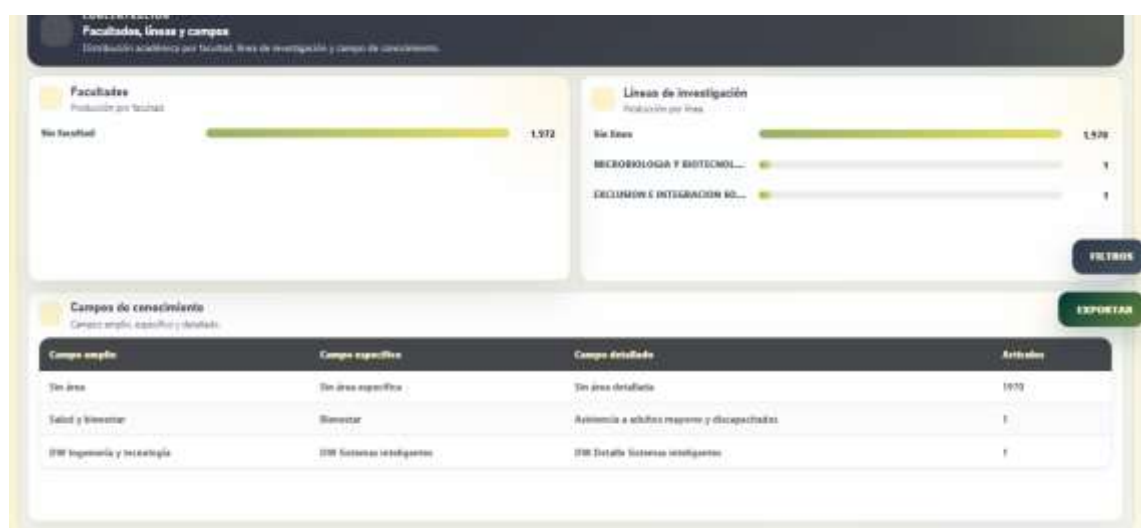
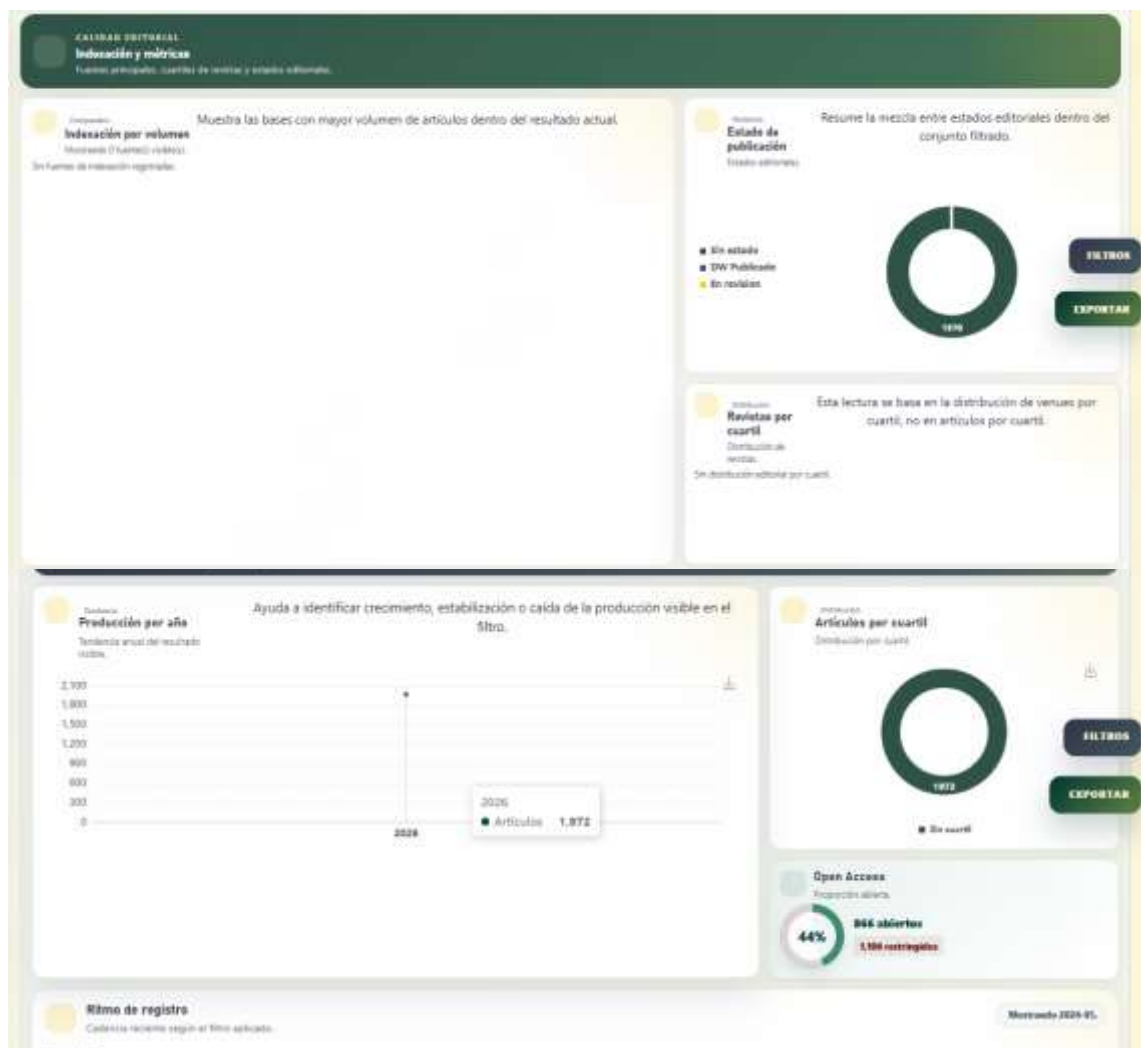
Página 1 de 160

Seguir

TÍTULO PUBLICACIÓN	ID#	ESTADO	REVIEWER - EXENTO	FECHA	TUBOS DE INDICACIÓN	ACCIONES
Competitionally tunable NiAl temperature in Mn _{1-x} Co _x : A route to enhance magnetoelectr...	15.10710(1) 2023-04-01	En proceso	Journal of Materials Science and Technology	2023-04-01	No expandido	Ver Editar Eliminar
A Conceptual Model to Integrate Operations Strategies Into Circular and Sustainable Global Supply Chains	15.10710(1) 2023-04-01	En proceso	Sustainable Development	2023-04-01	No expandido	Ver Editar Eliminar
A Modular BQS Framework for Enhanced KIRA YouTube Navigation	15.10710(1) 2023-04-01	En proceso	Lecture Notes in Networks and Systems	2023-04-01	No expandido	Ver Editar Eliminar
Academician's interest and technology acceptance model in University	15.10710(1) 2023-04-01	En proceso	Revista Colombiana De Ciencias	2023-04-01	No expandido	Ver Editar Eliminar
Accounting: Science, Technique or Technology?	15.10710(1) 2023-04-01	En proceso	Lecture Notes in Networks and Systems	2023-04-01	No expandido	Ver Editar Eliminar
Acoustic Performance of Ecuadorian Construction System: From Analysis to Regulatory Implementation	15.10710(1) 2023-04-01	En proceso	Lecture Notes in Networks and Systems	2023-04-01	No expandido	Ver Editar Eliminar

Reportería





BLVDOT ESPañOL
PEDI BIT
Artículos por PED BIT

Las filas incluyen indexaciones y pendientes.

PEDI BIT
Detalle reportable

1,972 artículos indexados : 1,972 filas reportables

Resumen Tabla

Título de artículo	Base de datos	Revista / ISSN	DOI	Autor	Última de actualización	Facultad	Mes de publicación	Proyecto	No. de
Competitively tunable NiO temperature in Mn1-xCoM: A route to enhance magnetic-ion exchange bias control	En indexación	Journal of Materials Science and Technology	10.1016/j.jmst.2026.08.018	López-Pérez M. Autor principal	Sin datos	Sin facultad	2027-01	No	1
The Effect of Core Stability Training on Shooting Accuracy and Balance in Paraquet Athletes	En indexación	International Journal of Sport Studies for Health	10.61835/ijsssh.2026.002	Ramírez E. Autor principal	Sin datos	Sin facultad	2026-12	No	1
Characterization and evaluation of a phage cocktail targeting <i>Salmonella enterica</i> in a Turkey farm	En indexación	BMC Microbiology	10.1185/12866-020-04043-8	Vela-Chavez M.G. Autor principal	Sin datos	Sin facultad	2025-12	No	1
Comparative cardiometabolic effects of high, moderate, and low intensity exercise in polygenic obesity: a systematic review and network meta-analysis of randomized controlled trials	En indexación	BMC Women's Health	10.1185/12866-020-04043-8	Sánchez M. Autor principal	Sin datos	Sin facultad	2026-12	No	1

FILTROS
EXPORTAR

BLVDOT ESPañOL
TOD Total
Publicaciones por TOD Total

TOD Total
Detalle reportable

1,972 artículos indexados : 1,972 filas reportables

Resumen Tabla

Título de publicación	Base de datos	Revista / ISSN	Autor	Estado	Facultad	Revisión pendiente	Monitoreo pendiente	Mes
Competitively tunable NiO temperature in Mn1-xCoM: A route to enhance magnetic-ion exchange bias control	En indexación	Journal of Materials Science and Technology	López-Pérez M. Autor principal	En estado	Sin facultad	No	No	2026-12
The Effect of Core Stability Training on Shooting Accuracy and Balance in Paraquet Athletes	En indexación	International Journal of Sport Studies for Health	Ramírez E. Autor principal	En estado	Sin facultad	No	No	2026-12
Characterization and evaluation of a phage cocktail targeting <i>Salmonella enterica</i> in a Turkey farm	En indexación	BMC Microbiology	Vela-Chavez M.G. Autor principal	En estado	Sin facultad	No	No	2026-12
Comparative cardiometabolic effects of high, moderate, and low intensity exercise in polygenic obesity: a systematic review and network meta-analysis of randomized controlled trials	En indexación	BMC Women's Health	Sánchez M. Autor principal	En estado	Sin facultad	No	No	2026-12
Extraction of bioactive compounds from natural variety cacao shell (Theobroma Cacao L.) obtained using supercritical CO2	En indexación	Multidisciplinary Science Journal	Carrero-Rojas W.X. Autor principal	En estado	Sin facultad	No	No	2026-08

FILTROS
EXPORTAR

BLVDOT ESPañOL
Seguimiento operativo
Revisión y estado de los artículos

Revisión institucional: filas según filtro.

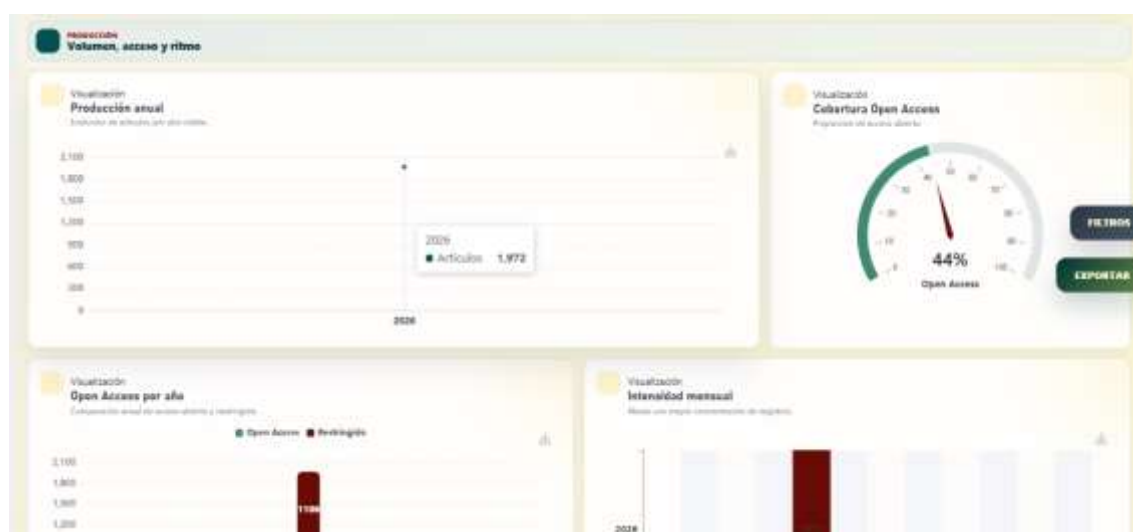
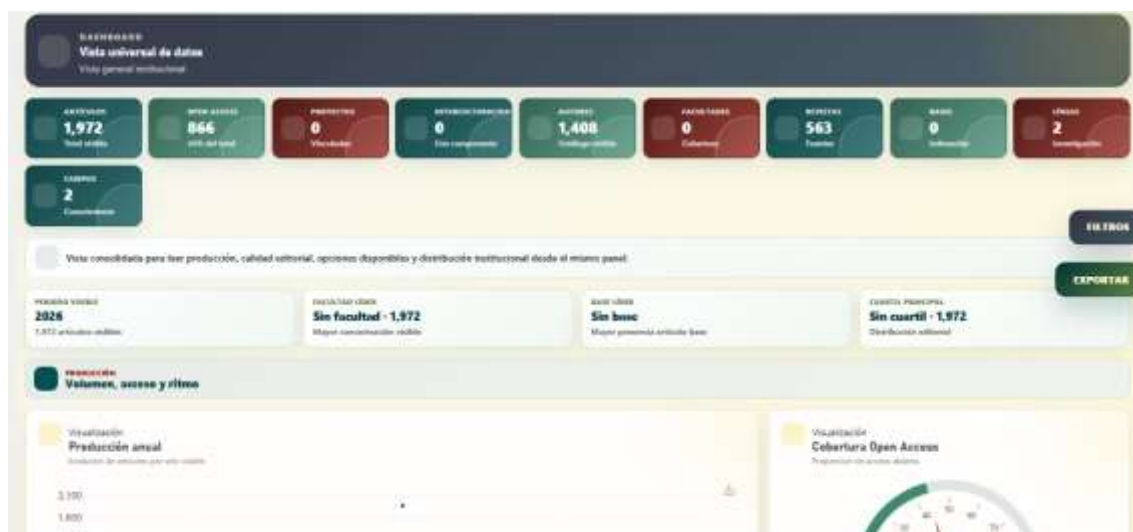
Seguimiento y detalle
Revisión, revisión y artículos

Revisión Revisión Artículo
Mostrando 12 artículos recientes:

Título	Año	Estado	Última	Revista	IA
Competitively tunable NiO temperature in Mn1-xCoM: A route to enhance magnetic-ion exchange bias control	2027	Sin estado	Sin datos	Journal of Materials Science and Technology	1
The Effect of Core Stability Training on Shooting Accuracy and Balance in Paraquet Athletes	2026	Sin estado	Sin datos	International Journal of Sport Studies for Health	1
Characterization and evaluation of a phage cocktail targeting <i>Salmonella enterica</i> in a Turkey farm	2025	Sin estado	Sin datos	BMC Microbiology	1
Comparative cardiometabolic effects of high, moderate, and low intensity exercise in polygenic obesity: a systematic review and network meta-analysis of randomized controlled trials	2026	Sin estado	Sin datos	BMC Women's Health	1
Extraction of bioactive compounds from natural variety cacao shell (Theobroma Cacao L.) obtained using supercritical CO2	2026	Sin estado	Sin datos	Multidisciplinary Science Journal	1
Effectiveness of telemedicine interventions for the management of gestational diabetes in primary care centers	2026	Sin estado	Sin datos	Admission Primaria	1

FILTROS
EXPORTAR

Dashboard



Bloque jerárquico de autores

The screenshot shows a user profile for 'Celia G.' with 27 publications. The page includes a header with the user's name and a 'Seleccionar' button. Below the header, there are tabs for 'Publicaciones' and 'Resumen'. The main content area displays a list of publications, each with a title, a 'Publicado' status, and a date. A sidebar on the right contains a 'FILTROS' section with a 'Seleccionar' button and an 'EXPORTAR' button. The publications are listed in a table format with columns for title, status, and date.

Publicaciones	Resumen
PUBLICACIONES REGISTRADAS	
27 publicaciones listadas	
Continuar	
FILTROS	
Seleccionar	
EXPORTAR	
Applying Deep Q Networks to Local Route Optimization	01/01/2020
Publicado	Celia G.
Navigating Governmental Choices: A Comprehensive Review of Artificial Intelligence's Impact on Decision-Making	01/01/2020
Publicado	Celia G.
Applying Deep Q Networks to Local Route Optimization	01/01/2020
Publicado	Celia G.
Lean Manufacturing Tools for Industrial Process: A Literature Review	01/01/2020
Publicado	Celia G.
Low-Cost Fuzzy Control for Faulty Hearing Systems	01/01/2020
Publicado	Celia G.
Problems with Health Information Systems in Ecuador, and the Need to Educate University Students in Health Informatics in Times of Pandemic	01/01/2020
Publicado	Celia G.

Filtros

Rango y referencia temporal

REFERENCIA TEMPORAL DE SERIES
Fecha de registro ▼

CREACIÓN DESDE
dd/mm/aaaa 📅

CREACIÓN HASTA
dd/mm/aaaa 📅

PUBLICACIÓN DESDE
dd/mm/aaaa 📅

PUBLICACIÓN HASTA
dd/mm/aaaa 📅

FILTROS

Contexto institucional
Estado, línea, facultad y base

Campos de conocimiento
Amplio, específico y detallado

Revista y calidad editorial
Revista, cuartil y acceso abierto

Participación autoral
Filtra toda la reportería por autor, coautor, filiación o rol

EXPORTAR

Cerrar Limpiar **Aplicar filtros**

Exportar reportería
PDF, Excel y dataset bruto para análisis externo.

PDF institucional

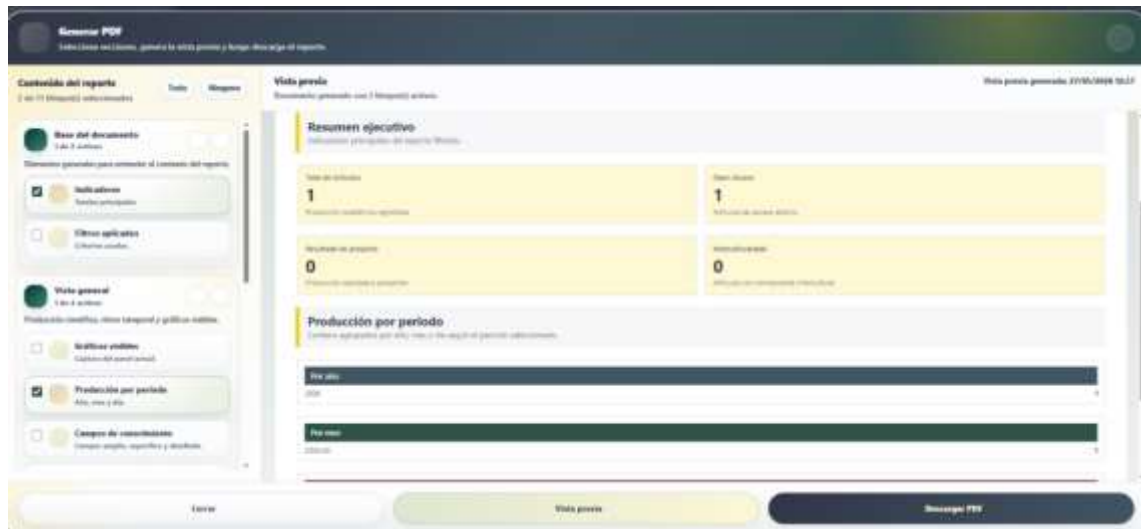
PDF de autores

Excel institucional

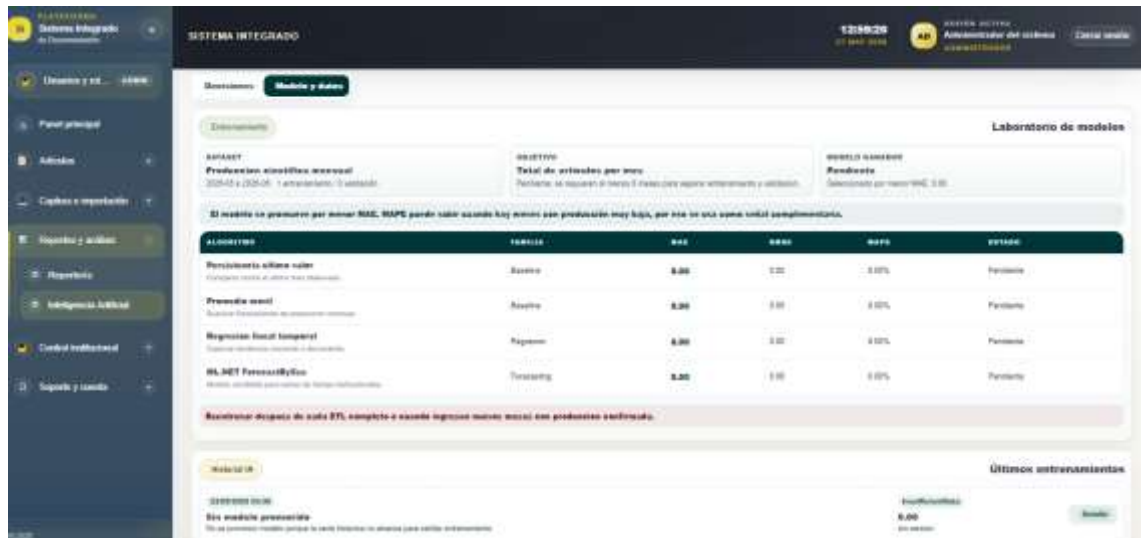
Dataset Excel

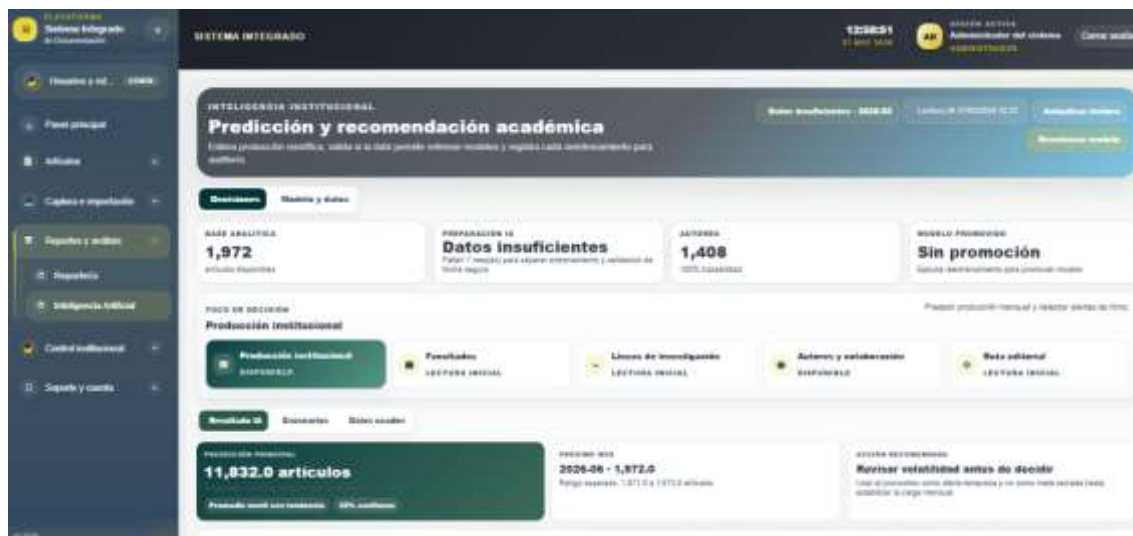
Dataset CSV ZIP

Modulo constructor de PDF



Módulo de inteligencia artificial





Anexo 12. Informe técnico del módulo de Inteligencia Artificial.

1. Propósito del módulo IA

El módulo de Inteligencia Artificial tiene como finalidad apoyar la toma de decisiones institucionales sobre la producción científica registrada. Su enfoque no reemplaza el criterio humano, sino que transforma los datos históricos de artículos, autores, indexaciones, facultades, líneas de investigación y workflow en señales útiles para planificación, seguimiento y mejora continua.

2. Componentes implementados

El módulo se integra al sistema mediante servicios backend, controladores API, objetos de transferencia de datos y una interfaz de visualización. Esta estructura permite consultar información consolidada, generar predicciones, presentar recomendaciones y registrar historial de entrenamiento.

Componente	Función dentro del sistema
InstitutionalIntelligenceService.cs	Servicio principal encargado de procesar datos, ejecutar predicciones, generar recomendaciones y coordinar entrenamiento.
IntelligenceController.cs	Controlador API que expone endpoints para dashboard IA, entrenamiento e historial.

InstitutionalIntelligenceDto.cs	DTOs utilizados para transportar predicciones, diagnósticos, métricas y recomendaciones.
IntelligenceInsights.razor	Interfaz frontend donde el usuario visualiza dashboard IA, tendencias, recomendaciones y resultados de entrenamiento.

3. Predicciones generadas por el sistema

El sistema realiza predicciones relacionadas con la producción científica institucional, utilizando información consolidada en el Data Warehouse y disponible para reportería institucional. El horizonte de predicción definido es de seis meses.

Predicción	Variable principal	Uso institucional
Producción científica mensual	Total de artículos por mes	Anticipar si la producción científica esperada sube, baja o se mantiene.
Predicción por facultad	Artículos históricos y tendencia por facultad	Identificar unidades académicas con mayor o menor ritmo de producción y priorizar seguimiento.
Predicción por línea de investigación	Producción histórica por línea	Detectar líneas activas, estables o en descenso para apoyar planificación investigativa.

4. Salidas del módulo predictivo

- Mes proyectado.
- Valor esperado de artículos.
- Límite inferior de confianza.
- Límite superior de confianza.
- Algoritmo utilizado.
- Porcentaje de confianza.
- Interpretación del resultado.

5. Recomendaciones generadas por el sistema

El módulo genera recomendaciones explicables basadas en reglas, indicadores y resultados predictivos. Estas recomendaciones convierten los resultados del análisis en acciones comprensibles para los usuarios institucionales.

Tipo de recomendación	Criterios utilizados	Ejemplos de acciones sugeridas
Calidad del dataset	Volumen histórico, meses disponibles, completitud de autores, ORCID e indexación.	Ampliar base histórica, completar filiación, normalizar ORCID e identificación.
Predicción	Confianza baja, alta volatilidad o pocos datos mensuales.	Usar pronóstico como alerta temprana y comparar producción real con rango esperado.
Editorial	Bases de indexación, cuartiles, Open Access, revistas y métricas editoriales.	Normalizar cuartiles, fortalecer estrategia Open Access y mejorar cobertura de indexación.
Autores y colaboración	Producción por autor, coautoría, filiación y trazabilidad autoral.	Identificar nodos de colaboración y completar información autoral.
Workflow	Devoluciones, etapas repetidas, duración de revisión y estados finales.	Preparar futuro modelo de riesgo de demora y priorizar casos complejos.

6. Algoritmos utilizados

Algoritmo	Tipo	Aplicación en el sistema
Persistencia del último valor	Baseline	Usa el último mes observado como referencia mínima de comparación.
Promedio móvil de tres meses	Baseline interpretable	Suaviza fluctuaciones recientes y funciona cuando la serie histórica es pequeña.
Regresión lineal temporal	Regresión	Identifica tendencia creciente, descendente o estable en la serie mensual.
ForecastBySsa de ML.NET	Pronóstico de series temporales	Modelo principal cuando existen suficientes meses

		históricos; genera predicción a seis meses con intervalos de confianza.
--	--	---

7. Entrenamiento, validación y selección del modelo

El entrenamiento se ejecuta desde el backend mediante el servicio de Inteligencia Artificial. La variable objetivo corresponde al total de artículos por mes. Para conservar la lógica temporal de los datos, el sistema utiliza validación temporal, separando los primeros meses para entrenamiento y los últimos para validación.

Meses disponibles		Meses de validación
18 o más		6
12 a 17		4
8 a 11		3
Menos de 8		0
Métrica	Significado	Uso
MAE	Error absoluto medio.	Métrica principal para seleccionar el mejor modelo.
RMSE	Raíz del error cuadrático medio.	Métrica complementaria y criterio de desempate.
MAPE	Error porcentual absoluto medio.	Señal complementaria; debe interpretarse con cautela en meses con valores bajos.
Score	Aproximación basada en 100 - MAPE.	Indicador complementario de desempeño relativo.

El sistema selecciona el modelo con menor MAE y, en caso de empate, el menor RMSE. Si existe un modelo evaluado, se promueve como versión activa; si los datos son insuficientes, el estado se registra como InsufficientData.

8. Reentrenamiento, historial y auditoría

La política definida establece reentrenar el modelo después de cada ETL completo o cuando ingresen nuevos meses con producción confirmada. El sistema reconstruye la serie mensual, separa entrenamiento y validación, evalúa algoritmos candidatos, selecciona el menor MAE, genera versión del modelo, guarda historial e invalida la caché del dashboard IA.

Endpoint	Función
GET api/intelligence/dashboard	Devuelve predicciones, recomendaciones, diagnóstico y escenarios.
POST api/intelligence/training/run	Ejecuta reentrenamiento manual.
GET api/intelligence/training/history	Consulta historial de entrenamientos.
Elemento auditado	Descripción
RunId, fechas y estado	Identifican cada corrida de entrenamiento y su resultado.
Algoritmo promovido y versión activa	Permiten conocer qué modelo fue seleccionado.
Filas de entrenamiento y validación	Evidencian el volumen utilizado en cada proceso.
Métricas por algoritmo	Registran MAE, RMSE, MAPE y Score obtenidos.
Usuario ejecutor	Permite trazabilidad institucional del entrenamiento.

9. Escenarios IA contemplados

Escenario	Estado	Algoritmo o enfoque recomendado
Predicción de producción	Implementado	ML.NET ForecastBySsa, regresión y modelos baseline.
Recomendación editorial	Implementado por reglas explicables	Ranking por reglas y clasificación futura.
Riesgo de demora	Planificado	Clasificación binaria.
Calidad de datos	Implementado parcialmente por reglas	Reglas explicables y clasificación futura.

10. Aporte del módulo IA

- ✓ Predicción institucional de producción científica mensual.
- ✓ Predicción por facultad y por línea de investigación.
- ✓ Recomendaciones editoriales, de calidad de datos, autores y colaboración.
- ✓ Diagnóstico de preparación de datos.
- ✓ Comparación de algoritmos y selección automática del mejor modelo mediante MAE.
- ✓ Historial auditable de entrenamientos.
- ✓ Integración con Data Warehouse y reportería institucional.

Anexo 13. Informe del módulo de Inteligencia Artificial.

1. Objetivo del informe

Este informe documenta la base matemática y técnica del módulo de Inteligencia Artificial aplicado al sistema de gestión de producción científica. Su finalidad es respaldar metodológicamente las fórmulas, indicadores, métricas de evaluación, predicciones y reglas de recomendación utilizadas por el sistema.

2. Preparación general para IA

El sistema calcula un indicador global denominado OverallScore, el cual representa el nivel de preparación de los datos para ejecutar predicciones y recomendaciones confiables. Este indicador combina cinco dimensiones ponderadas: volumen de datos, trazabilidad de autores, cobertura de indexación, calidad de carga y señales de workflow.

Fórmula: $\text{OverallScore} = (\text{DatasetVolumeScore} \times 0,30) + (\text{AuthorTraceScore} \times 0,22) + (\text{IndexingCoverageScore} \times 0,20) + (\max(\text{LoadQualityScore}, 55) \times 0,14) + (\max(\text{WorkflowSignalScore}, 45) \times 0,14)$

Indicador	Peso	Propósito
Volumen de datos	30 %	Evalúa si existen suficientes artículos para análisis predictivo.
Trazabilidad de autores	22 %	Mide disponibilidad de información autoral.
Cobertura de indexación	20 %	Evalúa existencia de datos sobre fuentes, bases y cuartiles.
Calidad de carga	14 %	Mide proporción de registros procesados correctamente.
Señales de workflow	14 %	Evalúa evidencia operativa del proceso de revisión y devolución.
OverallScore	Nivel de preparación	
80 % a 100 %	Alta	
60 % a 79 %	Media	
40 % a 59 %	Inicial	

Menor a 40 %	Baja
--------------	------

3. Indicadores de preparación de datos

Indicador	Fórmula	Interpretación
AuthorTraceScore	$(\text{Artículos con trazabilidad de autor} / \text{Total de artículos}) \times 100$	Mide completitud de información autoral.
LoadQualityScore	$(\text{Filas exitosas} / \text{Total de filas procesadas}) \times 100$	Evalúa calidad de importación, carga masiva o staging.
IndexingCoverageScore	$(\text{Artículos con indexación} / \text{Total de artículos}) \times 100$	Mide disponibilidad de información editorial y científica.
WorkflowSignalScore	$\min(100, \text{TotalStageExecutions} - \min(30, \text{ReturnedStages}))$	Evalúa señales operativas del flujo de revisión.
SharePercent	$(\text{ValorSegmento} / \text{TotalArticulos}) \times 100$	Calcula participación por facultad, línea, indexación, cuartil u Open Access.

4. Predicción temporal de producción científica

La predicción temporal utiliza una serie mensual donde la variable objetivo corresponde al total de artículos por mes. El horizonte actual es de seis meses futuros. Cada punto de predicción contiene periodo, valor esperado, límite inferior, límite superior, algoritmo utilizado e interpretación del resultado.

Representación: $\text{Forecast}(t + h) = \text{artículos esperados para el mes futuro } h$, donde $h = 1, 2, 3, 4, 5, 6$.

5. Modelos matemáticos utilizados

Modelo	Fórmula o criterio	Uso
Persistencia del último valor	$\text{Predicción}(t + 1) = \text{Valor}(t)$	Línea base mínima para comparación.
Promedio móvil de tres meses	$(\text{Valor}(t) + \text{Valor}(t - 1) + \text{Valor}(t - 2)) / 3$	Suaviza fluctuaciones recientes.
Regresión lineal temporal	$\text{Predicción} = \text{intercepto} + \text{slope} \times \text{mesFuturo}$	Estima tendencia creciente, descendente o estable.

ForecastBySsa de ML.NET	horizon = 6; confidenceLevel = 0,95; seed = 7	Pronóstico de series temporales con intervalos de confianza.
-------------------------	---	--

Pendiente de regresión: $\text{slope} = (n \times \text{sum}(xy) - \text{sum}(x) \times \text{sum}(y)) / (n \times \text{sum}(x^2) - (\text{sum}(x))^2)$

Intercepto: $\text{intercepto} = \text{promedio}(y) - \text{slope} \times \text{promedio}(x)$

Regla de tendencia	Interpretación
slope > 0,25	Tendencia creciente
slope < -0,25	Tendencia descendente
Caso contrario	Tendencia estable

6. Confianza y volatilidad de la predicción

El sistema calcula un puntaje de confianza considerando la cantidad de meses históricos disponibles y la volatilidad de la serie mensual. Una serie más extensa aumenta la confianza, mientras que una serie más volátil reduce el puntaje.

Meses históricos	Confianza base
24 o más	85 %
18 a 23	72 %
12 a 17	60 %
6 a 11	45 %
Menos de 6	25 %

$\text{VolatilityPenalty} = (\text{Volatilidad} / \max(1, \text{Promedio})) \times 45$

$\text{ConfidenceScore} = \text{VolumeScore} - \text{VolatilityPenalty}$, limitado entre 15 % y 95 %.

$\text{Volatilidad} = \sqrt{\text{sum}((x - \text{promedio})^2) / n}$

7. Separación entrenamiento-validación

El sistema utiliza validación temporal y conserva el orden cronológico de los datos. Esta decisión evita mezclar registros aleatoriamente, lo cual no es recomendable en series temporales.

Meses disponibles	Meses de validación	Justificación
18 o más	6	Permite validación más amplia.
12 a 17	4	Permite validación media.

8 a 11	3	Permite validación inicial.
Menos de 8	0	Datos insuficientes para validación temporal.

8. Métricas de evaluación predictiva

Métrica	Fórmula	Interpretación
MAE	$\text{sum}(\text{Real} - \text{Predicho}) / n$	Error promedio en unidades reales; métrica principal.
RMSE	$\text{sqrt}(\text{sum}((\text{Real} - \text{Predicho})^2) / n)$	Penaliza errores grandes; criterio de desempate.
MAPE	$(\text{sum}((\text{Real} - \text{Predicho}) / \text{Real}) \times 100) / n$	Error porcentual medio; sensible cuando Real es bajo o cercano a cero.
Score	$\text{max}(0, 100 - \text{MAPE})$	Indicador complementario de desempeño relativo.

La selección del mejor modelo se realiza ordenando por menor MAE y, en caso de empate, por menor RMSE. Formalmente: $\text{BestModel} = \text{argmin}(\text{MAE}, \text{RMSE})$.

9. Reglas de recomendación

Condición	Prioridad	Recomendación generada
TotalArticles < 250	Alta	Ampliar base histórica para mejorar modelos segmentados.
ForecastConfidence < 40	Alta	Revisar volatilidad antes de utilizar la predicción como referencia fuerte.
AuthorTraceScore < 90	Alta	Fortalecer trazabilidad autoral, filiación y ORCID.
IndexingCoverageScore < 80	Media	Normalizar fuentes, cuartiles y métricas editoriales.
OpenAccessShare < 60	Alta	Mejorar estrategia de acceso abierto institucional.

10. Valores clave actuales del módulo

Elemento	Valor
Horizonte de predicción	6 meses
Nivel de confianza SSA	95 %
Meses mínimos para validar	8

Meses recomendados	12
Semilla ML.NET	7
Métrica principal	MAE
Métrica de desempate	RMSE
Penalización por volatilidad	Hasta 45 puntos
Confianza mínima	15 %
Confianza máxima	95 %
Umbral Open Access	60 %
Umbral trazabilidad autores	90 %
Umbral cobertura indexación	80 %
Umbral baja confianza forecast	40 %
Umbral volumen bajo	250 artículos

11. Fórmulas principales consolidadas

Componente	Fórmula
OverallScore	$(\text{DatasetVolumeScore} \times 0,30) +$ $(\text{AuthorTraceScore} \times 0,22) +$ $(\text{IndexingCoverageScore} \times 0,20) +$ $(\max(\text{LoadQualityScore}, 55) \times 0,14) +$ $(\max(\text{WorkflowSignalScore}, 45) \times 0,14)$
AuthorTraceScore	$(\text{ArtículosConTrazabilidadDeAutor} /$ $\text{TotalArtículos}) \times 100$
LoadQualityScore	$(\text{FilasExitosas} / \text{TotalFilasProcesadas}) \times 100$
IndexingCoverageScore	$(\text{ArtículosConIndexación} / \text{TotalArtículos}) \times 100$
WorkflowSignalScore	$\min(100, \text{TotalStageExecutions}) - \min(30,$ $\text{ReturnedStages})$
PromedioMovil3	$(\text{Valor}(t) + \text{Valor}(t - 1) + \text{Valor}(t - 2)) / 3$
slope	$(n \times \sum(xy) - \sum(x) \times \sum(y)) / (n \times \sum(x^2) -$ $(\sum(x))^2)$
MAE	$\sum(\text{Real} - \text{Predicho}) / n$
RMSE	$\sqrt{\sum((\text{Real} - \text{Predicho})^2) / n}$
MAPE	$(\sum((\text{Real} - \text{Predicho}) / \text{Real}) \times 100) / n$
Score	$\max(0, 100 - \text{MAPE})$
SharePercent	$(\text{ValorSegmento} / \text{TotalArtículos}) \times 100$
ConfidenceScore	$\text{VolumeScore} - ((\text{Volatilidad} / \max(1, \text{Promedio}))$ $\times 45)$

Anexo 14. Modelo dimensional tipo constelación implementado en el Data Warehouse institucional.

