



Fortalecimiento del Sistema de Información de la Extensión Universitaria

Facultad de Economía – Universidad del Rosario

Consultores: Samuel Blanco Castellanos y Cristian Beltrán Mendez

Asignatura: Proyecto Práctico

Profesor: Juan David Vargas Velez

Fecha: Bogotá D.C., 19 de noviembre de 2025

*Documento elaborado en el marco del proyecto académico de Proyecto Práctico,
correspondiente al diseño y desarrollo del sistema de información para la extensión
universitaria de la Facultad de Economía.*

Índice

1	Contextualización	4
1.1	Descripción del cliente	4
1.2	Entorno: Facultad de Economía	5
2	Formulación del problema	8
3	Objetivos del proyecto	8
3.1	Objetivo general	8
3.2	Objetivos específicos	9
4	Alcance del proyecto	9
4.1	Alcance abordado en este semestre	10
4.2	Aspectos que se podrían abordar en etapas posteriores	10
4.3	Delimitación institucional	11
5	Metodología del proyecto	11
5.1	Enfoque general	11
5.2	Fases del proyecto	11
5.3	Técnicas e instrumentos utilizados	12
5.4	Proceso iterativo de trabajo	13
6	Diagnóstico del Estado Actual del Sistema de Información	13
7	Revisión de literatura	18
7.1	Sistemas de información universitarios	18
7.2	Experiencias institucionales: aprendizajes comparados	20
7.3	Referentes técnicos sobre arquitectura, calidad y gobernanza del dato	27
7.4	Síntesis de aportes	28
8	Visualización del sistema de información: punto de partida del piloto	29
8.1	Componentes esenciales del sistema	29
8.2	Lógica de funcionamiento del sistema	30
8.3	Fuentes de información	30
8.4	Proceso de extracción, transformación y carga	31
8.5	Almacenamiento estructurado y modelamiento de la base de datos	33
8.6	Diccionario institucional de datos	34
8.7	Consultas y mecanismos de recuperación de información	35
8.8	Visualización analítica e indicadores iniciales	36

9	Ejecución del piloto del sistema de información	37
9.1	Selección de tablas para la demostración del modelo	37
9.2	Requisitos funcionales del sistema	38
9.3	Requisitos Funcionales de ALTO nivel	39
9.4	Requisitos funcionales de Bajo Nivel	42
10	Modelo Entidad–Relación (MER)	49
10.1	Identificación de entidades	50
10.2	Relaciones principales del modelo	51
10.3	Cardinalidades y estructura lógica del modelo	52
10.4	Componentes del modelo MER	52
10.5	Derivación del diccionario de datos a partir del MER	53
10.6	Ejecución del piloto técnico	56
11	Visualización de Datos e Indicadores	57
12	Recomendaciones finales y siguientes pasos	58
12.1	1. Fortalecimiento estructural inicial (0–3 meses)	58
12.2	2. Construcción del proceso ETL institucional	58
12.3	3. Escalamiento del modelo relacional	59
12.4	4. Exploración de bases de datos no relacionales	59
12.5	5. Implementación de modelos analíticos y métricas de impacto	60
12.6	6. Escalamiento, automatización y evolución del sistema	60

Síntesis del Proyecto

El origen de esta consultoría se encuentra en los Términos de Referencia emitidos por la Dirección de Desarrollo y Extensión de la Facultad de Economía, en los que se planteó la necesidad de revisar, comprender y fortalecer el sistema de información que soporta las actividades de extensión de la Facultad. La motivación principal surge de una inquietud institucional: la extensión cumple un papel estratégico en la relación de la Facultad con su entorno, pero la información que la describe, la organiza y la documenta no se encuentra articulada de manera que permita analizarla, comunicarla o utilizarla con criterios de eficiencia y trazabilidad. El propósito de esta consultoría es, por tanto, aportar claridad sobre la situación actual, identificar oportunidades de mejora y avanzar hacia una propuesta que sienta las bases para un sistema de información más coherente, ordenado y orientado al uso estratégico de los datos.

De acuerdo con los Términos de Referencia, el proyecto se organizó en tres fases secuenciales: diagnóstico, análisis comparado y formulación de una propuesta de fortalecimiento. Esta estructura metodológica definió el desarrollo del trabajo a lo largo del semestre y permitió construir un proceso acumulativo donde cada fase aportó insumos para la siguiente. La primera fase se concentró en el diagnóstico inicial, orientado a reconocer el funcionamiento actual del sistema de información del área de Extensión, los datos que se recolectan, las fuentes utilizadas, los formatos de almacenamiento y los indicadores que hoy se reportan. Esta etapa permitió delimitar el enfoque de la consultoría a partir de una pregunta central: *¿en qué estado se encuentra hoy el sistema de información asociado a la extensión y cuáles son los factores que explican su configuración actual?*

El diagnóstico requirió examinar de manera sistemática las bases que emplea el área, los indicadores en uso, los mecanismos de registro de información y los flujos internos que sustentan los distintos procesos. La aproximación se basó en entrevistas semiestructuradas, revisión documental y análisis directo de la estructura de datos disponible. El resultado fue una caracterización de la situación actual que permitió identificar los principales retos para consolidar un sistema de información más estandarizado, articulado y apto para apoyar la gestión operativa de la Facultad.

Concluida la primera etapa, el proyecto avanzó hacia una segunda fase centrada en un análisis comparado y una revisión de literatura especializada. Esta revisión tuvo como finalidad comprender cómo abordan otras universidades la gestión de información de sus áreas de extensión, qué sistemas utilizan, qué principios de gobernanza de datos aplican y qué modelos conceptuales se han desarrollado para integrar procesos académicos, administrativos y misionales. Para ello se estudiaron documentos institucionales, investigaciones académicas y referentes internacionales sobre sistemas de información universitarios y prácticas de manejo y estandarización de datos. Este ejercicio permitió identificar elementos replicables, prácticas

ausentes en el contexto actual y oportunidades concretas de mejora.

La tercera fase del proyecto se dedicó a formular una propuesta inicial de fortalecimiento del sistema de información, integrando los hallazgos del diagnóstico y los insumos conceptuales derivados de la revisión de literatura. El propósito no fue establecer una solución definitiva, sino construir una ruta preliminar que articule las necesidades identificadas y trace posibles líneas de acción hacia un sistema más integrado, con mayor trazabilidad y con mejores condiciones para la medición del impacto de las actividades de extensión. Esta propuesta constituye un primer acercamiento técnico que podrá desarrollarse y profundizarse en fases posteriores.

En conjunto, el documento desarrolla con mayor detalle las actividades realizadas en cada fase, los insumos analizados y la manera en que estos elementos convergen en una recomendación alineada con los requerimientos de la Dirección de Desarrollo y Extensión. El proyecto constituye así un primer paso para comprender de manera más clara la información de extensión, organizarla de forma más eficiente y avanzar hacia una arquitectura que permita a la Facultad gestionar, analizar y comunicar de manera más adecuada el alcance de su labor.

1. Contextualización

1.1. Descripción del cliente

La consultoría se desarrolla para la **Dirección de Desarrollo y Extensión de la Facultad de Economía de la Universidad del Rosario**, unidad encargada de coordinar, registrar y articular las actividades de extensión de la Facultad. Esta dirección integra procesos relacionados con educación continuada, educación digital, proyectos de consultoría, relaciones con egresados, voluntariados, eventos académicos y acciones de proyección externa, constituyéndose en el principal enlace entre la actividad académica y el entorno institucional y social.

La Dirección de Desarrollo cumple una función estratégica al consolidar la información asociada a estas actividades y asegurar su articulación con aliados externos, dependencias de la Universidad y la comunidad académica. Para ello, administra un conjunto diverso de flujos de información y bases de datos que provienen tanto de la Facultad como de áreas centrales como Educación Continua, Educación Digital, Cancillería y el Sistema de Información Académica. Esta dependencia concentra un volumen creciente de información, lo que exige mecanismos claros de registro, actualización y análisis.

El equipo de trabajo está conformado por profesionales responsables de distintos componentes de la extensión: gestión de cursos y diplomados, laboratorio de finanzas, proyectos de consultoría, educación digital, iniciativas de apropiación social del conocimiento y relacionamiento con

egresados. Su labor implica integrar datos heterogéneos, mantener actualizados los registros operativos y producir información relevante para procesos de monitoreo, informes estratégicos y toma de decisiones.



CATALINA ZAPATA FLÓREZ

Directora de Desarrollo

catalina.zapataf@urosario.edu.co
Ext. 3002 - oficina 201 - Casa Pedro
Fermín



**MARÍA JOSÉ PINZÓN
DELGADO**

Gestora de Investigación

mariajos.pinzon@urosario.edu.co
Ext. 4157 - oficina 201 - Casa Pedro
Fermín



**ALEJANDRA NIÑO
QUINTERO**

**Coordinadora laboratorio de
finanzas y extensión**

maria.nino@urosario.edu.co
Ext. 4112, Laboratorio de Finanzas – CRAI
Mezzanine



**OLGA LUCÍA PINEDA
ALFONSO**

Asistente administrativo

olga.pineda@urosario.edu.co
Ext. 4109 - oficina 201 - Casa Pedro
Fermín

Figura 1: Equipo de trabajo de la Dirección de Desarrollo y Extensión

1.2. Entorno: Facultad de Economía

La **Facultad de Economía de la Universidad del Rosario** opera bajo una estructura administrativa en la que convergen áreas académicas, administrativas y de proyección externa. Esta estructura está compuesta por la Decanatura y cinco unidades principales: la Dirección de Pregrado, la Dirección de Posgrado, la Dirección Administrativa y Financiera, la Secretaría Académica y la Dirección de Desarrollo. Cada una cumple funciones diferenciadas, pero interdependientes, que influyen directamente en los flujos de información que utiliza y requiere la Dirección de Desarrollo y Extensión.

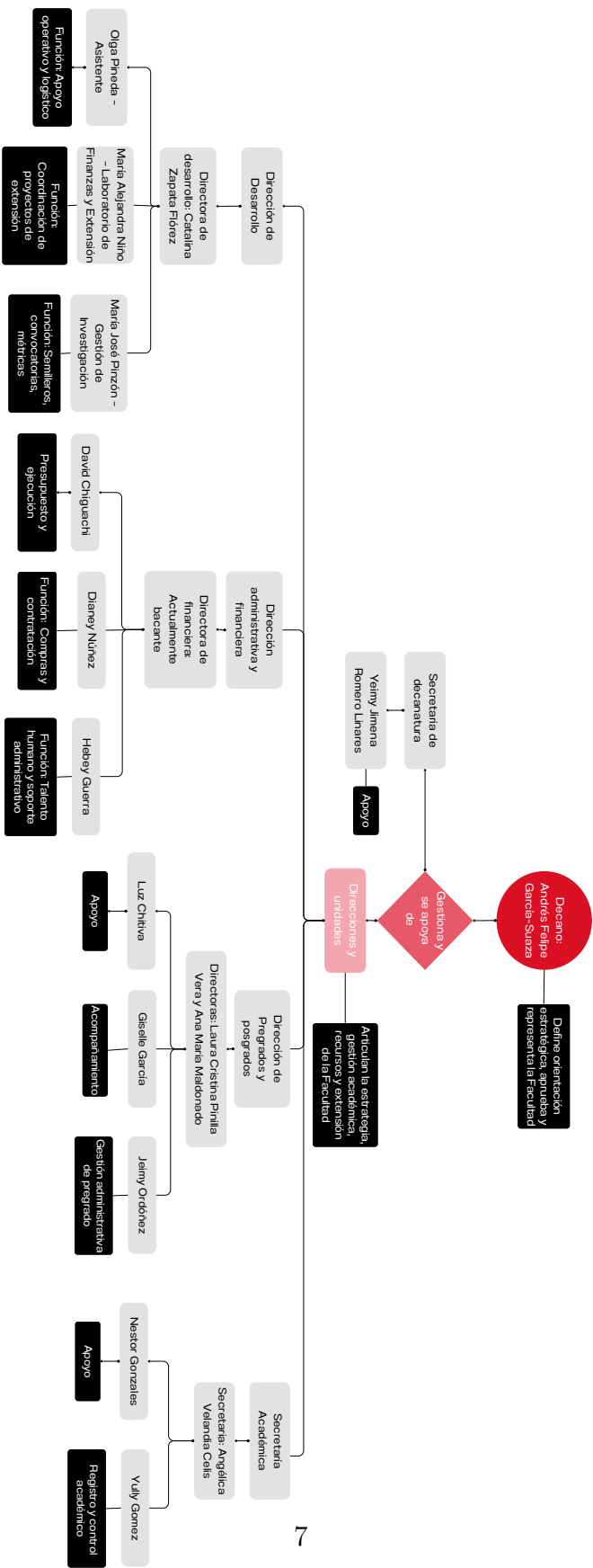
La Decanatura orienta la estrategia institucional y coordina la interacción entre las diferentes direcciones. La Dirección de Pregrado y la Dirección de Posgrado administran la planeación académica, la asignación de cursos y el seguimiento de los programas, generando información relevante para actividades de extensión como cursos libres, formación ejecutiva o eventos académicos. La Secretaría Académica gestiona los registros académicos, los trámites estudiantiles y otros controles institucionales. Por su parte, la Dirección Administrativa y Financiera supervisa procesos contractuales, presupuestales y operativos que inciden directamente en la ejecución de consultorías y en la operación financiera de la extensión.

En este entorno interno, la Dirección de Desarrollo funciona como un nodo de integración que recibe información de todas estas áreas y, al mismo tiempo, produce insumos estratégicos e información para ellas. Según las entrevistas realizadas en el proceso de recolección de datos, la unidad depende de bases enviadas por Educación Continua, reportes de Educación Digital, información de Cancillería para cursos internacionales, datos del Laboratorio de Finanzas, registros derivados de voluntariados y reportes de sistemas centrales como Pure o MinCiencias. A su vez, debe transmitir información consolidada hacia la Decanatura, direcciones académicas, áreas administrativas y procesos institucionales de seguimiento y acreditación.

Este flujo continuo de información configura un entorno operativo complejo, donde la heterogeneidad de fuentes y formatos genera desafíos importantes. La coexistencia de múltiples bases independientes, la ausencia de criterios de estandarización y la prevalencia de procesos manuales de registro y consolidación limitan la trazabilidad de los datos, dificultan la construcción de indicadores comparables y restringen la capacidad del área para analizar patrones o comunicar resultados de manera sistemática. Esta situación constituye la motivación central del cliente: comprender de forma clara la estructura operativa del área, analizar las dinámicas actuales de gestión de información y avanzar hacia una ruta técnica que permita fortalecer el sistema de información que soporta las actividades de extensión.

La Figura 2 presenta la estructura organizacional de la Facultad y permite visualizar las dependencias con las que la Dirección de Desarrollo interactúa de manera continua y de las cuales depende la calidad y disponibilidad de la información utilizada para sus procesos.

Figura 2: Mapa organizacional de la Facultad de Economía



2. Formulación del problema

A partir del entorno descrito, la Dirección de Desarrollo y Extensión se encuentra en un escenario donde la gestión de información es esencial para cumplir sus responsabilidades académicas y misionales. Sin embargo, la fragmentación de datos, la diversidad de fuentes, la ausencia de estandarización y la dependencia de procesos manuales limitan la trazabilidad, la consistencia y el uso estratégico de la información que soporta las actividades de extensión. Esta situación no solo dificulta el análisis interno, sino también la capacidad de la Facultad para comunicar resultados, demostrar impacto y orientar decisiones institucionales con base en evidencia.

En respuesta a este contexto, y conforme a los Términos de Referencia definidos para la consultoría, se estableció la necesidad de formular una *pregunta orientadora* que guiara el proceso de diagnóstico, análisis y propuesta técnica. De esta manera, el proyecto se organizó en torno a la siguiente pregunta problema:

¿Cómo puede la Dirección de Desarrollo y Extensión de la Facultad de Economía diseñar y consolidar un sistema de información que integre, estandarice y permita analizar de manera eficiente los datos generados por sus actividades de extensión, con el fin de mejorar la trazabilidad, la medición de impacto y la toma de decisiones institucionales?

Esta formulación recoge los elementos centrales del problema y define los criterios que orientan la consultoría. Desde una perspectiva técnica, implica construir una arquitectura de datos que articule las distintas fuentes de información, elimine reprocesos y permita la interoperabilidad entre unidades. Desde una perspectiva institucional, busca fortalecer los mecanismos mediante los cuales la Facultad visibiliza los resultados de la extensión, comunica su aporte a la comunidad y consolida su posición en la Universidad. La pregunta problema se convierte así en el eje articulador del proyecto, guiando el levantamiento de información, la revisión de literatura, el análisis de referentes y la construcción de una propuesta alineada con las necesidades de la Dirección de Desarrollo y Extensión.

3. Objetivos del proyecto

3.1. Objetivo general

Diseñar una propuesta técnica para el fortalecimiento del sistema de información de la Dirección de Desarrollo y Extensión de la Facultad de Economía, a partir de un diagnóstico de la situación actual, un análisis comparado de buenas prácticas y la construcción de un modelo conceptual y demostrativo que oriente la integración, estandarización y trazabilidad

de los datos asociados a las actividades de extensión.

3.2. Objetivos específicos

1. Realizar un diagnóstico detallado del sistema de información actualmente utilizado por la Dirección de Desarrollo y Extensión, identificando fuentes de datos, flujos de información, procesos de registro y principales problemas de integración, estandarización y trazabilidad.
2. Desarrollar una revisión de literatura y un análisis comparado de sistemas de información de extensión en instituciones nacionales e internacionales, con el fin de identificar prácticas, modelos y lineamientos que puedan ser aplicados o adaptados al contexto de la Facultad.
3. Analizar los puntos críticos y las oportunidades de mejora del sistema de información actual, definiendo necesidades técnicas, operativas y organizacionales necesarias para avanzar hacia un modelo más integrado y eficiente.
4. Formular recomendaciones y buenas prácticas para el fortalecimiento del sistema de información, orientadas a mejorar la gestión, consolidación, interoperabilidad y uso estratégico de los datos de extensión.
5. Diseñar un modelo conceptual y un esquema demostrativo en SQL que represente la estructura ideal de una base de datos integrada para la Dirección de Desarrollo y Extensión, permitiendo visualizar cómo podrían organizarse, relacionarse y consultarse las entidades clave del sistema.
6. Proponer una ruta de implementación que trace los pasos necesarios para evolucionar desde la situación actual hacia un sistema de información más articulado, estandarizado y alineado con las necesidades institucionales.

4. Alcance del proyecto

El proyecto se desarrolló como una primera aproximación al fortalecimiento del sistema de información de la Dirección de Desarrollo y Extensión. Su alcance se definió considerando el tiempo disponible, las capacidades técnicas del equipo y el marco establecido en los Términos de Referencia. En esta etapa, el objetivo fue comprender el estado actual de la información, identificar oportunidades de mejora y proponer lineamientos para avanzar hacia un sistema más integrado.

4.1. Alcance abordado en este semestre

Esta consultoría se concentró en las acciones que permiten sentar las bases para un proceso más amplio de fortalecimiento del sistema de información. En particular, esta fase incluyó:

- La caracterización del entorno institucional y de los flujos de información que intervienen en las actividades de extensión.
- La identificación de los puntos críticos del sistema de información actual y de los elementos que afectan la trazabilidad y el uso de los datos.
- La elaboración de un modelo conceptual y un diseño demostrativo de cómo podría estructurarse una base de datos integrada para la Dirección de Desarrollo.
- El planteamiento de lineamientos que orientan la evolución del sistema de información hacia un escenario más estandarizado y funcional.

El alcance de esta fase es, por tanto, analítico y propositivo: busca comprender, ordenar y proyectar posibles rutas de acción, sin desarrollar un sistema final ni intervenir operativamente en las plataformas institucionales.

4.2. Aspectos que se podrían abordar en etapas posteriores

Dado que la construcción de un sistema de información institucional requiere un horizonte más amplio, existen componentes que exceden esta fase y que podrían desarrollarse en una implementación futura. Entre ellos se encuentran:

- La integración efectiva entre las fuentes de datos internas y las plataformas centrales de la Universidad.
- La automatización de los procesos de captura, limpieza y consolidación de información provenientes de áreas como Educación Continua, Educación Digital, Cancillería o MinCiencias.
- El desarrollo de una base de datos completa en producción, con protocolos de gobernanza, seguridad, permisos y flujos de actualización definidos.
- La implementación de tableros, indicadores y sistemas de seguimiento que operen sobre una arquitectura de datos ya estandarizada.
- El fortalecimiento de los procesos internos de gestión documental y registro, incluyendo ajustes en prácticas de archivo y nomenclatura.

Estos elementos no forman parte del alcance del proyecto actual, pero representan pasos naturales para avanzar desde el diagnóstico y la propuesta técnica hacia una implementación escalable.

4.3. Delimitación institucional

La consultoría se circunscribe exclusivamente a los procesos, fuentes de información y necesidades de la Dirección de Desarrollo y Extensión de la Facultad de Economía. No se evaluaron sistemas de información de otras facultades, ni se abordaron componentes técnicos de las plataformas administradas por dependencias centrales de la Universidad.

5. Metodología del proyecto

El desarrollo del proyecto se estructuró bajo una metodología consultiva y de ingeniería de datos, guiada por los Términos de Referencia (TDR) entregados por la Dirección de Desarrollo y Extensión y por las orientaciones del profesor a cargo del curso. La metodología se diseñó considerando las restricciones de tiempo propias del espacio académico: una dedicación estimada de **6 horas semanales durante 16 semanas**, lo que requirió priorizar actividades, organizar fases de trabajo y establecer hitos claros para avanzar de manera gradual hacia los productos finales.

5.1. Enfoque general

El proyecto combinó técnicas cualitativas de levantamiento de información, análisis documental, revisión de literatura y modelación conceptual con herramientas propias de la ingeniería de datos. Este enfoque mixto permitió comprender el funcionamiento actual del sistema de información, identificar puntos críticos y desarrollar una propuesta técnica coherente con las necesidades de la unidad. Desde el inicio, el trabajo se organizó alrededor de una ruta metodológica construida a partir de los TDR y discutida periódicamente con el profesor encargado. Esta ruta orientó el alcance, los entregables y las fases del proceso.

5.2. Fases del proyecto

El proyecto se desarrolló en tres fases principales:

1. Diagnóstico inicial del sistema de información.

- Entrevistas semiestructuradas con actores clave de la Dirección de Desarrollo y otras dependencias asociadas al flujo de información.
- Revisión de bases de datos, formatos y documentos institucionales entregados por el cliente.
- Registro y análisis de las sesiones de levantamiento de información con el equipo de extensión.

- Uso de herramientas de IA para transcripción automática (GPTs) y para análisis preliminar del contenido de las entrevistas.

2. Análisis comparado y revisión de literatura.

- Revisión de estudios sobre sistemas de información universitarios, gobernanza de datos y buenas prácticas en extensión académica.
- Identificación de referentes nacionales e internacionales relevantes.
- Construcción de un marco conceptual que permitiera interpretar los retos de la Facultad en relación con experiencias de otras instituciones.

3. Propuesta técnica y modelación de solución.

- Elaboración del modelo conceptual de datos para un sistema integrado de extensión.
- Diseño de un modelo demostrativo en SQL que ejemplifica cómo podrían organizarse las entidades clave.
- Formulación de lineamientos, recomendaciones y una ruta de implementación.

5.3. Técnicas e instrumentos utilizados

Para la ejecución del proyecto se emplearon herramientas y técnicas complementarias que facilitaron el levantamiento, procesamiento y análisis de información:

- **Entrevistas semiestructuradas:** Realizadas con personal de la Dirección de Desarrollo, Secretaría Académica, Pregrado, Posgrado y Dirección Administrativa. Permitieron caracterizar flujos de información, identificar procesos y reconocer problemáticas operativas.
- **Reuniones periódicas con el profesor a cargo:** Funcionaron como espacios de validación de avances, ajuste metodológico, priorización de actividades y orientación técnica para cada etapa del proyecto.
- **Herramientas digitales de apoyo:**
 - **Trello** para la planeación y seguimiento del trabajo mediante tableros Kanban.
 - **Miro** para la elaboración de mapas conceptuales, diagramas y esquemas de estructura.
 - **Loom** para registro de ideas, explicaciones y documentación audiovisual.
 - **GPTs** para transcripción de entrevistas, síntesis preliminares y análisis exploratorio del contenido.
- **Modelación de datos y prototipado SQL:** Se emplearon herramientas de diseño

conceptual y gestores de bases de datos para construir el modelo demostrativo que acompaña la propuesta técnica.

- **Revisión documental:** Análisis de bases, reportes, correos y archivos proporcionados por la Dirección de Desarrollo, así como documentos institucionales relevantes para comprender los procesos operativos.

5.4. Proceso iterativo de trabajo

El proyecto avanzó mediante un proceso iterativo de formulación, retroalimentación y ajuste. Cada fase culminó con la presentación de avances al profesor, la discusión de hallazgos y la reformulación de elementos que requerían mayor precisión. Este ciclo permitió:

- Consolidar una comprensión progresiva del sistema de información.
- Integrar perspectivas del cliente y del docente.
- Refinar el modelo conceptual y la propuesta técnica.
- Mantener coherencia entre diagnóstico, literatura y recomendaciones.

6. Diagnóstico del Estado Actual del Sistema de Información

El levantamiento de información realizado en la Fase 1 permitió caracterizar con mayor precisión el estado actual del sistema de información asociado a las actividades de extensión de la Facultad. Los hallazgos se organizan en torno a los elementos solicitados en los Términos de Referencia: fuentes disponibles, procesos de captura, responsables, flujos, criterios de registro, herramientas utilizadas y condiciones institucionales para la gestión del dato. El diagnóstico refleja un escenario que requiere fortalecimiento, pero que ofrece insumos suficientes para orientar una solución técnica viable y alineada con las capacidades del equipo.

En primer lugar, se identificó que la Facultad no cuenta con un sistema de información consolidado, un repositorio único ni un mecanismo institucional para centralizar los registros de extensión. Esta ausencia no se limita a la inexistencia de una herramienta tecnológica; implica la falta de un modelo conceptual compartido sobre qué se considera un dato de extensión, qué unidades producen información relevante, bajo qué criterios se captura y cómo debería circular internamente. Actualmente, la información se encuentra distribuida en múltiples archivos independientes que cada unidad administra de manera autónoma según sus necesidades operativas inmediatas. Estos archivos se almacenan generalmente en hojas de cálculo sin estructura relacional, sin validaciones sistémicas y sin esquemas consistentes de control de versiones, lo que genera divergencias sobre la información disponible en momentos distintos del tiempo.

Esta situación tiene varias implicaciones. En términos de acceso, no existe un repositorio institucional donde se pueda consultar el estado de la información de manera unificada, lo cual obliga al equipo a solicitar archivos directamente a cada dependencia cada vez que se requiere un reporte, un indicador o una verificación puntual. En términos de contenido, no hay claridad sobre qué datos ingresan al proceso, quién los genera, en qué formato se capturan ni bajo qué reglas se clasifican. Por ejemplo, mientras algunas unidades registran datos de participantes, otras solo consignan información de actividades; algunas incluyen campos sobre responsables, costos o productos asociados, y otras trabajan con estructuras más reducidas. Esta heterogeneidad responde a necesidades operativas legítimas, pero dificulta identificar qué tipo de información constituye el núcleo mínimo institucional que debería estar disponible en cualquier sistema.

Además, la inexistencia de un repositorio consolidado implica que tampoco existe un inventario formal de fuentes. No es posible conocer, de manera sistemática, cuántos archivos relacionados con la extensión se han producido en un periodo, qué variables contienen, qué periodicidad de actualización manejan, ni qué procesos dependen de ellos. Tampoco se cuenta con un registro común que indique la procedencia de los datos, el responsable de su elaboración, el momento en que fueron actualizados por última vez o el propósito para el cual fueron capturados. Este vacío genera incertidumbre sobre la vigencia, completitud y confiabilidad de la información utilizada para procesos académicos, administrativos o estratégicos.

La ausencia de un mecanismo institucional de centralización también impide establecer flujos estandarizados para la captura. Cada unidad determina, según sus dinámicas internas, qué información recopila, cómo la solicita y en qué momento la registra. Esto produce variaciones en la oportunidad y granularidad de los datos, lo que dificulta sincronizar la información en momentos críticos como cierres semestrales, generación de informes o consolidación de actividades para procesos externos. Adicionalmente, al no existir una estructura formal de revisión, las unidades dependen de su experiencia previa para interpretar qué es relevante registrar, qué debe corregirse antes de compartir la información y cómo deben documentarse cambios o ajustes.

Este escenario limita severamente las posibilidades de análisis institucional. Al no contar con un sistema de información consolidado ni con una base relacional que permita vincular actividades con unidades, públicos, docentes, costos o resultados, las consultas transversales requieren reconstruir manualmente piezas dispersas de información. Esto afecta la capacidad de observar tendencias, identificar patrones, comparar unidades o construir indicadores de desempeño con validez institucional. En la práctica, el análisis no depende solo del dato, sino del esfuerzo de consolidación previo, lo cual reduce la posibilidad de generar conocimiento oportuno y dificulta el uso estratégico de la información para toma de decisiones.

En segundo lugar, se observó que la captura de información no sigue estándares comunes y que la Facultad no dispone de un proceso institucional para regular qué datos se reciben,

cómo se estructuran, quién los diligencia y bajo qué criterios se validan antes de su utilización. Actualmente, los datos que llegan a las unidades provienen de múltiples canales que responden a dinámicas operativas muy diferentes. Entre estas fuentes se encuentran formularios elaborados por cada dependencia, listados de inscripción generados por plataformas externas, solicitudes enviadas por correo electrónico, reportes suministrados por aliados o docentes, documentos administrativos asociados a la planeación de actividades y archivos que se consolidan únicamente con fines de reporte interno.

Esta heterogeneidad implica que la Facultad recibe información de naturaleza diversa, con estructuras distintas y con niveles variables de completitud. En la práctica, no existe un mecanismo institucional que determine qué variables deben capturarse para cada tipo de actividad, cuáles son los campos obligatorios, qué definiciones deben adoptarse o cómo deben manejarse los valores faltantes, inconsistentes o duplicados. Cada unidad procesa la información según sus necesidades operativas inmediatas, lo que da lugar a formatos que, aunque cumplen su función interna, no se articulan con los requerimientos institucionales de trazabilidad y análisis transversal.

Una vez recibida, la información se almacena directamente en hojas de cálculo. Estas tablas funcionan como repositorios operativos, pero no incorporan validaciones automáticas, reglas de integridad ni estructuras relacionales que permitan verificar las conexiones entre actividades, participantes, unidades académicas o responsables. Por ejemplo, valores como fechas, modalidades, programas o públicos pueden registrarse en distintos formatos o con nomenclaturas diferentes dependiendo del archivo o del usuario que lo diligencia. Esto afecta la consistencia interna y dificulta la integración posterior.

El uso directo de hojas de cálculo también implica que la información no pasa por procesos ETL, es decir, no se somete a etapas formales de extracción, transformación y limpieza orientadas a garantizar calidad antes del uso institucional. En ausencia de este proceso, los archivos se convierten en contenedores de información en bruto, acumulada tal como es recibida. Las inconsistencias se identifican únicamente cuando un dato genera un problema específico, como un error al consolidar un informe, un valor inesperado en un cálculo o una discrepancia entre unidades. En esos casos, las correcciones se realizan manualmente y de forma puntual, sin que queden registradas metodologías, reglas o procedimientos que permitan replicar el ajuste en futuros ciclos.

Además, la falta de un proceso ETL implica que no existe clasificación por tipo de dato, detección de registros duplicados, validación de dominios, normalización de categorías, estandarización de formatos de fecha, corrección de inconsistencias tipográficas o verificación de relaciones entre campos. Estas tareas recaen en quienes producen los informes y no en un proceso sistemático previo. Esto limita la capacidad de consolidar información de manera oportuna y confiable, ya que buena parte del esfuerzo se invierte en depurar archivos antes de analizarlos.

Otro efecto es que la información circula sin metadatos. No se documenta cuál es la versión vigente de un archivo, cuándo fue actualizado, qué criterios se emplearon para su construcción o qué decisiones se tomaron sobre la depuración. Sin metadatos, es difícil reconstruir el ciclo de vida del dato, evaluar su calidad o determinar qué registros son adecuados para un análisis longitudinal.

En tercer lugar, se constató que las hojas de cálculo utilizadas actualmente contienen estructuras heterogéneas que varían significativamente entre unidades y periodos. Cada dependencia define sus propios formatos de captura en función de sus necesidades administrativas inmediatas, lo que produce diferencias importantes en el tipo de información registrada, el nivel de detalle incluido, las categorías empleadas y la manera en que se nombran o codifican los campos. Esta variabilidad no obedece necesariamente a errores, sino a la ausencia de lineamientos institucionales que orienten qué datos deben recolectarse, cómo deben describirse y bajo qué principios deben organizarse.

La revisión de los archivos mostró que conceptos centrales para la gestión de la extensión, tales como actividad, modalidad, población participante, responsable o unidad académica, aparecen registrados con nomenclaturas distintas y con grados variables de granularidad. Por ejemplo, la categoría de modalidad puede aparecer como taller, capacitación, curso libre, formación continua o actividad académica sin que exista una definición institucional que establezca si estas denominaciones representan tipos diferentes, subcategorías de un mismo concepto o variaciones semánticas equivalentes. Esto mismo ocurre con campos como tipo de público, que en algunos casos se clasifica como estudiante, comunidad externa o participante general, mientras que en otros se registran descripciones más específicas como adulto, joven, docente o beneficiario directo.

En términos técnicos, esta heterogeneidad implica que no existe un diccionario de datos institucional que establezca definiciones formales, dominios permitidos, claves de referencia, relaciones entre entidades o reglas de diligenciamiento. La falta de un diccionario afecta tanto la captura como el procesamiento de la información, ya que impide validar si los valores registrados son consistentes con los dominios esperados, si los campos están completos, si las categorías corresponden a las definiciones institucionales o si las variaciones entre formatos reflejan diferencias reales o interpretaciones distintas del mismo fenómeno.

Además, la ausencia de definiciones comunes limita la interoperabilidad entre unidades. Aunque estas comparten objetivos académicos y administrativos, la falta de semántica transversal hace que cada archivo opere como un universo independiente. Esto dificulta integrar información entre unidades, consolidar reportes transversales, vincular registros con otras fuentes institucionales o construir indicadores comparables. En la práctica, cuando se requiere elaborar un informe que incluya datos de varias unidades, el equipo debe invertir tiempo considerable en homologar categorías, corregir inconsistencias, ajustar nomenclaturas y reconstruir equivalencias entre formatos.

Otro aspecto crítico es que las variaciones en estructura y semántica afectan la trazabilidad histórica. Si las definiciones cambian de un periodo a otro sin un registro documentado, los datos de años diferentes no pueden compararse con seguridad. Esta falta de estabilidad semántica impide construir series de tiempo confiables, monitorear la evolución de las actividades o evaluar el impacto de estrategias específicas aplicadas por la Facultad.

La heterogeneidad también limita la posibilidad de automatizar procesos. En ausencia de un diccionario de datos, cualquier intento de construir consultas dinámicas, flujos de integración o scripts de consolidación debe adaptarse manualmente a cada formato. Esto no solo incrementa el esfuerzo operativo, sino que aumenta la probabilidad de errores cuando se integran fuentes que no comparten una estructura mínima común.

Finalmente, por la experiencia adquirida a lo largo de la carrera se condiera un aspecto central del diagnóstico la ausencia de una arquitectura relacional que permita vincular los registros y sostenga procesos de validación consistentes a lo largo del tiempo. Actualmente, la mayor parte de la información relacionada con las actividades de extensión se encuentra en tablas planas, estructuradas de manera independiente y sin claves primarias o externas que permitan establecer relaciones estables entre actividades, unidades académicas, docentes, participantes, programas o recursos asociados. Esto significa que cada archivo opera como un contenedor autónomo de datos, sin puntos de conexión formales que posibiliten reconstruir trayectorias, identificar vínculos, combinar fuentes o realizar análisis transversales.

La inexistencia de esta estructura relacional influye directamente en la capacidad institucional para generar indicadores de valor. Sin una arquitectura diseñada para relacionar entidades, se dificulta calcular métricas como participación docente acumulada, distribución de actividades por tipo de público, relación entre modalidades y unidades académicas, desempeño por periodos, o la articulación entre extensión, formación y consultoría. Cada análisis implica reconstruir manualmente relaciones entre archivos, lo que no solo incrementa el esfuerzo operativo sino que introduce variabilidad dependiendo de la persona que realiza la consolidación.

Ligado a lo anterior, tampoco existen flujos formales de revisión, validación o consolidación que respalden la calidad del dato. Una vez las unidades registran la información en sus archivos, estos se almacenan en repositorios internos sin pasar por procesos estandarizados que verifiquen completitud, consistencia semántica, integridad referencial o conformidad con criterios institucionales. Las validaciones dependen de la experiencia de los equipos operativos y se realizan únicamente cuando un archivo va a ser utilizado para un informe específico. Como estas revisiones no están documentadas, las correcciones realizadas no quedan registradas, lo cual dificulta garantizar uniformidad entre periodos y reproducibilidad del proceso.

La ausencia de mecanismos de control de versiones y de registro de cambios refuerza esta problemática. No es posible identificar con precisión qué archivo constituye la versión vigente, qué modificaciones se realizaron respecto a versiones anteriores, por qué se efectuaron cambios

en definiciones o categorías o quién autorizó ajustes específicos en los datos. Esta falta de trazabilidad administrativa afecta la capacidad institucional para reconstruir la evolución histórica de la información y para sostener criterios comparables entre ciclos.

Además, la falta de validaciones sistemáticas amplifica los efectos de la ausencia de estructura relacional. Cuando no existen claves que conecten entidades, es más difícil detectar duplicados, identificar inconsistencias entre campos que deberían estar relacionados o validar que los valores registrados correspondan a dominios autorizados. Como consecuencia, la calidad de los datos depende en mayor medida del cuidado individual de quienes los registran y no de una arquitectura institucional que prevenga errores desde el diseño.

7. Revisión de literatura

La revisión de literatura se diseñó para comprender cómo distintas instituciones, marcos teóricos y referentes técnicos han abordado los retos que enfrenta la Dirección de Desarrollo y Extensión en relación con la captura, organización e integración de datos. La literatura consultada no sólo abarca estudios académicos sobre sistemas de información universitarios y gobernanza de datos, sino también experiencias prácticas de universidades que han implementado modelos de registro y seguimiento para actividades de extensión. Esto permitió construir un marco conceptual robusto que fundamenta el diagnóstico y orienta la propuesta técnica del proyecto.

7.1. Sistemas de información universitarios

La extensión universitaria en América Latina ha experimentado, durante las últimas décadas, una transformación marcada por dos tendencias simultáneas: su creciente institucionalización como función sustantiva y la demanda, cada vez más explícita, por medir y demostrar su impacto social e institucional. Lo que antes era un conjunto disperso de actividades de vinculación ahora constituye un eje estratégico para las universidades, tanto por su papel en la democratización del conocimiento como por su peso en procesos de acreditación, toma de decisiones y rendición de cuentas. La literatura reciente coincide en que las universidades se enfrentan a un entorno donde la presión por la calidad, la rendición de cuentas y la articulación con el entorno exige gobernanza de datos sólida, criterios homogéneos y mecanismos sistemáticos para registrar actividades, resultados e impactos.

La literatura sobre evaluación de la extensión insiste en que el impacto no se “descubre” al final, sino que se construye desde el diseño de la información. Rincón Zapata y Pérez Cadavid (2018) muestran que evaluar el impacto de la extensión implica, primero, identificar categorías de transformación social (productiva, cultural, social, ambiental, política) y, a partir

de ellas, definir variables observables que alimentan un modelo econométrico de regresión y un índice de impacto. Para ellos la metodología de impacto descansa sobre una condición previa: contar con un sistema de información capaz de producir series de datos consistentes sobre la extensión. Si las actividades se gestionan con archivos dispersos, criterios cambiantes o registros incompletos, la evaluación de impacto se vuelve frágil.

El vínculo entre datos e impacto se vuelve explícito cuando la literatura introduce el concepto de gobernanza de datos. En su diagnóstico universitario, Garrido Córdoba (2025) muestra que las instituciones suelen ubicarse en niveles intermedios de madurez, con ciertas fortalezas operativas básicas, pero con “debilidades en la integración de sistemas, la calidad de la información y la gobernanza de datos”, a lo que se suman barreras estructurales como “la falta de capacitación del personal y la ausencia de políticas claras” que impiden avanzar hacia un uso estratégico de la información. La literatura especializada es categórica: ningún proceso de evaluación o medición de impacto puede sostenerse sin una infraestructura informacional confiable. Por eso, Garrido Córdoba (2025) enfatiza que “la gestión efectiva de datos se ha convertido en un factor determinante para el éxito institucional en el ámbito educativo contemporáneo”, subrayando que la gobernanza, la integración y la calidad no son atributos complementarios, sino condiciones necesarias para producir evidencia robusta. Esta misma lógica aparece en la literatura administrativa aplicada a la extensión universitaria, donde Bravo Cedeño y Avilés Sotomayor (2020) señalan que la gestión académica debe organizar sus procesos y estructuras “a fin de medir el valor agregado que éstos le generan en su productividad”, lo que exige normas, tecnologías y procedimientos que reduzcan la ambigüedad y aseguren consistencia en la información. En conjunto, estos autores convergen en una afirmación central: sin datos gobernados, estructurados y coherentes, no es posible hablar de medición de impacto de manera sostenible.

Ahora bien, La literatura también muestra cuáles son las prácticas que hoy implementan las universidades para enfrentar estos retos. Una tendencia creciente consiste en adoptar marcos formales de gobernanza como el Data Management Maturity Model, utilizado para diagnosticar brechas de madurez informacional. Este modelo aparece explícitamente aplicado en instituciones latinoamericanas, donde se evalúa la calidad, interoperabilidad y gestión de los datos para orientar planes de mejora. Estas prácticas permiten transitar de registros dispersos hacia un enfoque sistemático, basado en criterios formales de gobernanza, calidad y responsabilidad sobre los datos.

En el plano tecnológico, se observa una migración hacia arquitecturas modernas de datos y pipelines escalables. Sarfo et al. (2025) advierten que las universidades “continúan dependiendo de sistemas heredados, alimentando una creciente necesidad de pipelines escalables de datos”, y proponen una arquitectura en capas con “mecanismos de gobernanza, seguridad y linaje de datos”. Este enfoque plantea que cualquier dato destinado a la analítica debe pasar por procesos de validación, limpieza y transformación antes de integrarse en indicadores estratégicos. La

adopción de estas arquitecturas fortalece la capacidad institucional para producir evidencia confiable y facilita la automatización de indicadores longitudinales.

Un segundo conjunto de avances se encuentra en los sistemas de información diseñados específicamente para la extensión universitaria. El caso de SIGEX, documentado por Barros da Silva et al. (2024), evidencia que las universidades están consolidando plataformas centralizadas que “centralizan la información, organizan la presentación de las acciones y automatizan su gestión”, permitiendo además “el seguimiento, evaluación y organización de las acciones extensionistas”. Estas soluciones reemplazan prácticas manuales y dispersas por flujos digitales estandarizados que mejoran la trazabilidad y la capacidad de análisis. La literatura chilena confirma esta tendencia. Insunza Peña (2020) describe la necesidad de un sistema capaz de “articular procesos, estructuras y tecnologías bajo un mismo estándar” ante la existencia de “registros desconectados y criterios distintos entre unidades”. Esta evidencia demuestra que la integración institucional es un requisito para producir información comparable y útil para la toma de decisiones.

A partir de esta evidencia, la revisión coincide en una conclusión estructural: antes de avanzar hacia modelos sofisticados de impacto, las universidades deben asegurar dos prioridades fundamentales. La primera es organizar y gobernar la información, lo que implica integrar registros dispersos, definir estándares, establecer responsabilidades institucionales y garantizar calidad, consistencia y trazabilidad en los datos que sostienen las acciones de extensión. La segunda es estructurar estos datos de manera que puedan sostener analíticas avanzadas, lo que incluye arquitecturas que validen, transformen y consoliden la información para convertirla en indicadores comparables en el tiempo. Solo cuando estas dos condiciones están aseguradas resulta viable introducir nuevas variables, ampliar dimensiones de análisis y construir indicadores robustos que permitan medir impacto social, institucional y académico de forma creíble. En otras palabras, la literatura traza una hoja de ruta clara: primero gobernar, luego estructurar y finalmente medir. Esta secuencia no es opcional, sino el fundamento técnico sobre el que descansa cualquier sistema moderno de evaluación de la extensión universitaria.

7.2. Experiencias institucionales: aprendizajes comparados

Para comprender cómo otras instituciones han enfrentado problemas similares a los de la Facultad, se analizaron experiencias nacionales e internacionales documentadas por el equipo. Estas experiencias muestran, con distintos niveles de avance, cómo las universidades han tratado de estandarizar plantillas, centralizar información y construir sistemas propios para registrar y medir extensión. El análisis permitió identificar patrones comunes y prácticas emergentes relevantes para el proyecto.

Institución / País	Estrategia aplicada	Resultado o aprendizaje	Referencia
Universidad del Rosario (Colombia)	Establece plantillas comunes para registrar actividades de extensión con campos estandarizados (actividad, modalidad, unidad, público) y responsables de validación por área. Las bases (Educon, Voluntariado, Consultoría) operan como archivos independientes, sin llaves únicas ni conexión entre responsables o registros, lo que impide integrar datos entre hojas o medir trazabilidad entre cursos, ingresos y docentes.	La estructura formal ayuda a reducir ambigüedad y a estandarizar el vocabulario, pero la fragmentación persiste y requiere consolidación manual para elaborar reportes. La falta de arquitectura relacional y herramientas automatizadas limita el potencial analítico y operativo del sistema.	Londoño Trujillo, Á. M. (2020). <i>Los indicadores de gestión en la extensión universitaria: un estudio en Instituciones de Educación Superior de Colombia durante el 2018</i> (Trabajo de grado, Maestría en Dirección). Universidad del Rosario.
Universidad Católica de Colombia	Adopta campos mínimos obligatorios con validaciones automáticas en Excel (año, periodo, modalidad, facultad). Las unidades diligencian las plantillas bajo supervisión del área de Extensión y se alojan en un repositorio común.	La estandarización mejora la coherencia del registro y facilita la consolidación interna. Aún no hay conexión con sistemas académicos ni trazabilidad entre responsables, lo que limita el análisis longitudinal y la interoperabilidad institucional.	Torres Naranjo, M. A., & Arias Hernández, J. S. (2017). <i>Modelo de medición de impacto para los proyectos sociales de la Facultad de Ingeniería en la Universidad Católica de Colombia</i> . Universidad Católica de Colombia.
Universidad de Chile (Chile)	Diseña el sistema SIEVC con jerarquías de categorías (tipo de actividad, unidad, público) y formularios digitales unificados por facultad. El modelo articula extensión, vinculación y comunicaciones, e incorpora procesos de validación distribuidos.	La unificación mejora trazabilidad, permite comparaciones entre unidades y fortalece la medición del impacto institucional. El caso demuestra la importancia de combinar tecnología con gobernanza académica para garantizar uso sostenido del sistema.	Insunza Peña, R. (2020). <i>Diseño de un sistema de información de extensión, vinculación con el medio y comunicaciones en el Campus Sur de la Universidad de Chile</i> (Memoria para optar al título de Ingeniero Civil Industrial). Universidad de Chile.
Universidad de Santander – UDES (Colombia)	Desarrolla un modelo institucional de evaluación y un banco de programas y proyectos, con fichas técnicas, procesos estandarizados, indicadores, teoría de cambio y lineamientos para la Vicerrectoría de Extensión.	Disminuye riesgos de pérdida, mejora trazabilidad documental y permite estructurar indicadores tangibles e intangibles para evaluar gestión, resultados e impacto. Los protocolos fortalecen la gestión del dato y permiten control de versiones y sistematización.	Universidad de Santander. (2019). <i>Modelo de medición de impacto social de la extensión universitaria</i> . Vicerrectoría de Extensión, Universidad de Santander.

Cuadro 1: M1. Organización de datos – síntesis comparada.

M2. Organización de datos

Institución / País	Estrategia aplicada	Resultado o aprendizaje	Referencia
Univasf (Brasil)	Implementó SIGEX , sistema que centraliza el registro, aprobación y seguimiento de actividades de extensión. Digitaliza formularios, automatiza procesos internos y permite generar informes sectoriales por unidad académica.	Redujo el uso de archivos dispersos, agilizó los procesos y mejoró la trazabilidad. El sistema permitió generar indicadores homogéneos y apoyar la toma de decisiones institucionales.	Barros da Silva, E., Cavalcanti Ramos, J. L., & Gomes Fonseca, P. (2024). <i>Desarrollo e implementación del SIGEX: un sistema de gestión para acciones de extensión universitaria</i> . Em Extensão, 23(2), 74–87.
UCI (Cuba)	Desarrolló un sistema de información gerencial que integra cursos, cátedras y proyectos en una única base. Incluye formularios unificados, reportes automáticos y exportación en PDF.	La integración de módulos eliminó duplicidades y fortaleció la coherencia entre procesos. El acceso a información consolidada en tiempo real facilitó la toma de decisiones académicas y administrativas.	Rodríguez, L., García, Y., & Alfonso, R. (2020). <i>Sistema de información gerencial para la gestión de la extensión universitaria</i> . Revista Cubana de Educación Superior, 39(3), 1–15.
Universidad de Chile – Campus Sur	Diseñó el sistema SIEVC mediante entrevistas institucionales y diagnóstico documental. Define cinco ejes: qué se registra, cómo se estructura, qué indicadores aplicar, cómo visualizar los datos y cómo validar institucionalmente.	Detectó fragmentación semántica y desconexión entre facultades. La propuesta busca estandarizar categorías, crear formularios comunes y definir responsables de carga y validación, mejorando la trazabilidad y comparabilidad.	Insunza Peña, R. (2020). <i>Diseño de un sistema de información de extensión, vinculación con el medio y comunicaciones en el campus sur de la Universidad de Chile</i> [Memoria de título]. Universidad de Chile.
Manual de Valencia – OEI/RICYT	Publicó un marco metodológico para caracterizar la vinculación universidad–entorno. Organiza indicadores en dimensiones y establece clasificaciones y fichas estandarizadas.	Facilita la comparación entre instituciones, mejora la rendición de cuentas y alinea los sistemas con estándares internacionales, siendo adoptado en múltiples países.	Observatorio Iberoamericano de la Ciencia, la Tecnología y la Sociedad (OCTS–OEI), & RICYT. (2017). <i>Manual iberoamericano de indicadores de vinculación de la universidad con el entorno socioeconómico: Manual de Valencia</i> .

tableM2. Organización de datos – Experiencias internacionales en sistemas de información para la extensión universitaria.

M3. Organización de datos – Fuentes IA

Autor / Fuente	Estrategia aplicada	Resultado o aprendizaje	Referencia
DAMA International (Global)	El marco DAMA-DMBOK 2.0 estructura la gestión de datos en 11 áreas clave, incluyendo calidad, arquitectura, integración, gobernanza y documentación. Propone lineamientos para construir un repositorio institucional estandarizado con procesos de control y trazabilidad de la información.	Permite pasar de registros desarticulados a un ecosistema de datos confiable y auditable. Su enfoque modular facilita identificar dueños de datos, mantener integridad entre fuentes y automatizar reglas de validación. Es especialmente útil para entornos como el de Extensión, con múltiples actores y fuentes.	DAMA International. (2017). *DAMA-DMBOK: Data Management Body of Knowledge* (2nd ed.). Technics Publications.
Sarfo, P., Owusu, E. & Ampadu, S. (Global)	Proponen una arquitectura de pipelines ETL escalables y modulares aplicados a educación superior. El diseño incluye zonas de staging, validaciones de integridad, integración de fuentes dispares (formularios, sistemas internos, hojas de cálculo) y catálogos de metadatos compartidos.	La solución reduce la intervención manual, mejora la coherencia semántica, detecta errores recurrentes y garantiza trazabilidad. El enfoque permite consolidar múltiples fuentes en una base relacional actualizable para análisis institucional en tiempo real.	Sarfo, P., Owusu, E., & Ampadu, S. (2025). *Architecting scalable data pipelines for learning analytics in higher education: A cloud-native approach*.
Piattini, M., De Miguel, A. & Marcos, E. (España)	Este clásico del diseño de bases de datos promueve el uso de modelos Entidad-Relación y normalización hasta tercera forma normal. Incluye diccionarios de datos y relaciones entre tablas para asegurar integridad referencial y compatibilidad futura.	Es útil para construir un modelo relacional sólido, donde cada dato tiene su estructura, su llave única y su relación con otras entidades. Esto facilita crear tableros conectados, analizar trayectorias (docentes, cursos, egresados) y evitar redundancias.	Piattini, M., De Miguel, A., & Marcos, E. (2000). *Diseño de bases de datos relacionales*. Madrid: RA-MA.

tableM3. Organización de datos – Referentes técnicos propuestos desde consultoría para fortalecer los sistemas institucionales de información.

M4. KPI – Indicadores

Institución / País	Indicador principal	Metodología o enfoque de medición
Universidad del Rosario (Col.)	Número de actividades de extensión por unidad	Conteo administrativo con corte semestral. Cada unidad diligencia un formato con categorías comunes (actividad, público, modalidad). Los datos se consolidan centralmente, aunque no se cruzan con ingresos ni responsables por falta de llaves únicas.
Universidad del Rosario (Col.)	Participantes únicos en cursos/eventos	Depuración manual de registros duplicados usando documento o correo. Se busca medir el alcance real evitando dobles conteos. El dato se reporta por facultad o tipo de evento, sin integración con otros sistemas.
Universidad Católica (Col.)	Ingresos por tipo de actividad	Reporte financiero desagregado por línea de trabajo. Se suman ingresos por centro de costo y se concilian con contabilidad general. Permite monitorear la sostenibilidad económica de las líneas de extensión.
Universidad de Chile (Chile)	Cobertura territorial de las actividades	Clasificación geográfica de las acciones de extensión según comuna o región. Se emplean mapas de calor en tableros institucionales para evaluar el impacto territorial y planificar nuevas ofertas.
Universidad de Santander (Col.)	Satisfacción de asistentes (%)	Encuestas breves post-actividad aplicadas en formato físico o digital. Los resultados (escala Likert) se ponderan por aforo y se promedian para evaluar la calidad percibida de los eventos.

tableM4. Indicadores de extensión – Prácticas de medición utilizadas en universidades con información sistematizada por Catalina.

M5. KPI – Indicadores

Institución / País	Indicador principal	Metodología o enfoque de medición
OEI-RICYT (Iberoamérica)	Ingresos por contratos con actores externos	Suma anual de ingresos según tipo de vínculo (I+D, asesoría, formación). Se estandarizan rubros y se cruzan con información institucional sobre personal, género, áreas y unidades para obtener indicadores comparables.
Univasf (Brasil) – SIGEX	Acciones de extensión aprobadas / ejecutadas	Cálculo automatizado a partir del sistema SIGEX. Se reportan tasas de aprobación y ejecución efectiva de actividades, con posibilidad de análisis por facultad, responsable o línea temática.
UCI (Cuba)	Participación (estudiantes / docentes / comunidad)	Consolidación de registros desde los módulos de cursos, cátedras y eventos. Se emplean paneles de control que permiten desagregar por tipo de actor, nivel de vinculación y frecuencia de participación.
Universidad de Chile – Campus Sur (Chile)	Alineación con ODS (% de actividades)	Cada actividad se clasifica según su contribución a los Objetivos de Desarrollo Sostenible (ODS). Se valida internamente y se consolida un informe porcentual anual con énfasis en impacto temático y poblacional.

tableM5. KPI Consultoría – Indicadores internacionales utilizados como referente en proyectos de fortalecimiento de la extensión.

M6. KPI – Propuesta desde IA (datos de extensión)

Indicador propuesto	Metodología o enfoque de medición
Índice de integridad relacional (% de llaves primarias/foráneas correctamente implementadas)	Evaluación de los esquemas de base de datos para verificar existencia y consistencia de relaciones entre tablas. Se calcula como el porcentaje de registros vinculados correctamente respecto al total esperado (p.ej., cursos sin responsable asociado, ingresos sin actividad origen). Indicador recomendado por Piattini et al. (2000).
Cobertura semántica de variables (% con dominio estandarizado)	Auditoría semántica de variables clave para verificar si cuentan con dominios definidos (catálogos o listas cerradas). Se mide el porcentaje de campos con diccionario actualizado y valores consistentes. Relevante en entornos con múltiples fuentes, como recomienda DAMA-DMBOK (2017).
Tasa de duplicación inter-fuente (% registros con coincidencias parciales)	Comparación entre bases de datos (eventos, cursos, egresados) para detectar registros que se repiten sin identificador común. El análisis se hace con técnicas de fuzzy matching (nombre, fechas, áreas temáticas) y permite estimar redundancia en sistemas no integrados.
Índice de trazabilidad (% registros con historial completo)	Proporción de actividades de extensión que pueden rastrearse desde su origen (formulación) hasta su cierre (evaluación/ingresos/resultados). Un valor bajo indica brechas en el ciclo de vida del dato. Inspirado en estándares de gobernanza de datos de NIST y OCDE.
Nivel de interoperabilidad institucional (escala 1–5)	Escala compuesta basada en criterios de interoperabilidad técnica (API, formatos abiertos), semántica (glosario compartido), organizacional (roles claros) y legal (autorizaciones). Recomendado por la Unión Europea para proyectos de datos abiertos (EIF, 2017).

tableM6. KPI – Indicadores propuestos desde IA para fortalecer la calidad y trazabilidad de los sistemas de datos de extensión universitaria.

A partir del análisis comparativo y los indicadores revisados en los distintos casos institucionales, se identifica un segundo patrón común: los sistemas especializados reducen la carga operativa y habilitan análisis más profundos solo cuando están respaldados por una gobernanza de datos clara, con roles definidos, estándares homogéneos y procesos de validación consistentes. Esta convergencia entre estandarización, automatización y trazabilidad proporciona un marco concreto para orientar el piloto propuesto: permite delimitar qué módulos deben priorizarse, cómo estructurar los flujos de información y qué criterios mínimos de calidad deben garantizarse desde el inicio.

7.3. Referentes técnicos sobre arquitectura, calidad y gobernanza del dato

Para complementar las experiencias institucionales, se revisaron marcos técnicos sobre gestión de datos aplicables a diseños de bases relacionales, integridad semántica y pipelines de integración. Estos referentes permiten traducir los principios conceptuales en lineamientos prácticos para una propuesta técnica. La revisión de referentes técnicos ofrece un marco robusto para comprender cómo deben diseñarse los sistemas de información cuando se busca asegurar integridad, consistencia y capacidad analítica. Estos textos permiten vincular los problemas diagnosticados en la Facultad con principios formales de arquitectura relacional, calidad del dato y gobernanza institucional, mostrando que la solución no depende únicamente de digitalizar formularios, sino de estructurar modelos conceptuales y procesos organizacionales que garanticen trazabilidad y coherencia en el tiempo

Piñeiro Gómez señala que una base de datos relacional solo puede garantizar integridad cuando existe una correspondencia clara entre entidades, atributos y restricciones, y cuando el modelo representa lo que denomina un universo de discurso consistente, es decir, un conjunto de definiciones estables y no ambiguas. Esta afirmación se conecta directamente con los problemas identificados en la Facultad: hoy cada unidad utiliza categorías diferentes para registrar actividades, modalidades, responsables y públicos, lo cual impide reconstruir la trazabilidad del dato e integrar información procedente de fuentes independientes. Tal como advierte Piñeiro, la ausencia de dominios bien definidos conduce a duplicidad, inconsistencias y dificultades para establecer relaciones confiables entre datos dispersos.

En este contexto, la arquitectura relacional no es una preferencia tecnológica, sino una necesidad derivada de la estructura real del problema. El piloto debe partir de entidades institucionales estables (actividades, unidades académicas, modalidades, responsables, públicos, proyectos vinculados) y de reglas explícitas de integridad basadas en claves primarias, claves externas y dominios controlados. De lo contrario, los problemas actuales se reproducirán incluso si se implementa una interfaz moderna.

Asimismo, el diagnóstico evidenció vacíos de roles, estándares y procesos comunes entre unidades. En esta línea, Gwen Thomas define la gobernanza como el ejercicio institucional de toma de decisiones sobre los asuntos relacionados con los datos y enfatiza que la calidad informacional solo se sostiene cuando existen reglas claras, responsables definidos y procesos repetibles en el tiempo. Thomas advierte que la ausencia de gobernanza produce definiciones paralelas, versiones inconsistentes y fricción operativa, lo que coincide con la situación observada en la Facultad, donde cada dependencia decide qué registrar, cómo registrarlo y qué validar.

El marco de Thomas subraya que los programas de gobernanza deben construir procesos estándar y repetibles que reduzcan la fricción operativa y protejan a los usuarios del sistema, y que la gestión de datos debe orientarse por servicios continuos y mecanismos de coordinación

institucional . Desde esta perspectiva, la Facultad no solo necesita un modelo técnico, sino lineamientos institucionales que definan responsabilidades, procesos de validación, reglas de cambio en los catálogos y mecanismos de control de calidad. Esto confirma que el piloto debe incluir componentes de gobernanza que garanticen que el sistema evolucione de manera controlada, evitando la proliferación de definiciones paralelas y la pérdida progresiva de integridad.

Por otro lado, el trabajo de Insunza Peña sobre el diseño del SIEVC en el Campus Sur de la Universidad de Chile resulta particularmente pertinente porque documenta un caso universitario con desafíos similares. Su propuesta se organiza en torno a cinco preguntas clave: qué se registra, cómo se estructura, qué indicadores se aplican, cómo se visualizan los datos y cómo se validan institucionalmente. El autor muestra que los problemas de fragmentación no se deben únicamente a la falta de tecnología, sino a la ausencia de categorías comunes y criterios homogéneos entre unidades, lo cual impedía obtener indicadores comparables y afectaba la capacidad de análisis estratégico. El aporte de este referente es doble. Por un lado, muestra que la estandarización semántica es indispensable para que un sistema pueda generar indicadores útiles y no solo registros aislados. Por otro, evidencia que la validación distribuida, cuando está soportada por reglas claras y definiciones comunes, reduce la carga operativa y fortalece la trazabilidad institucional. Este enfoque coincide plenamente con los hallazgos del diagnóstico, donde la falta de categorías comunes y validaciones sistemáticas ha llevado a dependencias altamente manuales y a la imposibilidad de reconstruir la información en el tiempo

7.4. Síntesis de aportes

La revisión conceptual, institucional y técnica permite articular un marco sólido para comprender qué tipo de arquitectura requiere la Facultad y por qué. Los referentes sobre bases relacionales muestran que las inconsistencias identificadas entre las unidades, tales como la fragmentación semántica, la duplicidad de categorías, la ausencia de identificadores únicos y la falta de trazabilidad, solo pueden resolverse mediante un modelo normalizado que asegure integridad estructural y definiciones comunes. La literatura especializada en gobernanza del dato complementa esta perspectiva al demostrar que la coherencia y la calidad informacional dependen de estándares institucionales, responsabilidades claramente definidas y procesos repetibles que orienten las decisiones sobre el ciclo de vida de la información. Por su parte, los estudios sobre sistemas universitarios, como el caso del SIEVC, evidencian que la estandarización semántica, la validación institucional y las arquitecturas integradas permiten transformar registros dispersos en información comparable, trazable y útil para la gestión académica.

En conjunto, estas perspectivas constituyen la estructura y el sustento teórico suficiente para orientar el diseño de un sistema basado en una base de datos relacional, acompañado de

un diccionario de datos que formalice reglas, dominios y definiciones institucionales y que permita ejercer una gobernanza efectiva del dato. Además, una vez consolidado un sistema robusto y estandarizado de registro, este marco habilita la posibilidad de visualizar de manera sistemática los indicadores clave de desempeño analizados en las matrices, fortaleciendo la capacidad analítica de la Facultad y su toma de decisiones. De este modo, los aportes revisados respaldan el diagnóstico y delinear con precisión los componentes necesarios para avanzar hacia un sistema sostenible, escalable y plenamente alineado con las necesidades institucionales.

8. Visualización del sistema de información: punto de partida del piloto

El primer paso del piloto consiste en visualizar el sistema de información como un conjunto articulado de componentes que interactúan de manera coherente para garantizar que los datos asociados a la extensión sean capturados, transformados, almacenados y utilizados de forma ordenada. Esta visualización inicial permite definir la estructura mínima necesaria del sistema y constituye la base conceptual que guía el diseño técnico y operativo posterior.

Un sistema de información no se limita a una herramienta o plataforma, sino que corresponde a un ecosistema compuesto por fuentes, procesos, arquitectura, reglas y mecanismos de consulta. Por ello, la construcción del piloto inicia con la identificación explícita de estos elementos y la definición preliminar de sus funciones dentro del flujo institucional de datos. La meta es avanzar desde un conjunto de archivos dispersos hacia un sistema integrado, confiable y capaz de soportar análisis estratégicos.

8.1. Componentes esenciales del sistema

La visualización del sistema establece que el piloto debe estructurarse alrededor de cinco componentes fundamentales:

1. **Fuentes de información.** Identificación de todas las fuentes primarias y secundarias que producen información de extensión. Incluye archivos actuales, formularios, reportes, bases complementarias y registros ya existentes en otras áreas de la institución.
2. **Proceso de extracción, transformación y carga.** Definición de un flujo organizado que permita recibir la información, depurarla, estandarizarla y prepararla antes de su incorporación al repositorio institucional.
3. **Almacenamiento estructurado.** Construcción de una base de datos relacional que funcione como núcleo del sistema. Este componente debe garantizar integridad, tra-

zabilidad y consistencia entre actividades, unidades, responsables, públicos y demás entidades relevantes.

4. **Mecanismos de consulta.** Diseño de consultas básicas que permitan recuperar información de manera sistemática. Estas consultas responden a preguntas prioritarias para la Facultad y aseguran que los datos puedan utilizarse sin necesidad de manipulación manual.
5. **Visualización analítica.** Definición de cómo se mostrarán los indicadores clave de desempeño mediante herramientas de análisis. Este componente permite traducir los datos estructurados en información lista para la toma de decisiones.

8.2. Lógica de funcionamiento del sistema

La visualización del sistema implica definir la función que desempeña cada componente en el flujo de información. Las fuentes deben clasificarse según su origen y relevancia. El proceso de extracción, transformación y carga debe ser claro y documentado. La base de datos debe cumplir principios relacionales para evitar duplicidad o inconsistencias. Las consultas deben alinearse con necesidades reales de gestión académica. Finalmente, la visualización debe entregar indicadores confiables que apoyen la planificación y el seguimiento institucional. Esta lógica permite anticipar la trazabilidad que se espera del sistema. La meta es construir un entorno integrado donde los datos fluyan de manera coherente, un repositorio confiable sostenido por llaves e integridad referencial y un sistema estratégico capaz de ofrecer información útil y oportuna para la toma de decisiones.

8.3. Fuentes de información

La identificación precisa de las fuentes de información constituye el punto de partida del piloto, ya que permite comprender qué datos ingresan al sistema, de dónde provienen y bajo qué condiciones se capturan. Este componente es fundamental para asegurar que la arquitectura posterior responda a la realidad operativa de la Facultad y pueda integrar eficientemente los insumos que hoy alimentan la gestión de extensión.

El diagnóstico realizado muestra que actualmente existen dos tipos de fuentes que deben ser incorporadas en el diseño del sistema:

1. **Tablas externas recibidas de terceros.** Corresponden a archivos que llegan a la Facultad desde unidades académicas, aliados, plataformas de inscripción o procesos administrativos externos. Estas tablas ya vienen estructuradas según las necesidades del actor que las produce y, por tanto, contienen campos, categorías y formatos definidos externamente. Estas fuentes operan como información primaria sobre actividades,

participantes o solicitudes, pero presentan variabilidad en su estructura y nivel de detalle.

2. **Tablas internas creadas y diligenciadas por el equipo de extensión.** Se trata de archivos construidos directamente por la Facultad para gestionar actividades, registrar información operativa, sistematizar comunicaciones con docentes, controlar procesos o consolidar insumos para reportes. Estas tablas son completadas manualmente por el equipo, lo que permite ajustar sus contenidos a las necesidades internas, aunque también genera variabilidad en nomenclaturas, categorías y criterios de diligenciamiento.

Ambos tipos de fuentes cumplen funciones esenciales, pero poseen naturalezas distintas. Las tablas externas constituyen información primaria cuya estructura no controla la Facultad, mientras que las tablas internas operan como mecanismos de registro y organización complementaria. Esta distinción es clave para el diseño del sistema, ya que implica que el proceso de extracción, transformación y carga deberá adaptarse a lógicas diferentes: en un caso, estandarizar formatos heterogéneos recibidos de múltiples actores; en el otro, formalizar definiciones y reglas para minimizar variaciones internas.

Reconocer y clasificar adecuadamente estas fuentes permite definir qué elementos deben priorizarse en la integración, qué campos requieren normalización y qué procesos deben documentarse para garantizar que la información fluya de manera coherente hacia el repositorio institucional que se construirá en etapas posteriores del piloto.

8.4. Proceso de extracción, transformación y carga

El proceso de extracción, transformación y carga constituye la columna vertebral del piloto, dado que permite organizar la manera en que las fuentes de información ingresan al sistema, se depuran y se preparan para alimentar la base de datos relacional. Su diseño debe considerar de manera explícita los dos tipos de fuentes identificadas: las tablas externas que llegan ya estructuradas por actores terceros y las tablas internas creadas y diligenciadas por el equipo de extensión. Un proceso ETL adecuado permite tratar ambas fuentes como insumos legítimos que, una vez transformados, se integran en un repositorio unificado.

El componente de extracción corresponde a la captura sistemática de estas fuentes. En el caso de tablas externas, la extracción implica recibir los archivos tal como son entregados por unidades académicas o aliados. En el caso de tablas internas, consiste en recuperar los archivos de trabajo que el equipo de extensión genera durante su operación regular. En esta etapa no se modifica el contenido, sino que se asegura que las fuentes ingresen al flujo de procesamiento bajo un esquema controlado y documentado.

La transformación es el núcleo técnico del proceso. En esta fase se estandarizan formatos, se normalizan categorías, se alinean definiciones, se depuran inconsistencias y se preparan los

datos para integrarlos en la base relacional. Este proceso incluye tareas como homogenizar fechas, corregir variaciones tipográficas, identificar duplicados, validar dominios, reestructurar columnas según la entidad correspondiente o separar información que llega combinada en un mismo campo. Esta etapa también permite incorporar reglas institucionales derivadas del diccionario de datos que se construirá en fases posteriores.

La carga es el cierre del proceso y consiste en ingresar los datos transformados en la base de datos institucional. Para el piloto, esta carga puede realizarse mediante scripts básicos que aseguren que cada entidad y cada relación se almacenen correctamente. En ciclos posteriores, este proceso podrá automatizarse para que la actualización de la base se realice con intervenciones mínimas y bajo criterios estandarizados.

Desde la perspectiva metodológica, los procesos ETL suelen estructurarse en dos modalidades: los flujos por lotes, donde las fuentes se procesan periódicamente en bloques, y los flujos incrementales, donde únicamente se incorporan los cambios detectados desde la última carga. Para el piloto, el enfoque por lotes es el más adecuado, dado que permite estabilizar la estructura y validar la calidad antes de avanzar hacia automatizaciones incrementales.

El proceso ETL también puede beneficiarse del uso de herramientas de inteligencia artificial. En la fase de transformación, la IA puede emplearse para identificar valores anómalos, sugerir estandarización de categorías, detectar duplicidades no exactas, proponer reglas de clasificación y generar documentación automática sobre el procesamiento realizado. En etapas avanzadas, puede utilizarse para crear agentes que supervisen el flujo de datos, identifiquen inconsistencias recurrentes y alerten al equipo sobre posibles mejoras en las reglas o los formatos.

Finalmente, un proceso ETL bien diseñado permite asumir las tablas que hoy utiliza la Facultad como insumos válidos que alimentan la base de datos. En lugar de reemplazar estas fuentes, el sistema las reconoce, las estructura y las integra para construir un repositorio confiable y preparado para análisis. De esta manera, el proceso ETL se convierte en un mecanismo que articula la operación actual con la arquitectura futura del sistema de información institucional.

Por razones de tiempo y alcance, el piloto no incluye la construcción completa del proceso ETL ni su automatización. En esta fase, los datos serán simulados directamente dentro de la base de datos para demostrar la estructura, las relaciones y el funcionamiento esperado del sistema. Sin embargo, el diseño propuesto está pensado para que la Facultad pueda, en etapas posteriores, desarrollar y automatizar este flujo mediante herramientas como scripts en Python o R, paquetes especializados en ingeniería de datos, conectores institucionales, cargas programadas desde servidores o incluso asistentes de inteligencia artificial que validen, limpien y estandaricen la información antes de ingresarla al repositorio. El ideal es que este proceso evolucione hacia un mecanismo continuo y conectado que permita que las fuentes actuales alimenten la base de datos de manera sistemática, estandarizada y con mínima intervención manual.

8.5. Almacenamiento estructurado y modelamiento de la base de datos

El modelamiento de la base de datos constituye el núcleo del piloto y la pieza central del sistema de información. En esta etapa se define la estructura fundamental que permitirá integrar las fuentes de datos, organizar la información de manera lógica y garantizar que pueda utilizarse para análisis transversales y toma de decisiones institucionales. A diferencia de otros componentes del piloto, este no es un paso adicional, sino el elemento que articula la visualización del sistema, el proceso ETL, el diccionario institucional de datos, las reglas de validación y los mecanismos de consulta.

El diseño del modelo requiere tomar decisiones sobre la definición de entidades, sus atributos, las claves primarias y externas, las relaciones entre entidades y los dominios que regirán cada uno de los campos. Las entidades representan unidades conceptuales estables del proceso de extensión, tales como actividad, unidad académica, modalidad, responsable, tipo de público o periodo. Determinar qué entidades deben existir implica analizar la naturaleza de la información, su función en los procesos institucionales y la necesidad de mantener definiciones y registros consistentes en el tiempo.

Cada entidad debe contar con una clave primaria que garantice unicidad. A su vez, el modelo debe establecer claves externas que conecten las entidades entre sí y permitan reconstruir las relaciones entre actividades, responsables, públicos y unidades académicas. Este proceso define la estructura relacional del sistema y permite evitar duplicidades, identificar vínculos y asegurar integridad en la organización de los datos. Las relaciones pueden ser de uno a uno, uno a muchos o muchos a muchos, y su correcta identificación es esencial para construir un sistema flexible y confiable.

La elección del enfoque relacional responde directamente a la situación actual de la Facultad, donde los datos se encuentran en archivos planos y aislados, sin relaciones formales ni claves que permitan integrarlos. Un modelo relacional es indispensable cuando se requiere organizar entidades claramente definidas, aplicar reglas de validación consistentes, evitar duplicidad, mantener trazabilidad y producir indicadores comparables entre periodos. Los modelos relacionales permiten aplicar principios de normalización que eliminan redundancias, consolidan definiciones y aseguran que cada registro dependa únicamente de la entidad a la que pertenece.

En este contexto, el modelamiento debe incluir una entidad central para las actividades y entidades complementarias para atributos estructurales como modalidades, unidades académicas, tipos de público, responsables y periodos. Se requieren también tablas intermedias para capturar relaciones complejas, especialmente aquellas de muchos a muchos, como las que se presentan entre actividades y responsables o entre actividades y ciertos tipos de públicos. La arquitectura inicial propuesta permite simular la integración de las fuentes actuales y demostrar cómo la información puede organizarse en un repositorio estructurado.

Si bien el enfoque relacional es el más adecuado para el contexto de la Facultad, también es importante comprender qué implicaría optar por un modelo no relacional. Los modelos no relacionales se basan en documentos, grafos o pares clave valor y ofrecen flexibilidad y escalabilidad en ciertos escenarios. Sin embargo, presentan limitaciones significativas en contextos donde la integridad referencial, la normalización y el control semántico son necesarios. En un entorno universitario con categorías definidas, estructuras organizacionales claras y necesidad de análisis comparativos, un modelo no relacional requeriría esfuerzos adicionales de gobernanza y no resolvería las necesidades centrales del proceso.

Por estas razones, el modelamiento relacional se convierte en el piloto en sí mismo. Es la estructura que permitirá ejecutar consultas, visualizar indicadores, incorporar reglas del diccionario institucional de datos y consolidar fuentes internas y externas. A través de este modelo, se establece la arquitectura base del sistema y se define la ruta para evolucionar desde un ecosistema fragmentado hacia un sistema institucional sólido, ampliable y gobernable. En fases posteriores podrán incorporarse validaciones automáticas, índices, vistas especializadas, auditorías y procesos de actualización continua que completen la transición hacia un sistema maduro y plenamente operativo.

8.6. Diccionario institucional de datos

El diccionario institucional de datos constituye el componente que permite formalizar las definiciones, dominios, estructuras y reglas que organizan la información dentro del sistema. Su función es otorgar coherencia semántica al modelo relacional, asegurar que todas las unidades utilicen conceptos homogéneos y proporcionar los lineamientos necesarios para la captura, validación y uso de los datos. En un contexto donde actualmente cada dependencia maneja formatos, categorías y nomenclaturas distintas, este diccionario se convierte en el mecanismo que unifica el lenguaje institucional y previene la proliferación de interpretaciones o estructuras divergentes.

El diccionario establece definiciones precisas para cada entidad, atributo y relación del modelo. Esto implica documentar qué se entiende por actividad, modalidad, unidad académica, tipo de público, responsable o periodo, así como identificar qué variables son obligatorias, cuáles son opcionales y bajo qué condiciones deben registrarse. Estas definiciones no solo describen el significado de cada campo, sino también su propósito dentro del sistema, su tipo de dato, su formato esperado y las reglas necesarias para garantizar integridad semántica.

Un componente central del diccionario corresponde a los dominios controlados. Cada atributo con valores categóricos debe estar asociado a un conjunto de valores válidos definidos institucionalmente. Esto incluye modalidades autorizadas, categorías de público, tipos de actividad, roles de los responsables o clasificaciones de programas. Establecer dominios comunes permite evitar variaciones tipográficas, duplicidades semánticas, diferencias entre unidades y problemas

de integración al momento de cargar la información en la base de datos.

El diccionario también documenta las reglas de diligenciamiento que deben aplicarse al momento de capturar la información. Estas reglas incluyen instrucciones básicas como formato de fecha, uso de identificadores institucionales, manejo de campos vacíos, estructura de nombres propios o codificación recomendada. Asimismo, define criterios de validación para garantizar que los datos capturados correspondan a dominios permitidos, cumplan con su tipo de dato y mantengan consistencia con otras entidades del sistema.

Otro elemento clave del diccionario es la incorporación de metadatos institucionales. Estos metadatos permiten registrar información sobre la fuente del dato, el responsable de su captura, la fecha de su actualización y las versiones del registro. La documentación de metadatos facilita la trazabilidad y permite reconstruir la evolución de los datos a lo largo del tiempo, lo cual es fundamental para análisis longitudinales y procesos de auditoría interna.

8.7. Consultas y mecanismos de recuperación de información

La capacidad de recuperar información de forma sistemática y confiable constituye uno de los objetivos centrales del sistema de información. Los mecanismos de consulta permiten transformar la base de datos relacional en una herramienta útil para la gestión académica y administrativa, facilitando la elaboración de reportes, el seguimiento de actividades y la toma de decisiones estratégicas. Esta etapa del piloto se concentra en definir las consultas mínimas necesarias y establecer criterios que aseguren consistencia entre periodos e independencia respecto a los formatos originales de las fuentes.

El diseño de las consultas parte de la estructura relacional definida en el modelamiento. Cada entidad representa un conjunto coherente de información que puede ser recuperada mediante consultas directas, mientras que las relaciones permiten vincular actividades con docentes, unidades académicas, modalidades y públicos. En este sentido, las consultas deben construirse a partir de combinaciones entre entidades que reflejen preguntas institucionales frecuentes, tales como número de actividades por unidad académica, distribución de modalidades en un periodo específico, volumen de participantes por tipo de público, participación docente o variaciones en la oferta de extensión entre semestres.

Un principio fundamental para la construcción de consultas es la independencia respecto a los formatos originales. El sistema no consulta directamente los archivos externos ni las tablas internas; en su lugar, opera sobre datos ya estandarizados y cargados en la base relacional. Esto asegura que los resultados no dependan de la estructura inicial de las fuentes y que puedan ser reproducidos de manera consistente a lo largo de distintos ciclos académicos.

Las consultas pueden clasificarse en dos grupos. Las consultas operativas permiten recuperar información puntual para apoyar procesos administrativos, tales como identificar actividades

activas en un periodo, verificar responsables asociados a una unidad o revisar listados de modalidades vigentes. Por su parte, las consultas analíticas integran varias entidades y permiten producir indicadores transversales que describen tendencias y comportamientos institucionales. En ambos casos, la base relacional proporciona la estructura necesaria para evitar duplicidades, inconsistencias o interpretaciones divergentes del dato.

Finalmente, el diseño de esta etapa debe considerar criterios de escalabilidad. Si bien el piloto define un conjunto de consultas básicas, el sistema debe permitir la incorporación progresiva de consultas más complejas, vistas especializadas y mecanismos avanzados como procedimientos almacenados, filtros dinámicos o consultas para análisis longitudinal. Estas capacidades permitirán ampliar las posibilidades del sistema y generar información cada vez más útil para la toma de decisiones.

8.8. Visualización analítica e indicadores iniciales

La visualización analítica constituye la fase en la que los datos estructurados en la base de datos se transforman en información accesible y comprensible para la gestión institucional. Su función es traducir consultas técnicas en representaciones que permitan identificar tendencias, patrones y comportamientos relevantes dentro de la extensión universitaria. Para el piloto, la visualización se enfoca en un conjunto de indicadores iniciales que pueden ser construidos a partir del modelo relacional y que reflejan preguntas prioritarias para la Facultad.

La visualización depende directamente de la calidad y consistencia de la base relacional. Al operar sobre datos ya depurados y estandarizados, los indicadores pueden expresarse de manera confiable y reproducible. Entre los indicadores que pueden desarrollarse en esta etapa se incluyen volumen de actividades por unidad académica, distribución de modalidades, participación por tipo de público, número de responsables asociados a la oferta, tendencias intersemestrales y proporciones de actividades según su naturaleza. Estos indicadores permiten describir la dinámica de la extensión y constituyen la base para análisis más complejos en fases posteriores.

Las visualizaciones deben diseñarse siguiendo criterios de claridad y utilidad. En lugar de replicar las estructuras de origen de los archivos, se enfocan en variables institucionales definidas en el diccionario de datos y alimentadas por el modelo relacional. Esto permite que los reportes no dependan de formatos específicos y que puedan mantenerse estables a lo largo del tiempo. La visualización puede apoyarse en herramientas como hojas de cálculo avanzadas, plataformas de inteligencia de negocios y lenguajes de programación especializados en análisis de datos.

En etapas más avanzadas, y una vez consolidada la estructura del sistema, es posible incorporar visualizaciones dinámicas y tableros interactivos que permitan a la Facultad explorar

información en tiempo real y generar análisis personalizados. Estas herramientas pueden integrarse con el proceso de carga de datos para construir entornos de monitoreo continuo, donde los indicadores se actualicen automáticamente con cada ciclo del sistema. El piloto, sin embargo, se concentra en demostrar la viabilidad de estos procesos y en generar un conjunto inicial de representaciones que facilite la toma de decisiones y la comprensión del comportamiento de la extensión.

Además, la visualización analítica permite incorporar los indicadores clave de desempeño identificados en la revisión de literatura institucional y técnica. Estos indicadores no solo describen la actividad operativa, sino que también reflejan prácticas comunes en sistemas consolidados como SIGEX y SIEVC, donde la estandarización y la trazabilidad habilitan métricas comparables entre periodos y unidades. En este piloto, dichos KPI pueden incluir medidas de volumen, cobertura, modalidades, participación, temporalidad y relación entre actividades y actores, lo que permite alinear la analítica del sistema con los referentes técnicos revisados y con los criterios de calidad y gobernanza que caracterizan a los modelos de referencia. Estos indicadores servirán como guía inicial para construir un tablero institucional que, en etapas posteriores, podrá ampliarse conforme evolucione la estructura del sistema y se integren nuevas fuentes de información.

9. Ejecución del piloto del sistema de información

9.1. Selección de tablas para la demostración del modelo

Para ejecutar el piloto se definió un conjunto acotado de tablas operativas que permitiera demostrar, de manera controlada y realista, el funcionamiento del modelo relacional propuesto. La selección se basó en dos criterios principales. En primer lugar, en las restricciones de tiempo establecidas para el piloto, que hacen necesario trabajar con un volumen manejable de información. En segundo lugar, en la necesidad de contar con tablas que compartieran variables susceptibles de relacionamiento, de modo que fuera posible mostrar la lógica del modelo relacional sin alterar la estructura original de los datos.

Bajo estos criterios, el piloto utiliza como insumo las tablas actualmente empleadas para gestionar las actividades de Moocs, Educación Continua, Consultorías, Eventos y Voluntariados. Estas tablas resultan especialmente útiles porque, aunque fueron diseñadas para fines operativos distintos, comparten elementos estructurales que permiten vincularlas dentro de la arquitectura relacional. Esto facilita demostrar cómo entidades como actividad, periodo, responsable, unidad académica o modalidad pueden integrarse en una base de datos centralizada.

En esta etapa del piloto, las tablas se emplean exclusivamente como ejemplos para ilustrar

cómo la información existente puede transformarse e incorporarse en el modelo relacional. No se modifica su contenido original ni se pretende replicar el sistema completo, sino mostrar cómo el modelo permite organizar la información de manera coherente, identificar relaciones entre registros y establecer una base estructurada a partir de fuentes que hoy se encuentran dispersas. La selección realizada permite así evidenciar el potencial del sistema y sentar las bases para la integración de un universo más amplio de datos en fases posteriores.

9.2. Requisitos funcionales del sistema

Una vez delimitado el alcance del piloto y visualizada la estructura general del sistema de información, el siguiente paso consiste en definir los requisitos funcionales. Estos requisitos describen lo que el sistema debe ser capaz de realizar desde la perspectiva del usuario y de los procesos institucionales. Su función es traducir la visión conceptual en comportamientos concretos que permitan orientar el diseño técnico, el modelamiento de la base de datos y los flujos de integración.

Los requisitos funcionales se clasifican en dos niveles:

Requisitos de alto nivel. Corresponden a capacidades generales del sistema y expresan qué se espera que el sistema logre sin detallar los procedimientos internos. Incluyen acciones amplias como centralizar la información, asegurar trazabilidad, permitir la integración de fuentes múltiples o generar información para la toma de decisiones. Este nivel es útil para comunicar la visión del sistema y garantizar alineación institucional.

Requisitos de bajo nivel. Describen acciones específicas que el sistema debe ejecutar, tales como validar campos obligatorios, normalizar categorías, aplicar reglas de integridad, ejecutar consultas estructuradas, realizar cargas periódicas o generar indicadores específicos. Estos requisitos son necesarios para diseñar el modelo relacional, construir el diccionario de datos y preparar los flujos operativos del sistema.

Ambos niveles son complementarios. Los requisitos de alto nivel definen la dirección del sistema, mientras que los de bajo nivel permiten implementarlo de manera precisa y controlada. En el contexto del piloto, esta definición inicial de requisitos constituye la base para estructurar el funcionamiento del sistema y preparar la construcción de los módulos de consulta y visualización.

9.3. Requisitos Funcionales de ALTO nivel

Código	RQF-AL01
Nombre	Gestión integral de actividades de extensión
Fecha	Octubre de 2025
Grado de necesidad	Alto
Usuario principal	Coordinación de Extensión, Dirección de Desarrollo y personal administrativo autorizado.
Descripción	El sistema debe permitir registrar, consultar y actualizar de manera centralizada la información asociada a las actividades de extensión gestionadas por la Facultad. La funcionalidad incluye la captura estructurada de datos esenciales como nombre de la actividad, tipo, modalidad, fechas relevantes, responsable, unidad académica y otros atributos definidos en el diccionario institucional de datos. El sistema validará la completitud de los campos, aplicará reglas de consistencia y garantizará que cada registro se almacene como una entidad única dentro de la base de datos relacional, preservando trazabilidad e integridad.
Entradas	Datos básicos de la actividad: nombre, tipo, modalidad, responsable, fechas, unidad académica, estado y demás atributos definidos para el módulo de actividades.
Fuente	Formularios internos del sistema e información enviada mediante las plantillas operativas utilizadas por unidades como Educación Continua, Consultorías, Eventos, Voluntariados o Moocs.
Salida	Registro estandarizado, validado y almacenado en la base relacional para su uso en consultas, indicadores y procesos analíticos.
Destino	Base de datos relacional del Sistema de Información de Extensión.
Regla de negocio	Solo usuarios con rol de Coordinador o Gestor podrán crear o modificar registros. La Dirección de Desarrollo validará actualizaciones de alto impacto o cambios estructurales.
Proceso general	Este requisito define la funcionalidad base del módulo de actividades, articulando la captura de información con el modelo relacional del sistema.

Cuadro 2: RQF-AL01 – Gestión integral de actividades de extensión

Código	RQF-AL02
Nombre	Gestión de actores y egresados
Fecha	Octubre de 2025
Grado de necesidad	Alto
Usuario principal	Coordinación de Extensión, Gestores Académicos y Dirección de Desarrollo.
Descripción	El sistema permitirá registrar, actualizar y consultar la información de los actores vinculados a las actividades de extensión, incluyendo docentes, estudiantes, egresados, aliados externos y demás participantes relevantes. La funcionalidad debe centralizar los datos de identificación, roles, unidades académicas y vínculos institucionales, garantizando que cada actor cuente con un registro único dentro del modelo relacional. El sistema permitirá asociar actores a una o varias actividades, registrar su nivel de participación y mantener sincronía con bases institucionales existentes cuando corresponda.
Entradas	Datos de identificación, contacto, rol, tipo de vínculo institucional y actividades en las que participa cada actor.
Fuente	Formularios internos de registro, bases institucionales como UR-Alumni y planillas suministradas por unidades académicas o validaciones de la Dirección de Desarrollo.
Salida	Registro único y estandarizado de cada actor, con información lista para análisis, vinculación y trazabilidad.
Destino	Tablas relacionales del sistema: Personas, Roles, Participaciones y Entidades .
Regla de negocio	Cada actor debe tener un identificador único. No se permiten registros duplicados. La edición o aprobación de cambios estará restringida a usuarios autorizados por la Dirección de Desarrollo.
Proceso general	El módulo organiza y mantiene la red institucional de actores vinculados a la extensión, permitiendo su asociación con las actividades registradas y asegurando trazabilidad, consistencia y control de la información personal e institucional.

Cuadro 3: RQF-AL02 – Gestión de actores y egresados

Código	RQF-AL03
Nombre	Generación estructurada de variables e indicadores derivados
Fecha	Octubre de 2025
Grado de necesidad	Alto
Usuario principal	Analistas institucionales y Dirección de Desarrollo.
Descripción	El sistema debe permitir la creación y gestión de variables derivadas e indicadores contruidos a partir de la información integrada en la base de datos relacional. La funcionalidad incluye definir fórmulas, parámetros y criterios de cálculo, así como registrar metadatos tales como unidad de medida, descripción, periodicidad y fecha de actualización. El módulo transformará datos operativos en información analítica estandarizada, apta para seguimiento, visualización y evaluación de la dinámica de extensión.
Entradas	Datos consolidados provenientes de los módulos de actividades, actores, unidades académicas y demás entidades definidas en el modelo relacional.
Fuente	Modelo de datos del sistema y parámetros analíticos definidos y aprobados institucionalmente.
Salida	Indicadores calculados, variables derivadas y estructuras analíticas listas para ser utilizadas en consultas y tableros de visualización.
Destino	Vistas analíticas y tablas intermedias del sistema destinadas al consumo por herramientas de visualización.
Regla de negocio	Los indicadores deben generarse exclusivamente a partir de fuentes oficiales del sistema, contar con metadatos completos y conservar trazabilidad respecto al dato original.
Proceso general	El módulo analítico convierte la información depurada y consolidada en resultados cuantitativos y cualitativos que facilitan el seguimiento institucional y la toma de decisiones estratégicas.

Cuadro 4: RQF-AL03 – Generación estructurada de variables e indicadores derivados

Código	RQF-AL04
Nombre	Visualización estratégica y tablero de control Power BI
Fecha	Octubre de 2025
Grado de necesidad	Alto
Usuario principal	Dirección de Desarrollo, Decanatura y autoridades institucionales.
Descripción	El sistema debe permitir la construcción de un tablero de control en Power BI que represente, de forma clara y accesible, los indicadores generados a partir de la base de datos relacional. Este tablero permitirá explorar la información por unidad académica, tipo de actividad, periodo o línea temática, aplicar filtros dinámicos y exportar reportes. El módulo proporciona una vista ejecutiva que facilita el monitoreo estratégico de la extensión y respalda la toma de decisiones basada en datos institucionales.
Entradas	Vistas analíticas e indicadores derivados producidos por el módulo de análisis del sistema.
Fuente	Base de datos relacional consolidada y estructuras analíticas validadas.
Salida	Tablero interactivo en Power BI con indicadores actualizados del sistema de extensión.
Destino	Usuarios estratégicos de la Facultad y autoridades institucionales.
Regla de negocio	El acceso a las visualizaciones estratégicas será restringido según el perfil del usuario. Solo perfiles autorizados podrán exportar reportes o ver información sensible.
Proceso general	El módulo constituye la capa final de visualización del sistema y traduce la información depurada y estructurada en herramientas de análisis para la gestión institucional.

Cuadro 5: RQF-AL04 – Visualización estratégica y tablero de control Power BI

9.4. Requisitos funcionales de Bajo Nivel

Código	RQF-BL01
Nombre	Insertión controlada de actividades
Descripción	El sistema registrará nuevas actividades validando que los identificadores de tipo, fuente, línea y entidad existan previamente.
Usuarios	Administrador, Coordinación de Extensión.

Cuadro 6: RQF-BL01 – Inserción controlada de actividades

Código	RQF-BL02
Nombre	Validación de fechas en actividades
Descripción	El sistema verificará que la fecha de finalización sea igual o mayor que la fecha de inicio al crear o modificar actividades.
Usuarios	Administrador, Coordinación de Extensión.

Cuadro 7: RQF-BL02 – Validación de fechas

Código	RQF-BL03
Nombre	Registro único por persona
Descripción	El sistema garantizará que no existan duplicados en CCPersona, manteniendo un único registro por documento.
Usuarios	Administrador.

Cuadro 8: RQF-BL03 – Registro único por persona

Código	RQF-BL04
Nombre	Asociación actividad-participante
Descripción	El sistema solo permitirá registrar participaciones si la actividad y la persona existen en la base de datos.
Usuarios	Administrador, Gestores de Extensión.

Cuadro 9: RQF-BL04 – Asociación actividad-participante

Código	RQF-BL05
Nombre	Cálculo automático de horas totales
Descripción	La vista de ranking sumará las horas por participante para análisis institucional.
Usuarios	Analistas institucionales.

Cuadro 10: RQF-BL05 – Cálculo automático de horas totales

Código	RQF-BL06
Nombre	Indicadores de riesgo por criterios
Descripción	El sistema calculará riesgo combinando criterios definidos y comparados con promedios institucionales.
Usuarios	Dirección de Desarrollo.

Cuadro 11: RQF-BL06 – Indicadores de riesgo

Código	RQF-BL07
Nombre	Identificación de actividades de alto riesgo
Descripción	Actividades con dos o más condiciones críticas serán clasificadas automáticamente como de alto riesgo.
Usuarios	Dirección de Desarrollo.

Cuadro 12: RQF-BL07 – Identificación de actividades de alto riesgo

Código	RQF-BL08
Nombre	Ranking de participación institucional
Descripción	El sistema ordenará los participantes por número de actividades y horas totales para análisis.
Usuarios	Coordinación de Extensión.

Cuadro 13: RQF-BL08 – Ranking de participación

Código	RQF-BL09
Nombre	Consolidación de roles en actividades
Descripción	El sistema consolidará los roles desempeñados por cada actor dentro de las actividades registradas.
Usuarios	Analistas institucionales, Dirección de Desarrollo.

Cuadro 14: RQF-BL09 – Consolidación de roles

Código	RQF-BL10
Nombre	Consulta de desempeño por entidad
Descripción	El sistema permitirá consultar el desempeño de cada entidad mediante agrupaciones y promedios.
Usuarios	Dirección de Desarrollo.

Cuadro 15: RQF-BL10 – Desempeño por entidad

Código	RQF-BL11
Nombre	Cálculo de avance promedio por entidad
Descripción	La vista calculará el promedio de avance de las actividades asociadas a cada entidad.
Usuarios	Analistas institucionales.

Cuadro 16: RQF-BL11 – Avance promedio por entidad

Código	RQF-BL12
Nombre	Agrupación de actividades por tipo
Descripción	Permitirá consultar actividades agrupadas por tipo de actividad para análisis del portafolio.
Usuarios	Coordinación de Extensión.

Cuadro 17: RQF-BL12 – Agrupación por tipo de actividad

Código	RQF-BL13
Nombre	Filtro de actividades por periodo
Descripción	Permitirá realizar consultas de actividades por periodo académico o administrativo.
Usuarios	Usuarios operativos.

Cuadro 18: RQF-BL13 – Filtro por periodo

Código	RQF-BL14
Nombre	Cálculo de rentabilidad financiera
Descripción	La vista calculará indicadores financieros como ingresos, egresos y ejecución para análisis.
Usuarios	Dirección Financiera, Dirección de Desarrollo.

Cuadro 19: RQF-BL14 – Rentabilidad financiera

Código	RQF-BL15
Nombre	Consulta de actividades por ODS
Descripción	La vista permitirá analizar las actividades asociadas a cada ODS a través de la línea temática.
Usuarios	Analistas institucionales.

Cuadro 20: RQF-BL15 – Actividades por ODS

Código	RQF-BL16
Nombre	Conteo de actividades por ODS
Descripción	Permitirá contar actividades por cada ODS asociado para monitoreo institucional.
Usuarios	Dirección de Desarrollo.

Cuadro 21: RQF-BL16 – Conteo de actividades por ODS

Código	RQF-BL17
Nombre	Avance promedio por ODS
Descripción	Permitirá calcular el avance promedio de actividades asociadas a cada ODS.
Usuarios	Dirección de Desarrollo.

Cuadro 22: RQF-BL17 – Avance por ODS

Código	RQF-BL18
Nombre	Clasificación automática del estado de actividades
Descripción	El sistema clasificará automáticamente las actividades como terminadas, en curso o retrasadas según la fecha y el avance.
Usuarios	Coordinación de Extensión.

Cuadro 23: RQF-BL18 – Estado de actividades

Código	RQF-BL19
Nombre	Validación automática de estados
Descripción	El sistema verificará que una actividad retrasada cumpla condiciones de fecha actual y avance menor al definido.
Usuarios	Administrador.

Cuadro 24: RQF-BL19 – Validación de estados

Código	RQF-BL20
Nombre	Consulta de actividades por fuente de financiamiento
Descripción	Permitirá agrupar y analizar actividades según su fuente de financiamiento para seguimiento financiero.
Usuarios	Dirección Financiera.

Cuadro 25: RQF-BL20 – Actividades por fuente de financiamiento

Código	RQF-BL21
Nombre	Cálculo de ejecución financiera promedio
Descripción	La vista calculará la ejecución financiera promedio en función de ingresos y egresos.
Usuarios	Dirección Financiera.

Cuadro 26: RQF-BL21 – Ejecución financiera

Código	RQF-BL22
Nombre	Generación de vistas optimizadas para Power BI
Descripción	Las vistas deben estar estructuradas de forma óptima para ser consumidas directamente por Power BI.
Usuarios	Analistas institucionales.

Cuadro 27: RQF-BL22 – Vistas para Power BI

Código	RQF-BL23
Nombre	Conteo de actividades por programa académico
Descripción	Permitirá consultar el número de actividades asociadas a cada programa académico.
Usuarios	Coordinación Académica.

Cuadro 28: RQF-BL23 – Actividades por programa académico

Código	RQF-BL24
Nombre	Cálculo de horas totales por programa
Descripción	La vista consolidará y sumará horas dedicadas a actividades clasificadas por programa.
Usuarios	Analistas institucionales.

Cuadro 29: RQF-BL24 – Horas por programa

Código	RQF-BL25
Nombre	Validación de llaves foráneas
Descripción	El sistema rechazará registros cuyo identificador foráneo no exista en las tablas correspondientes.
Usuarios	Administrador.

Cuadro 30: RQF-BL25 – Validación de llaves foráneas

Código	RQF-BL26
Nombre	Consulta de actividades por entidad
Descripción	Permitirá consultar las actividades organizadas o asociadas por cada entidad institucional.
Usuarios	Coordinación de Extensión.

Cuadro 31: RQF-BL26 – Actividades por entidad

Código	RQF-BL27
Nombre	Avance promedio del portafolio
Descripción	Permitirá calcular el avance promedio de todas las actividades para monitoreo institucional.
Usuarios	Dirección de Desarrollo.

Cuadro 32: RQF-BL27 – Avance general del portafolio

Código	RQF-BL28
Nombre	Análisis del comportamiento del portafolio
Descripción	Permitirá consultar volumen, tipo, modalidad, financiamiento y avance del portafolio total de actividades.
Usuarios	Dirección de Desarrollo.

Cuadro 33: RQF-BL28 – Comportamiento del portafolio

Código	RQF-BL29
Nombre	Consulta de riesgo financiero
Descripción	Permitirá identificar actividades con gastos atípicos o desproporcionados respecto a los promedios institucionales.
Usuarios	Dirección Financiera, Coordinación de Extensión.

Cuadro 34: RQF-BL29 – Riesgo financiero

Código	RQF-BL30
Nombre	Agrupación de actividades por modalidad
Descripción	Permitirá analizar actividades según modalidad (presencial, virtual, híbrida).
Usuarios	Coordinación de Extensión.

Cuadro 35: RQF-BL30 – Actividades por modalidad

Código	RQF-BL31
Nombre	Consulta por línea temática
Descripción	Permitirá consultar actividades asociadas a una línea temática específica.
Usuarios	Coordinación Académica, Docentes.

Cuadro 36: RQF-BL31 – Actividades por línea temática

Código	RQF-BL32
Nombre	Reporte de participación por rol
Descripción	Permitirá generar reportes consolidados de participación clasificados por rol institucional.
Usuarios	Analistas institucionales.

Cuadro 37: RQF-BL32 – Participación por rol

Código	RQF-BL33
Nombre	Consolidación presupuestal por actividad
Descripción	El sistema permitirá sumar ingresos y egresos para cada actividad de extensión.
Usuarios	Dirección Financiera.

Cuadro 38: RQF-BL33 – Consolidación presupuestal

Código	RQF-BL34
Nombre	Consulta por país o nacionalidad
Descripción	Permitirá consultar personas y entidades según país de origen o actividad.
Usuarios	Coordinación de Extensión.

Cuadro 39: RQF-BL34 – Consulta por país

Código	RQF-BL35
Nombre	Validación de avance
Descripción	El sistema validará que el avance declarado esté dentro del rango permitido por el modelo (0 a 1).
Usuarios	Administrador.

Cuadro 40: RQF-BL35 – Validación de avance

Código	RQF-BL36
Nombre	Agrupación temporal de actividades
Descripción	Permitirá analizar actividades según periodos definidos para evaluación temporal.
Usuarios	Dirección de Desarrollo.

Cuadro 41: RQF-BL36 – Agrupación temporal

Código	RQF-BL37
Nombre	Consulta de actividades por fuente original
Descripción	Permitirá consultar actividades según la unidad o fuente desde la que se originan.
Usuarios	Coordinación de Extensión.

Cuadro 42: RQF-BL37 – Actividades por fuente original

Código	RQF-BL38
Nombre	Reporte de tamaños del portafolio
Descripción	Permitirá consultar el tamaño del portafolio mediante conteos por tipo, línea y entidad.
Usuarios	Dirección de Desarrollo.

Cuadro 43: RQF-BL38 – Tamaño del portafolio

Código	RQF-BL39
Nombre	Identificación de entidades más activas
Descripción	Permitirá identificar qué entidades participan con mayor frecuencia en actividades de extensión.
Usuarios	Dirección de Desarrollo.

Cuadro 44: RQF-BL39 – Entidades más activas

Código	RQF-BL40
Nombre	Identificación de personas más activas
Descripción	Permitirá identificar a las personas que registran mayor participación y dedicación en actividades de extensión.
Usuarios	Coordinación de Extensión, Dirección de Desarrollo.

Cuadro 45: RQF-BL40 – Personas más activas

10. Modelo Entidad–Relación (MER)

El diseño del sistema de información propuesto requiere partir de un Modelo Entidad–Relación (MER) que permita representar de manera estructurada los elementos centrales del dominio de la extensión universitaria y las relaciones lógicas entre ellos. Este modelo constituye la base conceptual del sistema y orienta la construcción del modelo relacional físico utilizado en el piloto. A partir del análisis funcional, documental y operativo realizado con la Dirección de Desarrollo y Extensión, se identificaron las entidades principales, sus atributos y las relaciones necesarias para garantizar trazabilidad, consistencia y capacidad analítica.

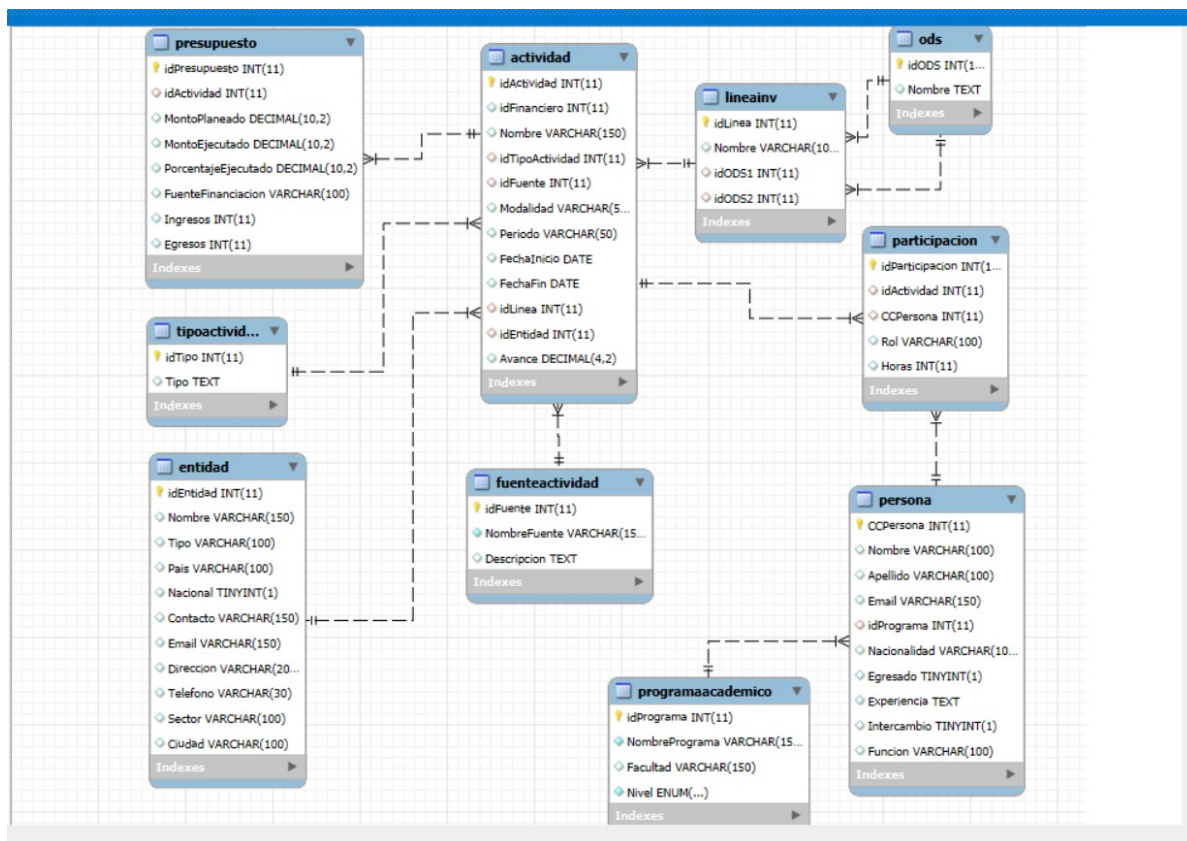


Figura 3: Modelo Entidad-Relación del Sistema de Información de Extensión

10.1. Identificación de entidades

A partir del levantamiento de información y del análisis de las tablas utilizadas por las unidades de extensión, se identificaron las siguientes entidades principales que estructuran el sistema:

- **Actividad**: representa cada acción de extensión registrada por la Facultad.
- **Persona**: actores participantes en las actividades (docentes, estudiantes, egresados, aliados externos).
- **Participación**: entidad intermedia que vincula personas y actividades, permitiendo registrar rol, dedicación y tipo de participación.
- **Entidad**: organizaciones externas o unidades aliadas que colaboran, aportan o desarrollan actividades.
- **Línea de investigación**: clasifica actividades y permite vincularlas con prioridades académicas y con objetivos de desarrollo sostenible.

- **ODS:** objetivos globales asociados a líneas e indicadores de impacto.
- **Fuente de actividad:** origen administrativo o académico desde donde se genera una actividad.
- **Tipo de actividad:** categoría estructural que define la naturaleza de la actividad (consultoría, evento, curso, voluntariado, etc.).
- **Presupuesto:** información financiera y operativa vinculada a cada actividad.
- **Programa académico:** programas asociados a personas o actividades para análisis curricular y de participación.

Estas entidades surgen de los procesos institucionales reales y permiten representar con fidelidad la dinámica operativa de la extensión.

10.2. Relaciones principales del modelo

El MER captura las relaciones fundamentales entre las entidades que participan en el sistema. Con base en el análisis del funcionamiento institucional y la estructura del piloto, se identificaron las siguientes relaciones:

- **Actividad–Tipo de actividad:** una actividad pertenece a un tipo específico (relación uno a muchos).
- **Actividad–Fuente de actividad:** cada actividad proviene de una unidad o fuente institucional definida (uno a muchos).
- **Actividad–Entidad:** vincula actividades con organizaciones o unidades aliadas que participan en ellas (uno a muchos).
- **Actividad–Línea de investigación:** cada actividad se clasifica en una línea temática (uno a muchos).
- **Línea–ODS:** una línea puede contribuir a uno o varios ODS (uno a muchos).
- **Actividad–Participación–Persona:** las personas pueden participar en múltiples actividades y una actividad puede tener múltiples participantes; esto se modela mediante una entidad intermedia que permite capturar horas, rol y contribuciones (relación muchos a muchos).
- **Actividad–Presupuesto:** cada actividad posee un conjunto de atributos financieros que permiten medir ingresos, egresos y porcentaje de ejecución (uno a uno o uno a muchos según la implementación).
- **Persona–Programa académico:** cada persona se asocia a un programa académico para análisis y segmentación.

Estas relaciones son esenciales para consultas institucionales como comportamiento del portafolio, análisis por ODS, redes de participación, segmentación por rol y riesgos operativos.

10.3. Cardinalidades y estructura lógica del modelo

El MER se estructura de acuerdo con principios clásicos de cardinalidad que garantizan integridad y coherencia:

- Relaciones **uno a muchos**: predominan en la clasificación de actividades por tipo, fuente, entidad u origen. Esto reduce duplicación y facilita la estandarización.
- Relaciones **muchos a muchos**: resueltas mediante entidades intermedias como Participación, necesarias para modelar roles, dedicación y redes de colaboración.
- Relaciones **uno a uno**: aplicables a componentes financieros cuando se desea mantener un registro independiente sin sobrecargar la entidad principal.

El uso adecuado de estas cardinalidades permite que el sistema sea consultable, escalable y compatible con procesos posteriores de automatización y gobernanza del dato.

10.4. Componentes del modelo MER

El modelo MER del sistema se compone de:

- **Entidades con claves primarias definidas**: cada tabla posee un identificador único que garantiza trazabilidad.
- **Atributos normalizados**: los atributos representan significados únicos y evitan redundancia, siguiendo principios básicos de normalización.
- **Relaciones explícitas mediante claves foráneas**: permiten unir información entre entidades sin pérdida de significado.
- **Entidad intermedia para relaciones complejas**: la tabla Participación resuelve la relación muchos a muchos entre personas y actividades.
- **Catálogos estructurales**: tablas como TipoActividad, FuenteActividad y ODS funcionan como referencia obligatoria para registro sistemático.

Este andamiaje conceptual guía la construcción del modelo relacional físico y asegura que el sistema pueda soportar análisis complejos, indicadores y visualización estratégica.

10.5. Derivación del diccionario de datos a partir del MER

La construcción del Modelo Entidad–Relación (MER) permite no solo identificar las entidades y relaciones fundamentales del sistema, sino también derivar de manera estructurada el **diccionario de datos**, que constituye la especificación formal de cada tabla, sus atributos, tipos de datos, llaves primarias y llaves foráneas. Este diccionario asegura integridad semántica, facilita la gobernanza del dato y sirve como referencia técnica para el desarrollo del modelo relacional, las consultas SQL y la construcción de vistas analíticas.

A continuación se presenta el diccionario de datos completo, replicado de manera idéntica al documento técnico original proporcionado por la Facultad, que define la estructura final del sistema de información. Este diccionario corresponde directamente a las entidades identificadas en el MER y desarrolla sus atributos de acuerdo con las necesidades operativas y analíticas del proyecto.

Campo	Tipo	Descripción	Llave
idODS	INT	Identificador único del Objetivo de Desarrollo Sostenible.	PK
Nombre	TEXT	Nombre oficial del ODS según la clasificación de Naciones Unidas.	

Cuadro 46: Diccionario de datos — Tabla ODS

Campo	Tipo	Descripción	Llave
idLinea	INT	Identificador único de la línea de investigación.	PK
Nombre	VARCHAR(100)	Nombre oficial de la línea según la clasificación institucional.	
idODS1	INT	Primer ODS asociado a la línea, según la orientación temática.	FK
idODS2	INT	Segundo ODS asociado para completar la relación conceptual de la línea.	FK

Cuadro 47: Diccionario de datos — Tabla LineaInv

Campo	Tipo	Descripción	Llave
idTipo	INT	Identificador único del tipo de actividad.	PK
Tipo	TEXT	Nombre del tipo de actividad (Consultoría, MOOC, Educación Continua, Evento, Voluntariado, etc.).	

Cuadro 48: Diccionario de datos — Tabla TipoActividad

Campo	Tipo	Descripción	Llave
idFuente	INT AUTO INCREMENT	Identificador único de la fuente de actividad.	PK
NombreFuente	VARCHAR(150)	Nombre de la fuente tal como aparece en la hoja de Excel original.	
Descripcion	TEXT	Descripción asociada a la fuente según archivo institucional.	

Cuadro 49: Diccionario de datos — Tabla FuenteActividad

Campo	Tipo	Descripción	Llave
idPrograma	INT AUTO INCREMENT	Identificador único del programa académico.	PK
NombrePrograma	VARCHAR(150)	Nombre oficial del programa académico según registro institucional.	
Facultad	VARCHAR(150)	Facultad, escuela o unidad académica correspondiente.	
Nivel	ENUM	Clasificación del nivel académico del programa (Pregrado, Posgrado, Maestría).	

Cuadro 50: Diccionario de datos — Tabla ProgramaAcademico

Campo	Tipo	Descripción	Llave
idEntidad	INT	Identificador único de la entidad.	PK
Nombre	VARCHAR(150)	Nombre oficial de la entidad.	
Tipo	VARCHAR(100)	Clasificación de la entidad (Pública, Privada, ONG, Multilateral, etc.).	
Pais	VARCHAR(100)	País de origen o registro de la entidad.	
Nacional	BOOLEAN	Indica si la entidad es colombiana (1) o internacional (0).	
Contacto	VARCHAR(150)	Nombre del contacto principal.	
Email	VARCHAR(150)	Correo electrónico del contacto.	
Direccion	VARCHAR(200)	Dirección física registrada.	
Telefono	VARCHAR(30)	Teléfono de contacto.	
Sector	VARCHAR(100)	Sector institucional o económico.	
Ciudad	VARCHAR(100)	Ciudad de ubicación principal.	

Cuadro 51: Diccionario de datos — Tabla Entidad

Campo	Tipo	Descripción	Llave
CCPersona	INT	Identificador único de la persona.	PK FK
Nombre	VARCHAR(100)	Nombres de la persona.	
Apellido	VARCHAR(100)	Apellidos de la persona.	
Email	VARCHAR(150)	Correo electrónico asociado.	
idPrograma	INT	Programa académico al que pertenece.	
Nacionalidad	VARCHAR(100)	País de nacionalidad.	
Egresado	BOOLEAN	Indica si la persona es egresada.	
Experiencia	TEXT	Experiencia académica o laboral.	
Intercambio	BOOLEAN	Indica si ha participado en intercambios académicos.	
Funcion	VARCHAR(100)	Rol o función que cumple dentro de la comunidad.	

Cuadro 52: Diccionario de datos — Tabla Persona

Campo	Tipo	Descripción	Llave
idActividad	INT AUTO INCREMENT	Identificador único de la actividad.	PK FK FK FK FK
idFinanciero	INT	Código financiero o contable asociado.	
Nombre	VARCHAR(150)	Nombre completo de la actividad.	
idTipoActividad	INT	Clasificación por tipo de actividad.	
idFuente	INT	Fuente administrativa de procedencia.	
Modalidad	VARCHAR(50)	Modalidad (Presencial, Virtual, Híbrida).	
Periodo	VARCHAR(50)	Periodo académico.	
FechaInicio	DATE	Fecha oficial de inicio.	
FechaFin	DATE	Fecha oficial de finalización.	
idLinea	INT	Línea temática asociada.	
idEntidad	INT	Entidad aliada asociada.	
Avance	DECIMAL(4,2)	Porcentaje de avance (0 a 1).	

Cuadro 53: Diccionario de datos — Tabla Actividad

Campo	Tipo	Descripción	Llave
idParticipacion	INT AUTO INCREMENT	Identificador único del registro de participación.	PK FK FK
idActividad	INT	Actividad en la que participa la persona.	
CCPersona	INT	Persona participante.	
Rol	VARCHAR(100)	Rol desempeñado en la actividad.	
Horas	INT	Total de horas dedicadas.	

Cuadro 54: Diccionario de datos — Tabla Participacion

Campo	Tipo	Descripción	Llave
idPresupuesto	INT AUTO INCREMENT	Identificador único del registro presupuestal.	PK
idActividad	INT	Actividad asociada al presupuesto.	FK
MontoPlaneado	DECIMAL(10,2)	Valor total del presupuesto proyectado.	
MontoEjecutado	DECIMAL(10,2)	Valor ejecutado al cierre o corte.	
PorcentajeEjecutado	DECIMAL(10,2)	Relación entre ejecutado y planeado.	
FuenteFinanciacion	VARCHAR(100)	Fuente de financiación declarada.	
Ingresos	INT	Total de ingresos asociados.	
Egresos	INT	Total de egresos asociados.	

Cuadro 55: Diccionario de datos — Tabla Presupuesto

10.6. Ejecución del piloto técnico

Una vez definido el Modelo Entidad–Relación (MER) y construido el diccionario de datos correspondiente, se procedió a la implementación técnica del sistema mediante la creación de la base de datos relacional. Para ello se desarrolló un conjunto de instrucciones SQL que materializan la estructura del modelo, crean las entidades, definen los tipos de dato, establecen las claves primarias y foráneas, y configuran las relaciones necesarias para garantizar integridad referencial y consistencia semántica.

El desarrollo se realizó en **MySQL Workbench**, donde se construyó el esquema completo y se desplegó en un **servidor local** para facilitar la ejecución del piloto y la validación de las consultas, vistas y reglas definidas. La base de datos puede ejecutarse en cualquier equipo que cuente con un motor MySQL, ya que el código fue construido siguiendo estándares portables y sin dependencias externas. Esto asegura que el sistema sea replicable, verificable y utilizable por la Facultad en iteraciones posteriores o bajo entornos controlados de prueba.

Todo el código SQL del piloto, incluyendo la creación de tablas, definición de llaves, inserción de datos simulados, vistas analíticas y consultas asociadas a los requisitos funcionales, fue almacenado en un repositorio público de control de versiones en GitHub. Esto permite trazabilidad, facilita el versionamiento y garantiza que la estructura del sistema pueda mantenerse, revisarse y mejorarse de manera colaborativa. El repositorio utilizado es el siguiente:

<https://github.com/samuelBlancoCastellanos/Piloto-Proyecto-Pr-ctico.git>

El repositorio contiene el archivo **PilotoV3.sql**, que incluye toda la definición del esquema y las vistas empleadas en el piloto. Gracias a esta estructura, cualquier usuario autorizado puede replicar el sistema completo en su computador local, cargar la base de datos y ejecutar las consultas necesarias para evaluar el comportamiento del modelo y su utilidad para la gestión de la extensión universitaria.

11. Visualización de Datos e Indicadores

Con la base de datos implementada en MySQL y las consultas analíticas ejecutadas en el servidor local, se desarrolló la fase de visualización, cuyo propósito es transformar los resultados del modelo relacional en tableros interactivos que apoyen la interpretación institucional y la toma de decisiones. Para ello, se ejecutaron en SQL las vistas y consultas derivadas de los requisitos funcionales, generando tablas analíticas consistentes con el diseño del sistema.

Dado que el piloto no contaba con un volumen suficiente de datos históricos reales, se emplearon técnicas asistidas por inteligencia artificial para generar datos sintéticos que simulan el comportamiento operativo de la extensión universitaria. Estos datos artificiales permiten reproducir patrones de participación, asignación de roles, financiamiento, avance y clasificación temática, garantizando que las consultas y visualizaciones puedan probarse en un entorno completo y coherente.

Una vez generadas las tablas analíticas, la base fue conectada directamente a **Power BI** mediante el conector nativo de MySQL. De esta manera, el tablero consumió las vistas y consultas del sistema sin necesidad de transformaciones adicionales, preservando integridad semántica e independencia respecto a capas externas. En Power BI se construyeron visualizaciones diseñadas específicamente para responder las cinco preguntas estratégicas definidas en el piloto:

1. ¿Cómo se comporta el portafolio total de actividades de extensión en términos de volumen, tipo, modalidad, avance y financiamiento?
2. ¿Qué actividades presentan mayor riesgo operativo o financiero y cómo se comparan con los promedios generales?
3. ¿Qué entidades aportan más, tienen mejor desempeño y concentran un mayor número de actividades?
4. ¿Qué roles asumen las personas dentro de las actividades de extensión y quiénes participan más activamente?
5. ¿Cómo se distribuyen las actividades de extensión según los Objetivos de Desarrollo Sostenible y las líneas de investigación?

El tablero resultante permite filtrar información, comparar unidades, identificar comportamientos atípicos y observar tendencias institucionales. Esta capa constituye el componente final del piloto, pues convierte la estructura del modelo relacional en información de valor estratégico para la gestión de la extensión.

Finalmente, tanto el archivo del tablero (**EXTENSION.pbix**) como el código SQL que alimenta sus visualizaciones fueron almacenados en el repositorio oficial del proyecto, lo que garantiza trazabilidad, replicabilidad y disponibilidad para futuras iteraciones y pruebas:

12. Recomendaciones finales y siguientes pasos

El piloto permitió demostrar que es posible estructurar un sistema de información coherente para la extensión universitaria mediante el uso de bases de datos relacionales, reglas de gobernanza del dato y visualizaciones estratégicas. Sin embargo, la implementación completa del sistema requiere un proceso de fortalecimiento progresivo, que combine mejoras estructurales, automatización y ampliación del alcance. A continuación se presentan las recomendaciones y siguientes pasos que la Facultad debería desarrollar para consolidar el sistema en el corto y largo plazo.

12.1. 1. Fortalecimiento estructural inicial (0–3 meses)

En esta primera fase es fundamental consolidar la estructura mínima viable del sistema de información existente. Las acciones recomendadas incluyen:

- **Definir los requisitos mínimos del sistema.** Formalizar el conjunto básico de funcionalidades que se requieren para que el sistema opere de forma consistente y replicable.
- **Depuración y estandarización de las fuentes actuales.** Esto implica unificar formatos, corregir duplicados, identificar llaves inconsistentes y normalizar categorías institucionales.
- **Definir llaves primarias, versiones y reglas de validación.** Esto permitirá garantizar trazabilidad y reducir la variabilidad entre unidades que reportan datos.
- **Diagnóstico de brechas de información.** Identificar limitaciones temáticas, poblacionales y temporales que afectan la capacidad analítica del sistema.
- **Evaluar la información disponible y su completitud.** Determinar qué tablas deben priorizarse en la fase de escalamiento a partir del nivel de madurez del dato.

Estas acciones preparan el terreno para la automatización y para la incorporación futura del sistema en los procesos regulares de la Facultad.

12.2. 2. Construcción del proceso ETL institucional

El piloto utilizó datos sintéticos generados mediante inteligencia artificial. Sin embargo, un sistema institucional requiere integrar de manera automática las fuentes reales mediante un proceso formal de extracción, transformación y carga (ETL). Los pasos sugeridos incluyen:

- **Diseño del ETL** que reciba tanto las tablas que otras unidades remiten a Extensión como las tablas que Extensión construye manualmente.
- **Limpieza, estandarización y validación automática** de datos al momento de la carga.
- **Automatización del flujo** mediante scripts programados, contenedores o servicios institucionales.
- **Posible apoyo de inteligencia artificial** para mapear columnas, detectar anomalías, normalizar categorías y documentar reglas del proceso.

Aunque su construcción completa excede el alcance del piloto, la definición conceptual ya está establecida y constituye uno de los pasos estratégicos más importantes para consolidar el sistema.

12.3. 3. Escalamiento del modelo relacional

El piloto demostró la viabilidad técnica del modelo en un subconjunto de tablas. El siguiente paso es extender la estructura relacional a todos los módulos de extensión para que el sistema pueda operar sobre la totalidad del portafolio. Este escalamiento implica:

- **Replicar el MER y diccionario de datos** para todas las fuentes institucionales restantes.
- **Unificar las tablas históricas** y consolidar las variantes existentes entre años, unidades y procesos.
- **Incorporar nuevas entidades** según necesidades futuras, como beneficiarios, productos académicos, certificaciones o resultados de impacto.
- **Ampliar las vistas analíticas y consultas SQL** para cubrir nuevas preguntas estratégicas.

Con este escalamiento, el sistema permitirá pasar del análisis de un subconjunto de actividades a la capacidad de monitorear y gestionar el portafolio completo con criterios estandarizados.

12.4. 4. Exploración de bases de datos no relacionales

Aunque el sistema puede implementarse completamente en un modelo relacional, es recomendable explorar alternativas no relacionales en etapas posteriores, especialmente cuando se integren datos no estructurados (documentos, textos, comentarios, formularios ampliados). Entre las recomendaciones se encuentran:

- **Explorar bases tipo NoSQL** para almacenar documentos, registros flexibles o histó-

ricos complejos.

- **Evaluar arquitecturas híbridas** en las que la base relacional gestione la información estructurada y una base no relacional almacene componentes documentales o semiestructurados.
- **Conectar ambas arquitecturas a un mismo sistema analítico** mediante vistas unificadas o pipelines de integración.

Esto permitirá soportar funcionalidades futuras como repositorios de documentos, cargas masivas y análisis semánticos.

12.5. 5. Implementación de modelos analíticos y métricas de impacto

Superada la fase de integración de datos, el sistema podrá incorporar modelos analíticos más avanzados. Las acciones sugeridas son:

- **Diseño de indicadores y métricas de impacto** basados en evidencia y alineados con los ODS y líneas de investigación.
- **Desarrollo de modelos de riesgo operativo y financiero** para priorizar actividades críticas.
- **Aplicación de analítica avanzada** para entender patrones de participación, desempeño, alianzas y resultados.

Esto permitirá que la extensión evolucione hacia un sistema de monitoreo continuo y evaluación institucional basada en datos.

12.6. 6. Escalamiento, automatización y evolución del sistema

Finalmente, el sistema de información deberá evolucionar de un piloto controlado a un componente permanente de la gestión de la Facultad. Para ello se recomienda:

- **Estandarizar el proceso de registro de actividades** mediante formularios institucionales conectados directamente al sistema.
- **Automatizar la actualización de dashboards** mediante cargas programadas desde el servidor.
- **Incorporar chatbots internos** para apoyar procesos administrativos y consultas de datos.
- **Documentar completamente el sistema** incluyendo uso, diccionario de datos, reglas, validaciones y procedimientos.

- **Asegurar sostenibilidad técnica** mediante roles definidos de administración del sistema y control de versiones en GitHub.

Estas acciones permitirán que el sistema crezca de manera ordenada, sostenible y alineada con las necesidades académicas, administrativas y estratégicas de la Facultad.

Referencias

- Asociación Colombiana de Universidades, ASCUN. (2006). *Política para el fomento de la calidad de la educación y el compromiso social a través de los egresados* [Documento de trabajo]. Red SEIS – Seguimiento de Egresados de Instituciones de Educación Superior.
- Barros da Silva, E., Cavalcanti Ramos, J. L., & Gomes Fonseca, P. (2024). *Desarrollo e implementación del SIGEX: un sistema de gestión para acciones de extensión universitaria*. Em Extensão, 23(2), 74–87.
- Bravo Cedeño, M. C., & Aviles Sotomayor, V. M. (2020). *Influencia de los procesos administrativos en la calidad productiva de las instituciones de educación superior y extensiones universitarias de Manabí*. RECIMUNDO, 4(3), 251–266. [https://doi.org/10.26820/recimundo/4.\(3\).julio.2020.251-266](https://doi.org/10.26820/recimundo/4.(3).julio.2020.251-266)
- Córdova Negrete, M. G., Domínguez Toala, G. del P., & Córdova Cabrera, D. J. (2025). *Retos y perspectivas de la gestión administrativa en la educación superior: fortalecimiento institucional, calidad educativa y liderazgo académico en el contexto globalizado*. Multidisciplinary Journal of Sciences, Discoveries, and Society, 2(2), e-207. <https://doi.org/10.71068/xzb5wn45>
- European University Association. (2021). *Universities without walls: A vision for 2030*.
- Garrido Córdoba, J. (2025). *Hacia un modelo de gobernanza de datos: Diagnóstico y recomendaciones para el Centro Regional Universitario de San Miguelito, 2025*. Revista Especializada de Ingeniería y Ciencias de la Tierra, 5(1). <https://doi.org/10.48204/reicit.v5n1.7682>
- Insunza Peña, R. (2020). *Diseño de un sistema de información de extensión, vinculación con el medio y comunicaciones en el campus sur de la Universidad de Chile* [Memoria de título de Ingeniero Civil Industrial]. Universidad de Chile.
- Januszewski, S. (2017). *La gestión de la información académica en la Universidad de Buenos Aires. Las dimensiones de los sistemas de información*. RELAPAE, (7), 96–105.
- Kochhann, A., & Monteiro, C. (Eds.). (2024). *Extensión universitaria en América Latina: construyendo una praxis crítica y emancipadora*. Kelps.
- Londoño Trujillo, Á. M. (2020). Los indicadores de gestión en la extensión universitaria: un estudio en Instituciones de Educación Superior de Colombia durante el 2018 (Trabajo de

- grado, Maestría en Dirección). Universidad del Rosario.
- Observatorio Iberoamericano de la Ciencia, la Tecnología y la Sociedad (OCTS–OEI), & Red Iberoamericana de Indicadores de Ciencia y Tecnología (RICYT). (2017). *Manual iberoamericano de indicadores de vinculación de la universidad con el entorno socioeconómico: Manual de Valencia*.
- Organisation for Economic Co-operation and Development (OECD). (2020). *Education at a glance 2020: OECD indicators*. OECD Publishing. <https://doi.org/10.1787/69096873-en>
- Pérez Martínez, A., Aguilar Hinojosa, J. J., & Rodríguez Fernández, A. (2018). *Gobernanza y gestión universitaria en Latinoamérica*. Revista Venezolana de Gerencia, Vol. Esp.(1). <https://www.redalyc.org/articulo.oa?id=29062781004>
- Piñeiro Gómez, J. M. (2013). *Bases de datos relacionales y modelado de datos*. Ediciones Paraninfo.
- Rincón Zapata, C., & Pérez Cadavid, A. (2018). *Metodología para la evaluación de impacto de la extensión en la educación superior*. Económicas CUC, 39(1), 137–152. <https://doi.org/10.17981/econuc.39.1.2018.09>
- Rodríguez, L., García, Y., Alfonso, R. (2020). Sistema de información gerencial para la gestión de la extensión universitaria. Rev. Cub. Educ. Sup., 39(3), 1–15.
- Sarfo, E. A., Adu-Twum, H. T., Mensah, P., Ababio, M. N., Ayannusi, A. O., Emuveyan, O., Afolayan, A. F., Bakare, O. I., & Tawo, O. E. (2025). *Architecting scalable data pipelines for learning analytics in higher education: A cloud-native approach*. Engineering Science & Technology Journal, 6(6). Fair East Publishers.
- Thomas, G. (2007). *The DGI Data Governance Framework*. The Data Governance Institute. <https://www.datagovernance.com>
- Torres Naranjo, M. A., Arias Hernández, J. S. (2017). Modelo de medición de impacto para los proyectos sociales de la Facultad de Ingeniería en la Universidad Católica de Colombia (Tesis de pregrado). Universidad Católica de Colombia.
- Universidad de Santander. (2019). Modelo de medición de impacto social de la extensión universitaria. Vicerrectoría de Extensión, Universidad de Santander.