Informe final de proyecto de investigación. Año 2023

Informe final

(carátula)

Programa Universitario de Investigación en Recursos Naturales y Ambiente (PUIRNA)

(nombre del programa universitario de investigación de la Digi)

Diseño de un Sistema de Información de Investigación Actual (CRIS) para la Universidad de San Carlos de Guatemala.

nombre del proyecto de investigación

4.8.24.4.74

Partida presupuestaria

Centro Universitario de Oriente CUNORI

unidad académica o centro no adscrito a unidad académica avaladora

Samy Eunice Pinto Castañeda - Coordinadora
César Emilio Casasola Miranda - Investigador
Hermas René Humberto Ramírez Rodríguez - Auxiliar de investigación I
Luis Diego Ambrocio Ramírez - Auxiliar de investigación I
Jeyson Federico Martínez López - Auxiliar de investigación I
Mayker Abiel Díaz Gudiel - Auxiliar de investigación I

nombre del coordinador del proyecto y equipo de investigación contratado por Digi

Chiquimula, Guatemala, 21/11/2023

lugar y fecha de presentación del informe final dd/mm/año

Informe final de proyecto de investigación. Año 2023

Autoridades

Dra. Alice Burgos Paniagua
Directora General de Investigación

Ing. Agr. MARN Julio Rufino Salazar
Coordinador General de Programas

MSc. Inga. Andrea Rodas Morán
Coordinadora del Programa de Investigación

Autores

M.A. Inga. Samy Eunice Pinto Castañeda
Coordinadora del proyecto

Ing. César Emilio Casasola Miranda
Investigador

Hermas René Humberto Ramírez Rodríguez
auxiliar de investigación I

Luis Diego Ambrocio Ramírez
auxiliar de investigación I

Jeyson Federico Martínez López
auxiliar de investigación I

Mayker Abiel Díaz Gudiel
auxiliar de investigación I

Colaboradores (si aplica): anotar nombres apellidos e institución que representa

Universidad de San Carlos de Guatemala, Dirección General de Investigación (Digi), 2023. El contenido de este informe de investigación es responsabilidad exclusiva de sus autores.

Esta investigación fue cofinanciada con recursos del Fondo de Investigación de la Digi de la Universidad de San Carlos de Guatemala a través de la partida presupuestaria 4.8.24.4.74 en el Programa Universitario de Investigación en Recursos Naturales y Ambiente (PUIRNA).

Los autores son responsables del contenido, de las condiciones éticas y legales de la investigación desarrollada.

1 Índice general

1 Índice general 5
Índice de tablas 9
Índice de figuras 10

2 Resumen 11
Abstract 12
3 Introducción 13
4 Antecedentes 14
5 Planteamiento del problema 16
6 Marco teórico 17
6.1 CRIS 17
6.2 CIENCIA ABIERTA 18
6.3 LEY DE ACCESO A LA INFORMACIÓN PÚBLICA 19
6.4 RESEARCH INFORMATION MANAGEMENT (RIM) 20
6.5 REPOSITORIOS INSTITUCIONALES 20
6.6 METADATA 20
6.7 EUROCRIS 21
6.8 CERIF 21
6.9 DSpace-CRIS 22
6.10 ORCID 22
6.11 OPENAIRE 23
6.12 HÉRCULES 23
7 Estado del arte 24
8 Objetivos (generales y específicos aprobados en la propuesta) 25
8.1 General 25
8.2 Específicos 26
9 Materiales y métodos 26
9.1 Delimitación en tiempo 26

9.2 Delimitación de área geográfica 26
9.3 Enfoque de la investigación 27
9.4 Método 27
9.5 Recolección de información 28
9.6 Técnicas e instrumentos 29
10 Resultados y discusión 31
10.1 Resultados objetivo 1 32
10.2 Resultados objetivo 2 44
10.3 Resultados objetivo 3 52
Requerimientos de Hardware para el sistema CRIS (de acuerdo a documentación de DSpace) 58
Instancias Familia T3a 63
EC2 64
RDS para MySQL 64
RDS para PostgreSQL 65
EBS 66
Recomendación de DLP 66
EC2 66
EBS 67
RDS 67
Requisitos y Características Adicionales a Considerar en el Ambiente Tecnológico para la Implementación del Sistema de Información de Investigación 68
Mecanismos esenciales a considerar que garantizan la seguridad de la API diseñada 68
¿Qué es la seguridad de las API web? 68
¿Por qué son tan cruciales la autenticación y la autorización para la protección de la API? 69
Configurar Recursos 70
Riesgos 70
Prácticas recomendadas más comunes en relación con la seguridad de las API 71
Servicios y herramientas para la gestión y la seguridad de las API 73
Red Hat 73
Akamai 74
Microsoft Azure 74
Endpoint DLP 76
SSO (Single Sign On) LastPass 77
Autorización con OAuth 78
Seguridad de la API de DSpace-CRIS 80
10.4 Resultados objetivo 4 81

Diseño arquitectónico del Sistema de Información de Investigación Actual para USAC 81
Descripción general del diseño arquitectónico del sistema 81
Diseño de arquitectura general del sistema (opción 1. Módulo de indicadores dentro de DSPACE CRIS) 81
Diseño de arquitectura general del sistema (opción 2. Módulo de indicadores dentro de RIMS) 82
DSPace 7 83
SISTEMAS A CONSIDERAR PARA LA INTEROPERABILIDAD 84
Información proporcionada por los stakeholders 85
Necesidades específicas de consumo y/o manipulación de datos entre los sistema 85
Alternativas para acceder y modificar la base de datos, considerando las características específicas del modelo relacional de la base de datos de DSpace CRIS. 85
Caracterización de las alternativas consideradas para el acceso y modificación de la base de datos 86
Diagrama de tecnologías utilizadas en DSpace CRIS y RIMS (módulo de investigadores) 88
Diagrama de organización interna de DSpace CRIS y RIMS (módulo de investigadores) 89
Arquitectura de Módulo de Indicadores Sistema CRIS 91
Arquitectura de Módulo de Indicadores Sistema RIMS 92
Registro de investigadores en el sistema RIMS 93
Consumo de datos del servicio ORCID 95
Definición de los Endpoints que se deberán proveer a terceros con los datos calculados de los indicadores de investigación 96
Estados de las Peticiones de los EndPoints 97
Correspondencia entre los datos que se obtendrán de los servicios de sistemas de terceros y los atributos de las entidades definidas en el modelo relacional del proyecto 98

11 Conclusiones 102
12 Recomendaciones 104
13 Referencias 105
14 Apéndice 113
15 Vinculación 229
16 Estrategia de difusión y divulgación 230
17 Aporte de la propuesta de investigación a los Prioridades Nacionales de Desarrollo (PND) identificando su meta correspondiente: 230
18 Orden de pago final (incluir únicamente al personal con contrato vigente al 31 de diciembre de 2023) 231
19 Declaración del Coordinador(a) del proyecto de investigación 231

20 Aval del Director(a) del instituto, centro o departamento de investigación o Coordinador de investigación del centro regional universitario 231
21 Visado de la Dirección General de Investigación 232

Índice de tablas

Tabla 1. Criterios de inclusión y exclusión para el personal administrativo y de informática de la Dirección General de Investigación. 33
Tabla 2. Cantidad de individuos por cargo que fueron incluidos en la muestra del personal administrativo y de informática. 34
Tabla 3. Criterios de inclusión y exclusión para los investigadores. 34
Tabla 4. Unidades académicas de los investigadores incluidos en la muestra. 36
Tabla 5. Criterios de inclusión y exclusión para el personal con experiencia en la generación de indicadores de investigación para la Dirección General de Investigación. 38
Tabla 6. Cuestionarios aplicados a los individuos de la muestra. 39
Tabla 7. Descripción de las columnas del diccionario de metadatos. 44
Tabla 8. Esquemas empleados en el diccionario de metadatos. 48
Tabla 9. Descripción de las columnas del diccionario de metadatos. 51
Tabla 10. Resumen del análisis de la demanda del portal DIGI. 57
Tabla 11. Requerimientos mínimos DSPACE. 58
Tabla 12. Requisitos sistema DSPACE CRIS. 59
Tabla 13. Requisitos sistema RIMS. 59
Tabla 14. Requerimientos de Hardware Sistema DSPACE CRIS. 59
Tabla 15. Requerimientos de Hardware Sistema RIMS. 60
Tabla 16. Cotización de Costos. 63
Tabla 17. Características Amazon EC2. 64
Tabla 18. Características RDS para MySQL. 65
Tabla 19. Características RDS para PostgreSQL. 65
Tabla 20. Características Amazon EBS. 66
Tabla 21. Endpoints para obtención de datos del SIIF. 99

Índice de figuras

Figura 1. Diagrama de flujo de la aplicación 37
de criterios de inclusión y exclusión para investigadores. 37
Figura 2. Vista parcial del diagrama entidad-relación. 40
Figura 3. Vista parcial del diagrama relacional que incluye las tablas que se usarán en el RIMS. 41
Figura 4. Vista parcial del diagrama relacional que incluye las tablas que apoyarán el cálculo de indicadores. 43
Figura 5. Vista parcial del diccionario de datos. 46
Figura 6. Modelo de contenidos de DSpace. 47
Figura 7. Vista parcial del diccionario de metadatos. 50
Figura 8. Solicitudes atendidas por el portal de la Web DIGI. 53
Figura 9. Valor promedio de request para cada hora del día. 54
Figura 10. Máximo valor de request para cada hora del día. 55
Figura 11. Máximo registro de requests/segundo para cada hora. 56
Figura 12. Arquitectura AWS DSPACE CRIS. 61
Figura 13. Arquitectura AWS RIMS. 62
Figura 14. Cómo funciona una API web. 69
Figura 15. Security (OAuth 2.0). 75
Figura 16. Azure API Management. 76
Figura 17. Endpoint DLP Plus. 77
Figura 18. SSO Lastpass. 78
Figura 19. Consumo de datos del Prototipo OAuth 2.0. 79
Figura 20. Endpoints de OAuth 2.0 en Spring Boot. 80
Figura 21. Diseño de arquitectura general del sistema opción 1. 82
Figura 22. Diseño de arquitectura general del sistema opción 2. 83
Figura 23. Diagrama de tecnologías utilizadas en DSpace CRIS y RIMS. 89
Figura 24. Diagrama de organización interna de DSpace CRIS y RIMS. 90
Figura 25. Arquitectura de Módulo de Indicadores Sistema CRIS. 91
Figura 26. Arquitectura de Módulo de Indicadores Sistema RIMS. 92
Figura 28. Consumo de datos del servicio ORCID. 95

2 Resumen

La información asociada y generada por los procesos de investigación de la Universidad de San Carlos de Guatemala (USAC) no se gestiona de manera óptima. Esto repercute en dificultades para la colaboración e intercambio de información y en trabajo científico con poca visibilidad y reconocimiento. Universidades e institutos de investigación a nivel mundial han experimentado mejoras en la administración de la información de investigación a través de la implementación de sistemas CRIS (Current Research Information System, por sus siglas en inglés).

Un sistema CRIS beneficiará a la USAC al agilizar la administración de la información de los investigadores y la construcción de indicadores; potenciando la visibilidad y reconocimiento en el contexto nacional e internacional. Los objetivos propuestos fueron: diseñar el modelo de base de datos basado en el estándar CERIF, definir un diccionario de datos a partir del modelo, identificar requerimientos técnicos y características necesarias para implementar el sistema y diseñar mecanismos de integración que aseguren la interoperabilidad. La investigación aplicó el método investigación-acción, con un alcance descriptivo, y tuvo como resultado el diseño del sistema formalizado en cuatro productos: diseño entidad-relación ER y diseño físico de la base de datos, diccionario de datos, especificación de requerimientos y diseño arquitectónico del sistema.

Palabras clave: visibilidad, interoperabilidad, repositorio, administración, indicadores.

Abstract

The information associated with and generated by the research processes at the Universidad de San Carlos de Guatemala (USAC) is not managed optimally. This results in difficulties for collaboration and information exchange, as well as limited visibility and recognition of scientific work. Universities and research institutes worldwide have experienced improvements in research information management through the implementation of CRIS (Current Research Information System) systems.

A CRIS system will benefit USAC by streamlining the management of researchers' data and indicator construction; thereby enhancing visibility and recognition at both the national and international levels. The proposed objectives were: to design a database model based on the CERIF standard, define a data dictionary based on the model, identify technical requirements and necessary features for system implementation, and design integration mechanisms to ensure interoperability. The research applied the action research method with a descriptive scope, resulting in the formalized design of the system, which included four products: both entity-relationship (ER) and physical database design, a data dictionary, requirement specifications, and architectural system design.

Keywords: visibility, interoperability, repository, management, indicators.

3 Introducción

La Universidad de San Carlos de Guatemala se caracteriza por su compromiso con los ejes estratégicos: investigación, docencia y extensión. A pesar de los avances significativos en materia de investigación, existen aspectos que limitan la agilidad de los procesos de investigación, así como el alcance y visibilidad de los resultados. De manera que se identifica como problema, el manejo ineficiente de la información relacionada a las investigaciones que desarrolla la USAC.

Para resolver el problema expuesto, se propone el diseño de un sistema CRIS. Los sistemas de investigación CRIS (Current Research Information System, por sus siglas en inglés) son herramientas que permiten la gestión de la investigación en una institución académica o científica. Estos sistemas son de gran importancia ya que facilitan el seguimiento y la gestión de los proyectos de investigación, las publicaciones y las colaboraciones de los investigadores. Según Jetten, et al. (2019), los sistemas CRIS se han convertido en una opción relevante para los investigadores en cuanto al registro, la gestión y la exposición en línea de información sobre su investigación.

De acuerdo con Azeroual, et al. (2021), los sistemas CRIS están ganando importancia en diversas universidades internacionales. También resaltan las características únicas de este sistema, en comparación con otros, como su capacidad para integrar sistemas internos de la universidad con diversas fuentes de datos externas en una sola plataforma. De manera que se facilita el seguimiento de los proyectos en curso, incluyendo su financiación, el personal involucrado y los plazos de entrega.

Un sistema CRIS permitirá la gestión eficiente de los proyectos de investigación y puede tener un impacto positivo en la productividad y en los resultados de la investigación. La información sobre la investigación puede utilizarse para identificar áreas de fortaleza y debilidad en la institución, y para detectar oportunidades para el crecimiento y la colaboración. La información sobre la investigación puede ser compartida con la comunidad académica y científica a través de la publicación en bases de datos y repositorios en línea. La mejora en

visibilidad y el impacto de la investigación pueden contribuir a atraer financiación de otras fuentes y a otros investigadores de alto nivel.

La propuesta de diseño se enfoca en cuatro objetivos específicos, siendo éstos: el diseño del modelo de base de datos acoplado al estándar CERIF, la definición de un diccionario de datos a partir del modelo, la identificación de los requerimientos técnicos y características para el sistema y el diseño de la arquitectura del sistema, con mecanismos de integración que aseguren la interoperabilidad. Considerando los objetivos y la naturaleza de la investigación, se plantea emplear un diseño mixto.

4 Antecedentes

Hasta la fecha, la gestión de la información vinculada a la investigación se ha restringido principalmente a la existencia de formularios web de las convocatorias para proyectos de investigación, biblioteca virtual y el portal oficial. Estos canales han servido como medios de comunicación unidireccionales y depósitos de información; sin embargo, han carecido de la capacidad necesaria para gestionar de manera eficiente y centralizada la diversidad de datos generados por las actividades de investigación.

Adicionalmente, se implementó un esquema operativo conocido como Registro Universitario de Investigadores (RUI). Este sistema implica la asignación de un código único a cada coordinador de investigación y la recopilación de datos relacionados para almacenarlos en un sistema de la Dirección General de Investigación. Este registro sirve como referencia en algunos procesos de seguimiento o estadísticas.

Hasta el momento, no se ha logrado una consolidación de datos relacionados con la investigación, lo que dificulta el acceso actual a información relevante. El seguimiento de proyectos de investigación se lleva a cabo mediante herramientas como correo electrónico, formularios de Google y carpetas en Google Drive, lo que se traduce en incrementos de tiempo en la gestión.

Informe final de proyecto de investigación. Año 2023

Aunque existen otros sistemas que podrían considerarse como aliados para la gestión completa de la investigación, actualmente operan de manera aislada, sin mecanismos de interoperabilidad entre ellos. La colaboración entre investigadores y equipos, con el fin de fomentar un entorno de trabajo colaborativo, ha sido impulsada por iniciativas individuales debido a la ausencia de una plataforma centralizada para este propósito.

La iniciativa actual busca llenar este vacío mediante la implementación de un Sistema de Información de Investigación Actual (CRIS), un sistema especializado diseñado para recopilar, organizar y facilitar el acceso a la información actualizada sobre la investigación en curso. Este sistema no solo permite el seguimiento de investigadores, proyectos de investigación y publicaciones sino que también actúa como un punto central para la gestión integral de datos relacionados con la investigación.

Simultáneamente, se integra el Sistema de Gestión de Información de Investigación (RIMS) en el ecosistema para proporcionar una infraestructura sólida para la gestión de información a lo largo del desarrollo de proyectos de investigación. El RIMS no solo respalda la captura eficiente de datos, sino que también facilita la integración con otros sistemas, asegurando la interoperabilidad y la fluidez en el intercambio de información.

En resumen, este proyecto representa un avance significativo en la gestión de la investigación, superando las limitaciones de enfoques tradicionales y estableciendo las bases para la gestión más eficiente y efectiva de la información relacionada con la investigación. Este ecosistema no solo aborda las necesidades actuales, sino que también sienta las bases para la evolución continua en la gestión de la investigación en el futuro.

5 Planteamiento del problema

Los investigadores dedican un porcentaje considerable de su jornada laboral en tareas administrativas, tales como la preparación de reportes dirigidos a los inversionistas. La organización Ex Libris (2020) condujo una investigación en la cual entrevistaron a 314 investigadores de diversas disciplinas en Estados Unidos, Reino Unido y Australia. El propósito era conocer detalles de la experiencia en el desarrollo de investigaciones. Encontraron que un 71% de los investigadores encuestados manifestaron que se ocupan en actividades administrativas de un 11% a 50% de su tiempo laboral.

La producción del conocimiento debe desarrollarse como una actividad colaborativa para lograr mejores resultados. De Castro (2018) señala una creciente tendencia de formar redes de soporte de investigación entre instituciones nacionales e internacionales. Pero también reconoce que las diferencias de las estructuras organizacionales, recursos humanos y políticas de las instituciones constituyen obstáculos para desarrollar y fortalecer las estructuras de colaboración.

El trabajo científico no culmina al hacer publicaciones, ya que el impacto de los resultados dependerá de cuántas personas los conocerán y emplearán posteriormente para generar más conocimiento. Para Basantes-Andrade, et al. (2022), la visibilidad de las publicaciones científicas influye en el posicionamiento y reconocimiento de la universidad o institución que las genera. Como consecuencia, se alcanzan mejores puestos en listas que clasifican universidades y se atrae inversión. Además, dichos autores identifican algunas razones por las que las publicaciones pierden visibilidad, por ejemplo: el uso de formatos análogos que las confinan a bibliotecas físicas. También, el empleo de métodos tradicionales de difusión y divulgación, tales como revistas, reportes, congresos y conferencias cumplen con su objetivo, pero no resultan suficientes para la sociedad contemporánea.

La Dirección General de Investigación de la USAC ha publicado dos informes de indicadores del sistema de investigación de la universidad. Los informes han abarcado los períodos: 2012-2014 y 2015-2018 (Sosa, et al., 2016; Sosa y García, 2019). El objetivo del primer informe fue "llevar a cabo un diagnóstico situacional de la investigación y así determinar los indicadores que ayudaron a detectar las potencialidades y necesidades de mejora" (Sosa, et

al., 2016). Según lo expresa Sosa, et al. (2016), se utilizó un cuestionario de 29 preguntas dirigido a las autoridades de las distintas unidades académicas.

Resumiendo, se identificaron cuatro problemas específicos, los cuales se sintetizan en las siguientes oraciones. Las tareas administrativas consumen tiempo que podría aprovecharse en actividades de investigación *per se*. Las características heterogéneas de las universidades e institutos dificultan la colaboración e intercambio de información. Muchas publicaciones científicas tienen bajo impacto por tener escasa o nula visibilidad. La integración y procesamiento manual de diferentes fuentes de información relacionadas a proyectos de investigación es un trabajo lento y extenso. Generalizando, se identifica el problema como una gestión ineficiente de la información relacionada a las investigaciones que desarrolla la USAC.

6 Marco teórico

6.1 CRIS

Rodríguez (2015) describió que, los CRIS (Current Research Information Systems o Sistemas de Información de Investigación Actual) se han convertido en el modelo preferido para integrar y tratar toda la información académica y de investigación de diferentes instituciones, con el objetivo de estudiar y evaluar los resultados de las investigaciones.

En esencia, los CRIS permiten la gestión global de todos los procesos relacionados con la investigación, proporcionando acceso y difundiendo la información científica para ayudar a los usuarios en el registro, notificación y toma de decisiones en relación con el proceso de investigación.

Según Asserson y Jeffery (2010), por lo general, en un CRIS se pueden encontrar datos relevantes acerca de proyectos, personas, unidades organizativas, programas de financiación, resultados de investigación, instalaciones, equipos y eventos.

Para Rodríguez (2015), existen cuatro categorías de CRIS basadas en su alcance geográfico:

- Institucionales, que manejan la información científica de una sola institución, como una universidad o un centro de investigación, como el CSIC.
- Regionales, que administran la información científica en un nivel geográfico determinado, como SICA2 en Andalucía, España.
- Nacionales, que manejan la información científica de múltiples instituciones dentro de un país, como Sweden ScienceNet en Suecia o CRIStin en Noruega. Estos CRIS también suelen mostrar las cifras de investigadores, proyectos, organizaciones y resultados.
- Internacionales, como IST World o EuroRIS-Net + Research Infrastructures Observatory, que se encargan de la información científica de más de un país.

Además, se pueden clasificar según otros criterios, como el tipo de información I+D+i o por áreas temáticas. Los CRIS proponen evitar la fragmentación de sistemas que existe en la administración de las instituciones, uniendo toda la información relativa a la actividad científica en un solo emplazamiento, lo que permite evaluar el impacto de la ciencia en la sociedad y comparar la actividad científica entre instituciones similares.

6.2 CIENCIA ABIERTA

Según Vázquez Tapia (2019), como resultado del movimiento de Acceso Abierto (Open Access) establecido por la Declaración de Berlín en 2003, el concepto de Ciencia Abierta (Open Science) o Ciencia Digital (Digital Science) es definido por la Comisión Europea (2016) como "un nuevo enfoque del proceso científico basado en el trabajo cooperativo y nuevas formas de difundir el conocimiento mediante el uso de tecnologías digitales y nuevas herramientas colaborativas" (p. 33). Como consecuencia de esta nueva perspectiva, surgen los sistemas CRIS (Current Research Information System) como un modelo para la gestión de toda la información relacionada con la actividad científica.

Para De Castro (2018), los financiadores de investigación, como la Comisión Europea o el Consejo de Investigación de Ingeniería y Ciencias Físicas del Reino Unido (EPSRC), suelen establecer políticas de Ciencia Abierta para regular tanto el Acceso Abierto como la

gestión de datos de investigación. Posteriormente, delegan la tarea de implementar estas políticas a las instituciones debido a su cercanía con los investigadores financiados y a su capacidad para diseñar y ejecutar estrategias de promoción que aborden la complejidad de las políticas que tienen diferentes requisitos según el financiador. Las bibliotecas de investigación institucionales se convierten en el punto de contacto principal para cumplir con las políticas de Ciencia Abierta al tener un conocimiento profundo del panorama político y acceso a los sistemas CRIS institucionales para validar los metadatos.

6.3 LEY DE ACCESO A LA INFORMACIÓN PÚBLICA

La Ley de Acceso a la Información Pública, Decreto Legislativo Número 57-2008, tiene por objeto, entre otros, garantizar el derecho al acceso a información pública y favorecer la rendición de cuentas. En su artículo 6, se establece que la Universidad de San Carlos de Guatemala es uno de los sujetos obligados por dicha ley. El artículo 9 indica que los reportes y estudios de los sujetos obligados constituyen información pública, por lo que son sujetos de la aplicación del principio de máxima publicación. El principio de máxima publicación hace referencia a que toda información pública "no podrá ser reservada ni limitada sino por disposición constitucional o legal" (LAIP. Art. 9. 22 de octubre de 2008). De manera que, resulta imperativo aplicar el principio de máxima publicación a toda la información de los procesos de investigación de la DIGI, investigadores y resultados.

El capítulo sexto de la Ley de Acceso a la Información Pública brinda directrices para el manejo de datos personales en el contexto de la publicación y visibilidad de la información generada por los sujetos obligados según el artículo 6 de la misma ley. Se enfatiza la protección de datos personales y se requiere garantizar que los mismos en posesión de los sujetos obligados sean "adecuados, pertinentes y no excesivos" (LAIP. Art. 30. 22 de octubre de 2008). También ordena adoptar medidas para proteger la seguridad e integridad de los datos personales.

6.4 RESEARCH INFORMATION MANAGEMENT (RIM)

Los sistemas CRIS forman parte de la clasificación de sistemas de gestión de información de investigación, RIM, por sus siglas en inglés. Para Bryant, et al. (2017), los sistemas de gestión de información de investigación son muy útiles para las instituciones académicas, ya que permiten una combinación entre lo local y lo global, generando nuevas oportunidades para obtener perspectivas en diferentes niveles, desde los departamentos hasta la institución en su totalidad, y la posibilidad de compartir esta información a nivel regional, nacional e incluso transnacional.

6.5 REPOSITORIOS INSTITUCIONALES

Rodríguez (2015) describe que antes de la implementación de los CRIS, las universidades y centros de investigación crearon los repositorios institucionales, que sirvieron como un primer paso hacia el acceso abierto de la producción científica de estas instituciones. Los repositorios institucionales tienen una función clave en la comunicación científica, ya que permiten a las instituciones aumentar la visibilidad de sus investigadores, compartiendo el acceso a las obras de sus miembros.

También, las organizaciones científicas y académicas de todo el mundo se benefician de un acceso más fácil y rápido a la investigación desarrollada en otras instituciones, lo que permite una mayor difusión de la información. Además, la generación e inserción de un trabajo científico en un repositorio institucional aumenta la proyección del perfil del autor a nivel mundial y el impacto de la investigación realizada, compitiendo con las publicaciones periódicas y el JCR (Journal Citation Reports). Los RI también permiten realizar estudios de producción científica e impacto de la investigación, y favorecen la preservación de las investigaciones producidas a largo plazo.

6.6 METADATA

Según la Organización Nacional de Estándares de Información (NISO, 2004, pp. 1-2), citado por Díaz (2014, p.28) los metadatos son información estructurada que describe,

interpreta y define una fuente especial de información para facilitar su recuperación. El objetivo principal de los metadatos utilizados en los archivos institucionales es describir la información buscada. Esto aumenta la visibilidad y la influencia del recurso.

Jetten, et al. (2019) definen que los metadatos deben ser ricos, comprensibles y completos para que otros investigadores puedan encontrarlos, comprenderlos y utilizarlos.

Miller (2022) propone una clasificación de metadatos, identificando tres tipos: descriptivos, administrativos y estructurales. Los metadatos descriptivos especifican detalles de un recurso o elemento. Tales datos pueden ser útiles para buscar el recurso dentro de un repositorio o base de datos. Los metadatos administrativos proveen elementos útiles para respaldar la gestión del recurso. Miller (2022) identifica como subtipos del tipo de metadatos administrativos: elementos técnicos y de preservación, metadatos de derechos de autor, acceso y reproducción y metadatos de uso del recurso. Por último, se define como metadatos administrativos a los que se orientan a describir la estructura interna de algún objeto digital complejo o las relaciones entre diferentes recursos.

6.7 EUROCRIS

Joint (2008) expone que EuroCRIS es una asociación de expertos en CRIS dedicados a mejorar la disponibilidad de información de investigación. La comunidad CRIS se rige por un estándar de datos recomendado, lo que asegura la cohesión de los sistemas de información. Surla, et al. (2013) añaden que euroCRIS es una organización no lucrativa y su objetivo principal es trabajar en el avance y la compatibilidad de los sistemas.

6.8 CERIF

Rodríguez (2015) indica que, el formato CERIF (Common European Research Information Format o Formato Común Europeo de la Información Científica), es muy popular y ampliamente utilizado en el ámbito científico. La Unión Europea lo ha recomendado a sus Estados miembros como un modelo estándar para la gestión e intercambio de datos de investigación. La Comisión Europea creó un grupo de expertos para

desarrollar el formato, cuya primera versión tenía una estructura simple y un enfoque de entrada único. Sin embargo, los requisitos de calidad y estructura de los datos han llevado a su evolución.

En el año 2000, la Comunidad Europea entregó la responsabilidad de CERIF a euroCRIS, que lo ha promovido desde entonces. Las conferencias euroCRIS, que se celebran cada dos años, se centran en temas como CRIS, repositorios y evaluación de la investigación. La última versión de CERIF aplica una capa semántica, y la Unión Europea quiere liderar el ejemplo al usarlo.

Además Nikolić et al. (2015) refieren al sistema CRIS UNS como un caso exitoso de implementación CRIS en conformidad con CERIF. Dicho sistema se ha estado desarrollando desde 2008 en la Universidad de Novi Sad en la República de Serbia.

6.9 DSpace-CRIS

De acuerdo a lo expresado por Nurminen (2014), DSpace-CRIS proporciona un modelo de datos personalizable, entidades CRIS como autoridad para los metadatos de los artículos y un marco para recuperar datos de publicación de las bases de datos bibliográficas.

6.10 ORCID

De acuerdo con Vázquez, et al. (2017), ORCID es una entidad que opera sin fines de lucro y colabora con diversas comunidades de investigación compuestas por investigadores, universidades, editoriales, agencias de financiación, repositorios de datos y asociaciones internacionales. A través de su integración con sistemas, ORCID otorga la posibilidad a los autores de recibir reconocimiento por su trabajo. Además, promueve la rapidez en la obtención de información y análisis, gracias a la habilitación de conexiones claras y seguras entre los investigadores.

6.11 OPENAIRE

Dijk y Van Meel (2010) mencionan que OpenAIRE es un proyecto financiado por la Comisión Europea, que inició en diciembre de 2009 y su desarrollo tuvo una duración de tres años. El proyecto OpenAIRE establecerá la infraestructura necesaria para ayudar a los investigadores en la publicación de investigaciones y para cumplir con las directrices ERC sobre acceso abierto.

Según De Castro et al. (2019), la inclusión gradual de los sistemas CRIS en la lista de proveedores de datos de OpenAIRE tendrá un impacto significativo en la cantidad y calidad de metadatos contextuales disponibles para los registros de resultados de investigación, incluyendo datos de investigación, publicaciones y otros productos basados en texto. A diferencia de otras plataformas, los registros CRIS a menudo contienen información detallada sobre aspectos como la financiación y la afiliación de la organización. Esta información adicional permitirá a OpenAIRE mejorar la calidad de la información proporcionada por el agregador.

6.12 HÉRCULES

De acuerdo con Hernández (2022), el proyecto Hércules - Semantics of University Research Data - es un proyecto CRUE (Conferencia de Rectores de las Universidades Españolas) con el objetivo de crear un Sistema de Gestión de la Investigación (RMS, por sus siglas en inglés) basado en datos semánticos abiertos que proporcione una visión global de los datos de investigación del Sistema Universitario Español (SUS, por sus siglas en inglés), para mejorar el control, análisis y potencial colaboración entre las universidades y la sociedad.

Según De-Castro (2022), el proyecto Hércules ha mostrado una clara intención de ofrecer el sistema CRIS a todas las universidades españolas que estén interesadas en esta propuesta. El proyecto Hércules, presentado en los eventos euroCRIS, está dirigido principalmente a una plataforma "interna", que se dice que puede interactuar y coexistir con otras soluciones disponibles en las universidades.

7 Estado del arte

Actualmente, los sistemas CRIS están evolucionando para satisfacer las necesidades cambiantes de las instituciones de investigación y de los usuarios de estos sistemas. Una de las tendencias más innovadoras y actuales es la capacidad de integrar información de diferentes fuentes y sistemas, tanto internos como externos, con otros sistemas institucionales, como los de gestión de recursos humanos y bibliotecas, para una gestión más eficiente de la información y la creación de perfiles más completos de los investigadores.

En línea con esta evolución, Beigel (2021) propone la creación de nuevos servicios de comunicación de la ciencia para fomentar la colaboración en la producción de conocimiento con sectores sociales y crear resultados que no se publiquen en revistas académicas. La creación de una infraestructura de tipo CRIS permitiría desarrollar indicadores apropiados para evaluar la producción científica que vaya más allá de las publicaciones convencionales.

Paredes, et al. (2022) describen del mismo modo, que estos están siendo utilizados cada vez más para recopilar y analizar datos sobre la investigación, lo que permite a las instituciones tomar decisiones más acertadas sobre el uso de los recursos y el apoyo a la investigación. Además, estos sistemas están siendo diseñados para ser interoperables con otros sistemas de gestión de la investigación y satisfacer las necesidades a nivel internacional, permitiendo la gestión de información en varios idiomas y el cumplimiento de los requisitos de los diferentes sistemas de financiamiento de la investigación en todo el mundo.

Vázquez, et al. (2017) mencionan que, en los últimos años, los sistemas CRIS han avanzado en la integración de nuevos estándares y protocolos, como CERIF (Common European Research Information Format), ORCID (Open Researcher and Contributor ID), DOI (Digital Object Identifier) y OpenAIRE (Open Access Infrastructure for Research in Europe), lo que ha mejorado su interoperabilidad y capacidad para compartir información entre diferentes sistemas y comunidades científicas.

De acuerdo a Vázquez (2019), el desarrollo de sistemas CRIS en Latinoamérica, todavía se encuentra en una etapa temprana. A pesar de que algunas instituciones cuentan con sistemas de información de investigación, estos no se basan en un enfoque CRIS ni forman parte del modelo de servicio de su repositorio. El desarrollo de un sistema CRIS podría ser beneficioso para la población de la producción científica, y se considera como una oportunidad para las instituciones que buscan aumentar su visibilidad e impacto en este ámbito.

Según Azeroual, et al. (2021), en universidades europeas e internacionales, el sistema CRIS está cobrando mayor relevancia. Las características exclusivas de este sistema se pueden explicar para destacar las diferencias que presenta en comparación con otros sistemas. CRIS tiene la capacidad de integrar los sistemas internos de la universidad, tales como los registros del personal, la administración de estudiantes, las finanzas y los sistemas de gestión de precios, además de una amplia variedad de fuentes de datos externas. De igual manera, se puede incluir la información existente en una única plataforma. El personal universitario utiliza la información disponible en CRIS para diversos propósitos, permitiendo que los investigadores, administradores y delegados ingresen los datos solamente una vez.

Los sistemas de gestión de la investigación CRIS están evolucionando continuamente para satisfacer las necesidades cambiantes de las instituciones de investigación y de los usuarios de estos sistemas. Para Azeroual y Schöpfel (2019) la precisión y fiabilidad de los datos es crucial para las estrategias de seguimiento y evaluación basadas en sistemas CRIS. Si la información sobre personas, organizaciones, proyectos y resultados no es confiable y precisa, entonces el sistema no será útil.

8 Objetivos (generales y específicos aprobados en la propuesta)

8.1 General

Diseñar un Sistema de Información de Investigación Actual (CRIS) para la Universidad de San Carlos de Guatemala que gestione efectivamente los datos sobre los investigadores y

permita la construcción de indicadores de investigación, presentando una alta cohesión y un bajo acoplamiento entre los módulos del sistema, integrando la interoperabilidad para compartir datos de manera efectiva con otros sistemas y herramientas utilizadas en la institución.

8.2 Específicos

- Diseñar una base de datos para almacenar la información sobre los investigadores y los elementos requeridos para la construcción de indicadores; adaptando el estándar de modelo de datos CERIF a las necesidades y características específicas de los procesos de investigación de la USAC.
- Definir un diccionario de datos y metadatos que faciliten la organización de la información gestionada por el sistema, permitiendo la visibilidad del trabajo de investigación a través de indicadores pertinentes.
- Identificar los requerimientos y características mínimas que debe cumplir el ambiente tecnológico en que se implementará el Sistema de Información de Investigación Actual.
- Diseñar los mecanismos de integración con fuentes y sistemas existentes en la universidad, garantizando la interoperabilidad, brindando alta cohesión y un bajo acoplamiento; para mejorar la eficiencia y la colaboración entre los diferentes entes e individuos que participan en actividades de investigación.

9 Materiales y métodos

9.1 Delimitación en tiempo

La investigación se llevó a cabo durante los meses de junio a noviembre del año 2023.

9.2 Delimitación de área geográfica

La investigación se llevó a cabo en las instalaciones del Centro Universitario de Oriente, en el municipio de Chiquimula, Guatemala ubicada sobre el kilómetro 169 carretera CA-10, latitud 14°48'10.2"N y longitud 89°31'50.9"W respectivamente. Dicha

investigación se realizó en colaboración con personal de la Dirección General de Investigación DIGI-USAC ubicada en Ciudad Universitaria, zona 12 Edificio S-11, 3er. nivel. Guatemala, Guatemala latitud 14°34'58.6"N y longitud 90°33'21.3"W

9.3 Enfoque de la investigación

El enfoque mixto de investigación, para Ramírez y Lugo (2020), se define tradicionalmente como la utilización de enfoques cuantitativos y enfoques cualitativos en un mismo estudio. Sin embargo, afirman que un método mixto "no es la suma de lo cuantitativo con lo cualitativo, sino que es la combinación estratégica lo que marca su diseño de aporte a la investigación" (Ramírez y Lugo, 2020).

Se empleó un enfoque mixto, esto permitió realizar, con base en la recolección de datos numéricos y estadísticos, análisis cuantitativos y modelar la estructura de los datos en un CRIS. Por otro lado, los métodos cualitativos ayudaron a entender la experiencia de los usuarios y a identificar factores humanos que podrían afectar el uso del sistema. Además se obtuvo una visión completa y detallada del problema y de los datos relacionados con él.

A través de los métodos cuantitativos se proporcionpo información sobre la frecuencia y la distribución de ciertos comportamientos, mientras que los métodos cualitativos ayudaron a obtener, por medio de entrevistas, información más profunda sobre las percepciones, actitudes y experiencias, permitiendo construir un diseño útil y efectivo para todos los usuarios.

9.4 Método

Ruiz, et al (2022) definen que el método de investigación-acción se originó en el contexto de las ciencias sociales, aplicando observación y análisis a procesos existentes con el objetivo de proponer cambios y transformaciones. También indican que la investigación-acción se ubica en el conjunto de métodos que combinan enfoques cualitativos y cuantitativos. En la revisión sistemática, Ruiz, et al (2022) concluyen que las especialidades de ingeniería con mayor incidencia en la aplicación de investigación-acción

son: la informática, el desarrollo de software y sistemas de información. También encontraron que entre las técnicas más empleadas por dicho método, se encuentran: grupo focal, observación directa, entrevista, encuesta (cuestionario), análisis documental, análisis estadístico, entre otras.

Dada la naturaleza de esta investigación, se utilizó un método de investigación acción. Esto se debe a que, en primera instancia, se realizó un análisis de la situación actual de los sistemas y fuentes de información relacionadas a los procesos de investigación de la USAC. Luego del análisis realizado, se diseñaron cambios a nivel tecnológico que también tendrán impacto en los procesos y dinámica organizacional.

El trabajo de análisis, para cumplir con los primeros dos objetivos específicos, fue orientado por el marco de diseño de bases de datos relacionales.

El diseño de la arquitectura del sistema especifica los mecanismos de integración para lograr la interoperabilidad requerida. Se empleó el método de diseño de arquitectura de software denominado Attribute-Driven Design (ADD). Según Wojcik, et al. (2006), el método ADD propone un acercamiento iterativo al diseño de la arquitectura de un sistema, considerando un módulo a la vez y orientando el trabajo al cumplimiento de requerimientos y atributos de calidad que se deben cumplir.

También, definen una serie de pasos para aplicar el método ADD, que consisten en: confirmación de los requerimientos, elegir el módulo en el que se enfocará la iteración, identificar los atributos que orientarán el diseño (interoperabilidad para esta investigación), elegir los patrones arquitectónicos adecuados, ubicar los elementos específicos en el patrón, definir las interfaces para esos elementos y verificar o refinar los requerimientos.

9.5 Recolección de información

Para seleccionar informantes se tomaron en cuenta los siguientes criterios: experiencia con los sistemas y/o elementos actuales (los que tengan una experiencia significativa con el uso de los elementos a incluir en el sistema CRIS ya sea como

investigadores, administradores o personal técnico), rol en la institución (rol clave), diversidad de disciplinas (informantes que provengan de diferentes áreas para evaluar cómo el sistema CRIS se puede adaptar a las diversas necesidades). Toda esta labor se realizó en primera instancia con personal de DIGI y con investigadores de USAC.

Los parámetros a usar fueron: el número de usuarios (para evaluar el nivel de adopción por parte de la comunidad de investigación) y la interoperabilidad (la cual es importante para evaluar la capacidad que puede llegar a tener el sistema para integrarse con otros sistemas y mejorar la eficiencia y colaboración en las actividades de investigación).

Se utilizó una muestra por conveniencia siendo estos los investigadores de la convocatoria diferenciada 2023 e investigadores del Centro Universitario de Oriente CUNORI. Esto permitió seleccionar personas que estén trabajando actualmente para la universidad y que pudieron dedicar tiempo para responder las preguntas del estudio.

9.6 Técnicas e instrumentos

Se emplearon entrevistas semiestructuradas para obtener información pertinente sobre las necesidades y características específicas que debe cumplir el sistema. Estas entrevistas se realizaron al personal de la DIGI para comprender mejor los procesos y flujos de trabajo e información involucrados en los procesos de investigación científica. Para la aplicación de las entrevistas, se diseñaron cuestionarios que orienten su desarrollo.

El modelado entidad-relación (ER) es una técnica de modelado de datos que permite representar las entidades y relaciones entre ellas en un sistema. El modelado ER ayudó a definir la estructura de los datos del sistema, sus relaciones y sirvió como punto de partida para la creación del diccionario de datos. Como instrumentos, se usó el modelo CERIF como referencia y se creó una tabla para organizar las entradas del diccionario de datos.

Las peticiones a un servidor web, mejor conocidas por el término en inglés "*requests*", son eventos en los que los usuarios de un sistema solicitan un recurso, regularmente una página web, un archivo o el envío de datos a través de un formulario.

Gowandy (2020), expone dos fórmulas para calcular la capacidad de un servidor para atender peticiones a partir de características básicas del hardware: núcleos de la unidad central de procesamiento (CPU) y memoria RAM. A continuación se plantean las fórmulas desde la capacidad de procesamiento y desde la cantidad de memoria. Las variables usadas son: RPS: requests por segundo, NN: número de núcleos, DP: media de la duración del proceso (en segundos), MT: memoria total y MUP: media de la memoria usada por cada proceso.

$$RPS = NN * \frac{1}{DP}$$

Cálculo de requests por segundo a partir del número de núcleos (NN).

$$RPS = \frac{MT}{MUP} * \frac{1}{DP}$$

Cálculo de requests por segundo a partir de la memoria.

Ambas fórmulas pueden ser planteadas en sentido contrario. Partiendo de establecer la capacidad de resolver cierta cantidad de peticiones por segundo, determinar la cantidad de núcleos del procesador y capacidad de la memoria RAM. Para ello, Gowandy (2020) propone las fórmulas descritas a continuación.

$$RPS = \frac{Total\ de\ peticiones}{Tiempo}$$

Determinación de la carga del sistema, en términos de peticiones por segundo, a partir de

$$NN = RPS * DP$$

$$MT = RPS * DP * MUP$$

Cálculo del número de núcleos (NN) a partir de las peticiones por segundo.	*Cálculo de la capacidad de memoria, a partir de peticiones por segundo.*

La duración del proceso DP, será un parámetro propuesto de manera que aporte un equilibrio entre costo y una experiencia favorable para los usuarios.

Para crear la especificación de requisitos de hardware del sistema, se realizó una estimación de la media de peticiones por segundo y la media de memoria requerida para resolver cada solicitud. La estimación de la media de peticiones por segundo se basó en los datos históricos de tráfico y visitas a los sistemas informáticos relacionados con la investigación realizada en USAC. Se aplicaron las fórmulas previamente expuestas. El instrumento correspondiente fue una tabla que presenta los cálculos realizados con los resultados obtenidos.

10 Resultados y discusión

A. Jiménez, en una presentación gestionada por la DIGI como actividad de vinculación, compartió las experiencias de la Universidad Nacional de Costa Rica (UNA) en procesos de implementación de sistemas informáticos relacionados con la investigación (comunicación personal, 25 de septiembre 2023). Especialmente, relató las fases de diseño e implementación del Sistema de Información Académica (SIA). El SIA es un software abarcante, que cubre actividades de docencia, investigación y extensión. Ofrece instrumentos para formular y evaluar proyectos, favorece el control del flujo de procedimientos y permite la generación de indicadores.

En la recolección de requerimientos, se identificaron procesos y requisitos específicos y particulares, lo cual daba lugar a considerar el diseño de un sistema informático a la medida. A su vez, diferentes software CRIS *open source* fueron evaluados, destacando DSpace CRIS y VIVO. Además, A. Jiménez, refirió que en la UNA, existe una serie de sistemas que interactúan con el SIA, incluidos una instancia de VIVO y un portal de revistas (comunicación personal, 25 de septiembre 2023).

Considerando lo anterior, se plantea el diseño de dos sistemas informáticos vinculados: un Sistema de Gestión de la Información de Investigación (RIMS) y un Sistema de Información de Investigación Actual (CRIS). El primero se encargará de la administración y mantenimiento de la información, mientras que el segundo ofrecerá principalmente visibilidad. En conjunto, cumplirán con la característica esencial de interoperabilidad.

A continuación, se describen los resultados obtenidos para cada objetivo específico.

10.1 Resultados objetivo 1

Para alcanzar el primer objetivo, resultó necesario definir los requerimientos de datos y procesos. Con el propósito de comprender y delimitar las necesidades o requerimientos del sistema, se elaboró un documento de especificación de requerimientos. El documento pretende proveer una base sólida y clara para el diseño, desarrollo e implementación de un Sistema de Información de Investigación Actual (CRIS). Dicho documento surgió luego del análisis de la información proveída por investigadores y personeros de la DIGI. Dicha información se obtuvo a través de cuestionarios y entrevistas.

Se utilizó muestreo por conveniencia. En primer lugar se identificaron cuatro perfiles de individuos de los que se necesitaba obtener información, siendo éstos: personal administrativo de la DIGI (directora general y coordinadores de programas de investigación), encargado del área de informática y cómputo, investigadores y la persona que ha trabajado indicadores de investigación en el pasado.

Para aplicar el muestreo por conveniencia, se establecieron criterios de inclusión y exclusión correspondientes a los perfiles identificados, tal como se describe en la tabla 1.

Tabla 1. Criterios de inclusión y exclusión para el personal administrativo y de informática de la Dirección General de Investigación.

Criterios de inclusión	
No.	**Criterio**
1	Pertenecer al grupo de profesionales que desempeñan roles administrativos de la DIGI.
2	Cumplir con el rol de encargado de la sección de informática de DIGI.
3	Tener como atribución la gestión de asuntos relacionados con la propiedad intelectual.
4	Estar activos en sus funciones durante los meses de julio y agosto de 2023.
5	Tener conocimiento previo del proyecto de investigación.
Criterios de exclusión	
No.	**Criterio**
1	No contestar el cuestionario correspondiente.
2	No pertenecer al grupo de profesionales que desempeñan roles administrativos de la DIGI.

De los criterios de inclusión y exclusión para el grupo de personal administrativo y de informática, fueron seleccionados 9 individuos. El desglose de los mismos se detalla en la tabla 2.

Tabla 2. Cantidad de individuos por cargo que fueron incluidos en la muestra del personal administrativo y de informática.

No.	Cargo en la DIGI	Cantidad de individuos
1	Directora General de Investigación	1
2	Coordinador General de Programas	1
3	Coordinador de Programa	5
4	Encargado del Área de Informática y Cómputo	1
5	Encargada del Centro de apoyo a la tecnología e innovación CATI-DIGI	1
	Total	**9**

Se abordaron 28 investigadores de la Universidad de San Carlos, a los cuales se les preguntó detalles sobre los procesos y datos relacionados con la administración de proyectos de investigación. Los criterios de selección para los investigadores son expuestos en la tabla 3.

Tabla 3. Criterios de inclusión y exclusión para los investigadores.

Criterios de inclusión	
No.	**Criterio**
1	Contar con dos o más años de experiencia como investigador.
2	Estar afiliado por lo menos a una unidad avaladora de proyectos de investigación de la USAC.
3	Ser coordinador de algún proyecto de investigación de la Convocatoria Diferenciada 2023 o figurar en la base de datos de investigadores del Centro Universitario de Oriente.

Criterios de exclusión	
No.	**Criterio**
1	No contestar el cuestionario respectivo.
2	Pertenecer al grupo de personal administrativo o de informática de la DIGI y haber respondido el cuestionario correspondiente a uno de esos grupos.
3	Tener menos de dos años de experiencia como investigador y no contar con artículos publicados en revistas indexadas.
4	No contar con identificador ORCID y no estar dispuesto a crear uno.

El hecho de no contar con identificador ORCID y no tener disponibilidad de crearlo se interpreta como una actitud de resistencia a la incorporación de tecnologías de información en el ejercicio de actividades relacionadas al rol de investigador. Por lo tanto, se constituyó como un criterio de exclusión.

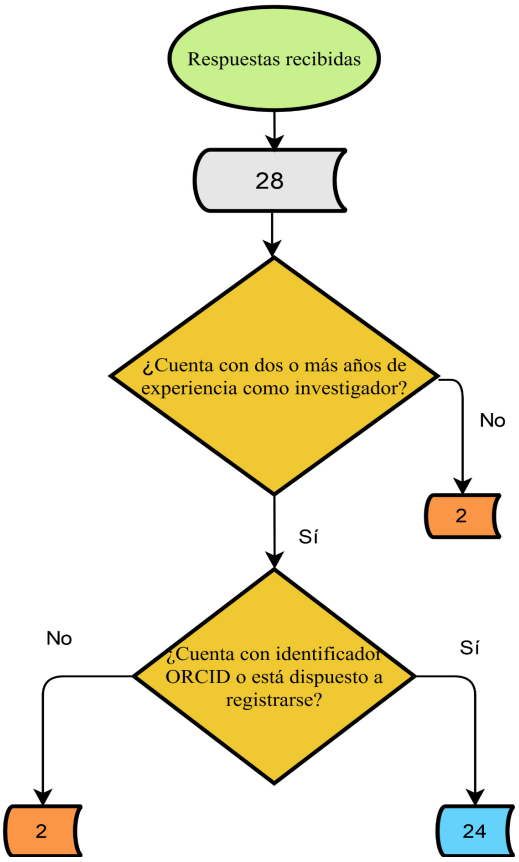
Se contó con el apoyo de la Coordinación General de Programas de Investigación Digi para distribuir el formulario a los coordinadores de investigaciones con financiamiento de DIGI y que se encontraban en desarrollo durante julio y agosto de 2023. Además, el Instituto de Investigación del Centro Universitario de Oriente colaboró con la distribución de cuestionarios con los investigadores afiliados al mismo. De forma general, se logró obtener respuestas de coordinadores de investigación pertenecientes a 12 unidades académicas. La tabla 4 lista las unidades académicas a las que se encontraban afiliados los coordinadores de investigación incluidos en la muestra.

Tabla 4. Unidades académicas de los investigadores incluidos en la muestra.

No.	Unidad académica
1	Centro de Estudios Urbanos y Regionales -CEUR-
2	Centro de Estudios de las Culturas en Guatemala -CECEG-
3	Instituto de Investigaciones Humanísticas -IIH-, Facultad de Humanidades
4	Instituto de Investigación de Centro Universitario de Oriente -IIC-, Chiquimula -CUNORI-
5	Centro de Estudios Urbanos y Regionales -CEUR-
6	Departamento de Investigaciones Generales del CUNSOL -DICUNSOL-, Centro Universitario de Sololá -CUNSOL-
7	Centro de Investigaciones en Psicología -CIEPS-, Escuela de Ciencias Psicológicas
8	Centro de Estudios en Comunicología -CICC-, Escuela de Ciencias de la Comunicación
9	Unidad de Investigación de la Escuela de Formación de Profesores de Enseñanza Media -EFPEM-
10	Dirección de Investigación de la Facultad de Arquitectura -DIFA-, Facultad de Arquitectura
11	Dirección General del Sistema de Investigación -DICUNOC-, Centro Universitario de Occidente, Quetzaltenango -CUNOC-
12	Coordinación de Investigación del CUNSAC -CDI-, Centro Universitario de Sacatepéquez -CUNSAC-

Se recibieron 28 respuestas correspondientes a investigadores que cumplían con los criterios de inclusión 2 y 3. Luego, a partir de los datos proveídos, se aplicaron los criterios de selección restantes. El diagrama de flujo de la figura 1 ilustra los criterios aplicados, dando como resultado una muestra de 24 investigadores.

Figura 1. Diagrama de flujo de la aplicación de criterios de inclusión y exclusión para investigadores.



Fuente: Elaboración propia.

Para conocer sobre los indicadores que ha utilizado la DIGI para evaluar su trabajo, se diseñó un cuestionario específico para los individuos que han trabajado en su formulación y cálculo. La tabla 5 muestra los criterios de selección usados para el personal con experiencia en la generación de indicadores.

Tabla 5. Criterios de inclusión y exclusión para el personal con experiencia en la generación de indicadores de investigación para la Dirección General de Investigación.

Criterios de inclusión	
No.	**Criterio**
1	Formar parte de las autoras de los documentos: Indicadores del sistema de investigación de la Universidad de San Carlos de Guatemala, 2015 - 2018 e Indicadores del sistema de investigación de la Universidad de San Carlos de Guatemala, 2019 – 2021.
2	Estar activas en sus funciones durante los meses de julio y agosto de 2023.
Criterios de exclusión	
No.	**Criterio**
1	No contestar el cuestionario correspondiente.

Se identificó que la persona encargada de Investigación, Desarrollo e Innovación ha sido designada en años anteriores para formular, generar y documentar los indicadores de investigación.

Se emplearon cuestionarios semi estructurados en formato digital, haciendo uso de Google Forms. El diseño de los cuestionarios fue específico para los segmentos de individuos según su perfil, de la manera en que se describió en el inciso 2 del presente informe. La tabla 6 detalla los cuestionarios utilizados con la cantidad de ítems correspondiente.

Tabla 6. Cuestionarios aplicados a los individuos de la muestra.

No.	Cuestionario	Cantidad de ítems
1	Cuestionario dirigido a investigadores	19
2	Cuestionario orientado a las regulaciones y normativas (propiedad intelectual)	4
3	Cuestionario dirigido al personal administrativo de la DIGI	18
4	Cuestionario orientado a la generación de indicadores	18

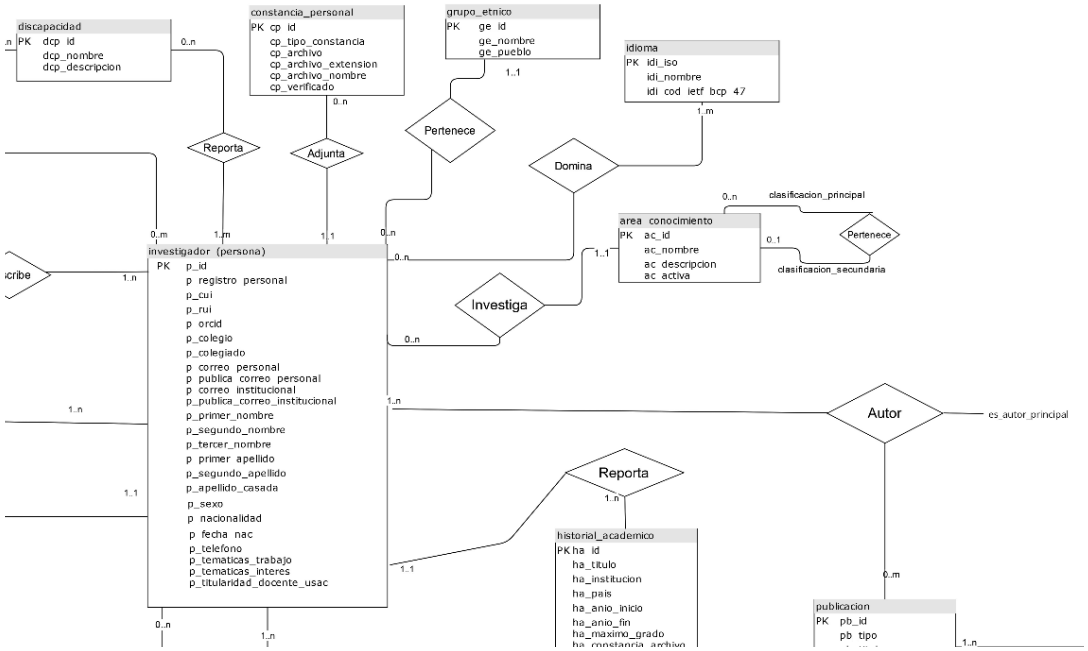
Para confirmar la comprensión de las respuestas proveídas, se compartió la especificación de requerimientos con los coordinadores de programas de la DIGI. Se aplicaron las observaciones recibidas y se desarrolló una reunión virtual para confirmar la retroalimentación recibida y la forma en que se adaptó al documento.

El documento de especificación de requerimientos, las respuestas a los cuestionarios y análisis consolidado de cuestionarios a investigadores y al personal administrativo de la DIGI, se encuentran en la carpeta digital que se entregó con el presente informe.

Se diseñó un diagrama entidad-relación (ER), que se plantea como modelo lógico de la base de datos. Se incluyeron entidades como: investigador, dependencia, proyecto (para poder calcular indicadores) y publicación (para completar la información de los investigadores). Las entidades referidas corresponden a las entidades CERIF: person, organization unit, project, publication (research output) respectivamente. El diagrama ER también podrá ser utilizado en la definición de entidades y relaciones para DSpace CRIS.

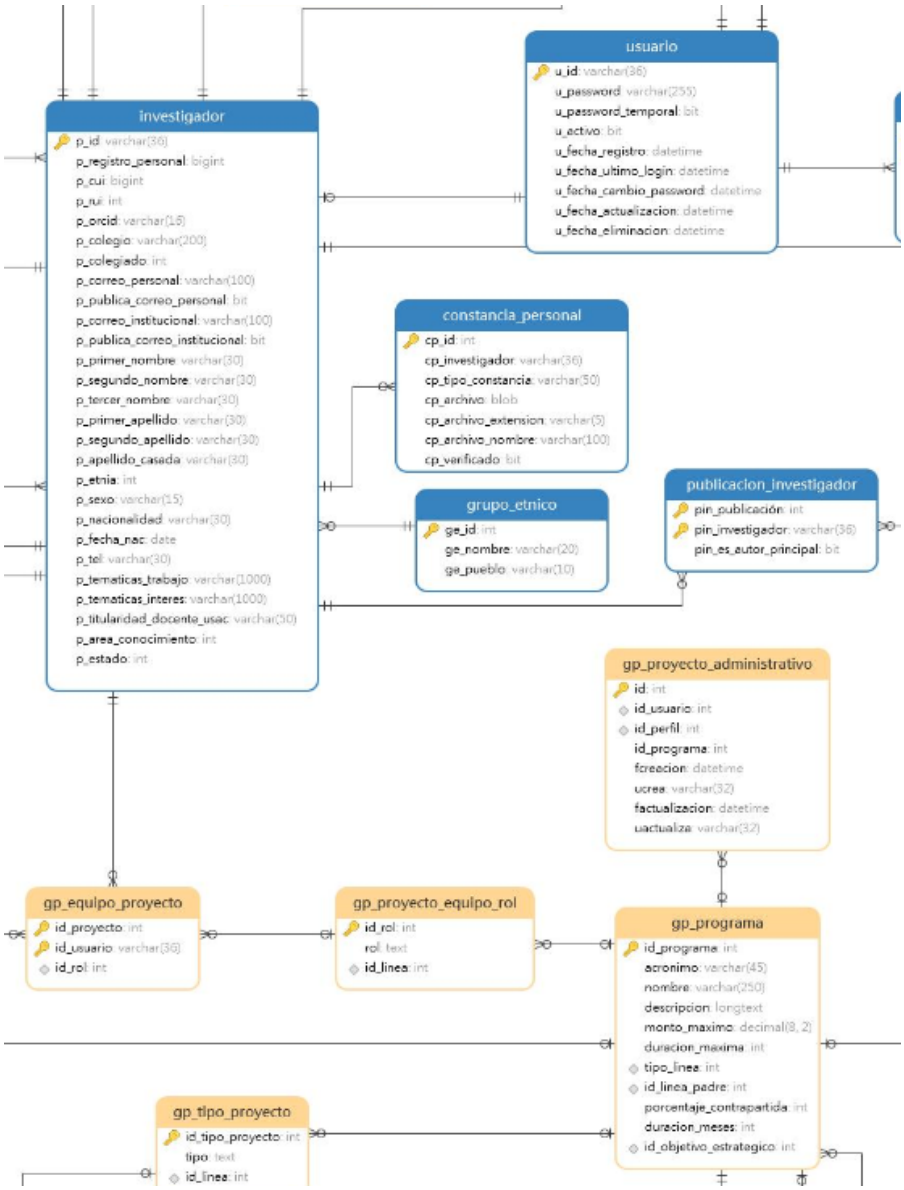
La figura 2 muestra una parte del diagrama ER descrito. En la carpeta digital presentada con el informe, se adjunta el diagrama completo con el nombre "ModeloER.jpg".

Figura 2. Vista parcial del diagrama entidad-relación.



Además, se diseñaron tres vistas del diagrama relacional. Dicho diagrama provee detalles específicos de implementación, tales como tipos de datos y longitudes. La primera vista del diseño relacional, se puede ubicar en el contenido de la carpeta digital, en el archivo "Diagrama-relacional-RIMS.jpg". Dicha vista contiene las tablas necesarias para la gestión de la información que se concentrará en el RIMS. Se distinguen dos colores de tablas, las azules corresponden a las diseñadas para el módulo de investigadores y las de color naranja fueron diseñadas por el equipo encargado del módulo de proyectos. La figura 3 muestra una parte del diagrama en mención.

Figura 3. Vista parcial del diagrama relacional que incluye las tablas que se usarán en el RIMS.



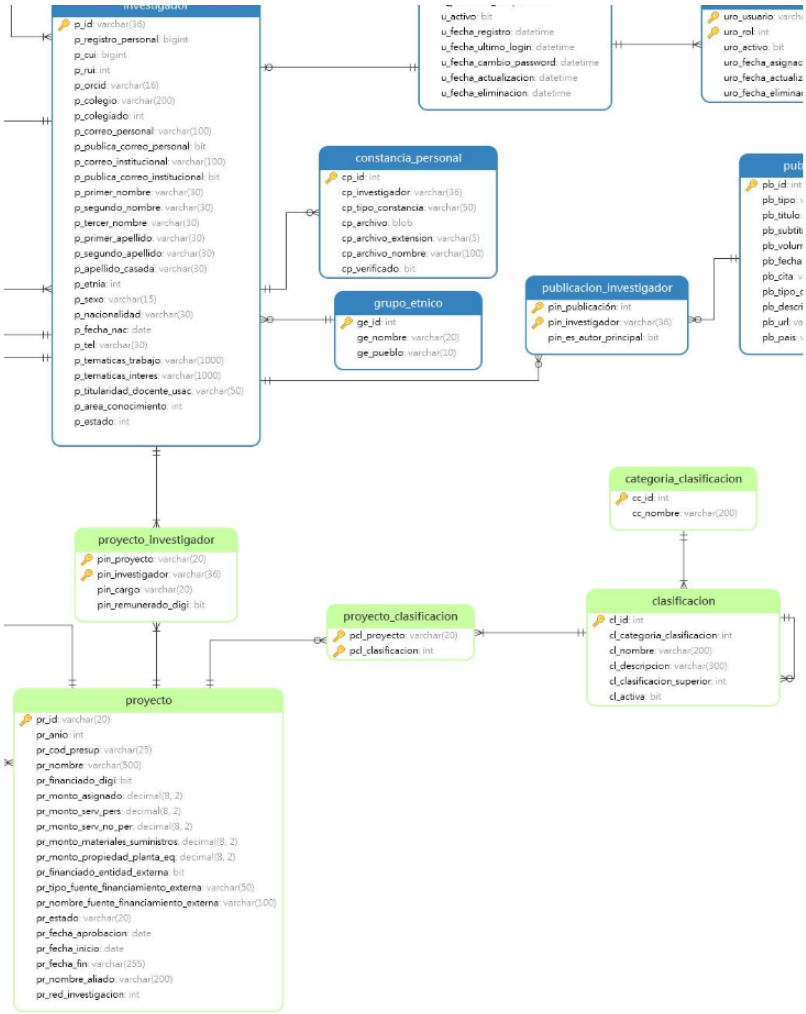
Nota. Las tablas delimitadas con color azul corresponden al módulo de investigadores.

La segunda vista del diseño relacional, se puede ubicar dentro de la carpeta digital, en el archivo "Diagrama-relacional-RIMS-Investigadores.jpg". Dicha vista muestra solamente las tablas que corresponden al módulo de investigadores en el diagrama anterior.

Por último, se puede ubicar a la tercera vista del diseño relacional, en el archivo "Diagrama-relacional-Indicadores.jpg" de la carpeta digital. Dicha vista muestra las tablas que corresponden al módulo de investigadores, usando color azul; y de color verde a las tablas que apoyarán el cálculo de indicadores. Dichas tablas equivalen a las entidades que propone el diagrama ER en relación a los proyectos. La figura 4 presenta un acercamiento a un sector del diagrama. Las tablas delimitadas con color verde pueden considerarse como un resumen de los datos de proyectos que se gestionarán con las tablas marcadas de color naranja en la primera vista del diagrama relacional.

Informe final de proyecto de investigación. Año 2023

Figura 4. Vista parcial del diagrama relacional que incluye las tablas que apoyarán el cálculo de indicadores.



Nota. Las tablas delimitadas con color verde, se contemplan para completar la información requerida en el módulo de investigadores.

Las columnas de la tabla "proyecto", fueron consideradas a partir del documento de parámetros a requerir en investigaciones financiadas por la DIGI o por terceros; según Sosa (2023). De dichas columnas, las relacionadas al aspecto financiero, se fundamentan en las clasificaciones presupuestarias reglamentadas para el sector público de Guatemala (Ministerio de Finanzas Públicas, 2018).

La tabla "colaborador" se agrega para las publicaciones que se carguen de ORCID. Cuando se manejen publicaciones en un módulo propio, solamente servirá para colaboradores no registrados en la tabla de investigador.

El monto ejecutado se calcula con la sumatoria de los cuatro grupos: servicios personales (pr_monto_serv_pers), servicios no personales (pr_monto_serv_no_per), materiales y suministros (pr_monto_materiales_suministros) y propiedad, planta, equipo e intangibles (pr_monto_propiedad_planta_eq).

Las clasificaciones por tipo de investigación, área del conocimiento, programa DIGI, tema regional y las que surjan en el futuro se manejarán con las tablas "categoria_clasificación" y "clasificación".

10.2 Resultados objetivo 2

El diccionario de datos se presenta en formato tabular. Las tablas de la base de datos se describen con detalle explicando el tipo, longitud, visibilidad, naturaleza y propósito de cada columna o atributo. Para conocer la estructura y organización de las columnas del diccionario de datos, se plantea una explicación del contenido del diccionario en la tabla 7.

Tabla 7. Descripción de las columnas del diccionario de metadatos.

Columna	Descripción
Nombre del elemento	Proporciona el nombre de la columna o atributo de la tabla que se está especificando.
Descripción	Ofrece una descripción clara y concisa de la columna, explicando su propósito y significado en relación con los datos.

Columna	Descripción
Tipo de dato	Indica si el dato es de tipo alfanumérico, número entero, número decimal, fecha, bit, entre otros.
Longitud	Especifica la longitud máxima permitida para el elemento de datos, especialmente relevante para campos de texto o caracteres. En algunos casos, se consignan dos números separados por coma. Lo anterior se usa para definir la longitud y precisión, en ese orden, para datos de tipo decimal.
Requerido	Especifica si el elemento de datos es obligatorio o es opcional, lo cual ayuda a comprender su importancia y necesidad al capturar datos.
Origen	Indica la fuente o el origen del dato, es decir, dónde se obtiene o se genera la información.
Criterio de Edición	Describe si el dato puede ser modificado después de ser almacenado la primera vez.
Público	Indica la visibilidad del dato al público en general en el sistema (sí/no), al tener el valor sí indica que el dato podrá ser visible de manera pública por todos los usuarios.

El documento digital que contiene al diccionario de datos, se adjunta en la carpeta presentada junto al informe final. El documento distribuye las tablas en las hojas que lo componen. En la figura 5 se muestra una vista parcial de la forma en que el diccionario de datos presenta a los atributos que componen la tabla que almacenará las áreas de conocimiento.

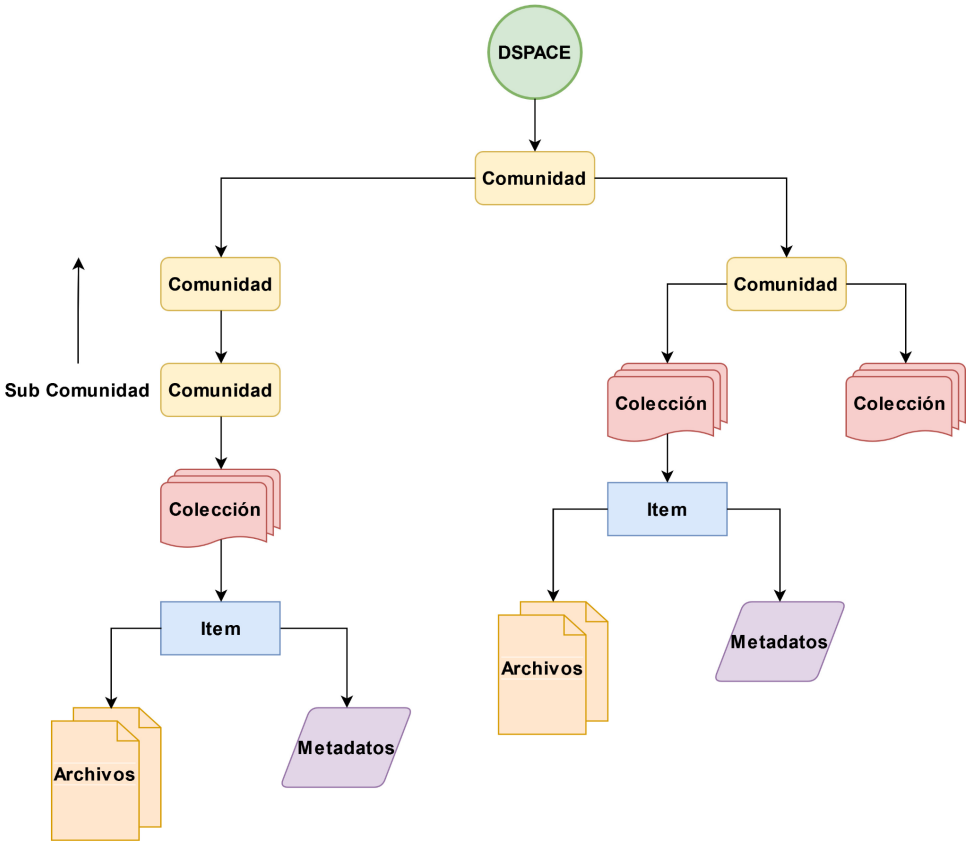
Figura 5. Vista parcial del diccionario de datos.

Diccionario de datos							
Nombre de la tabla:	Area_conocimiento				**Fecha de creación:**	Agosto de 2023	
Descripción:	Tabla que registra un catálogo de valores de referencia para las áreas de conocimiento según Frascati (2015).						
Nombre del elemento	**Definición**	**Tipo de dato**	**Longitud**	**Requerido**	**Origen**	**Criterio de edición**	**Público**
ac_id	Identificador del registro del área de conocimiento.	Integer	3	Sí	Generado automaticamente	No editable	No
ac_nombre	Nombre del área de conocimiento, ejemplo: Ciencias Naturales, Ciencias Sociales, etc.	Varchar	90	Sí	Formulario	Editable	No
ac_descripcion	Descripción del área de conocimiento.	Varchar	150	Opcional	Formulario	Editable	No
ac_activa	Valor booleano que servirá para realizar eliminación lógica. Indicará si el registro está activo.	Bit	1	Sí	Formulario	Editable	No
ac_clasificación_principal	Idenfiticador de otro registro de la misma tabla que corresponda al área de clasificación principal. Ejemplo: Ciencias Químicas -> Ciencias Naturales.	Integer	3	Opcional	Formulario	Editable	No
Relaciones:	Tabla "area_conocimiento" relación recursiva.				**Campos clave:**	ac_id	

Los metadatos son elementos concisos que describen o contextualizan a un dato. Según Donohue (2023), DSpace CRIS emplea diversos metadatos contemplados por defecto y también ofrece la posibilidad de añadir nuevos campos. Para cumplir con el segundo objetivo del proyecto, se diseñó un diccionario de metadatos. Dicho diccionario, es una referencia para poder utilizar apropiadamente los metadatos contemplados en DSpace CRIS.

De Albuquerque, et al. (2021), en una presentación creada para la Universidad Nacional de la Patagonia San Juan Bosco (UNPSJB), indican que los contenidos de DSpace se organizan en un modelo jerárquico. En el nivel superior, los elementos se agrupan en comunidades, las cuales pueden representar instituciones, facultades o carreras. Las comunidades pueden estructurarse en subcomunidades para reflejar las relaciones de dependencia entre las entidades que representan. Cada comunidad contiene colecciones de ítems. Los ítems son recursos básicos que se componen de archivos y metadatos que los describen. La figura 6 presenta el modelo jerárquico de organización de contenidos en cuestión.

Figura 6. Modelo de contenidos de DSpace.



Nota. Adaptada de "Curso de capacitación informática para UNPSJB" por P. de Albuquerque, S. Tettamanti y A. Lira, 2021 (http://sedici.unlp.edu.ar/handle/10915/126108). CC BY-NC-SA 4.0

Das y Sutradhar (2018) establecen que DSpace emplea metadatos para importar y exportar metadatos usando el protocolo para la recolección de metadatos de la Iniciativa Open Archives, OAI-PMH por sus siglas en inglés. Ese intercambio de los metadatos es clave para garantizar la interoperabilidad. También añaden que el uso de esquemas estandarizados de metadatos es un aspecto esencial que se debe considerar para dicho fin. Pero además, se subraya que DSpace es independiente del esquema de metadatos, lo cual implica que el software es flexible al recomendar estándares y también permitir la incorporación de esquemas propios o de terceros.

De Albuquerque, et al. (2021) mencionan que es posible enriquecer los campos de los metadatos aplicando una extensión o calificación a un esquema básico. También, cabe resaltar, que "un repositorio puede utilizar un perfil de metadatos basado en esquema propio, estándar y/o mixto" (de Albuquerque, et al., 2021). Como ejemplo, se puede referir al metadato dc.date, del estándar Dublin Core. Es posible extender el campo para darle mayor riqueza descriptiva. Para lograrlo, se concatena un punto seguido de un calificador, que estaría dando contexto o especificidad al elemento base. De manera que, dc.date.available describiría la fecha a partir de la cual un recurso está disponible. Y en la misma lógica, dc.date.issued indicaría la fecha de publicación del recurso.

Wood (2023) indica que DSpace ofrece una amplia gama de campos de metadatos y mecanismos para aumentar el catálogo. Donohue (2018) establece que, Dublin Core, es el esquema estándar predeterminado que usa DSpace. Además, 4Science (2023), en la rama específica para DSpace CRIS de su repositorio de código, define los registros de campos de metadatos que incorpora en su configuración básica inicial. Tales registros se encuentran agrupados según los esquemas a los que pertenecen, dentro del directorio /dspace/config/ registries. La tabla 8 detalla los registros de campos de metadatos utilizados por DSpace CRIS, que fueron considerados en el diccionario de metadatos elaborado.

Tabla 8. Esquemas empleados en el diccionario de metadatos.

Nombre del esquema	Prefijo de los campos	Archivo
Dublin Core (DCMI Metadata Terms)	dc	dublin-core-types.xml
Schema.org Person	person	schema-person-types.xml
OpenAIRE Guidelines for CRIS	oairecerif	openaire-cerif-types.xml
DSpace Cris researcher's page Types	crisrp	crisrp-types.xml

Nombre del esquema	Prefijo de los campos	Archivo
OpenAIRE Guidelines for CRIS	oairecerif	openaire-cerif-types.xml
DSpace Cris OrgUnit Types	crisou	crisou-types.xml
Schema.org Organization	organization	schema-organization-types.xml
OpenAIRE4 Guidelines for Literature Repositories	datacite	openaire4-types.xml
Schema.org Project	project	schema-project-types.xml
DSpace Internal types	dspace	dspace-types.xml

Nota. Datos adaptados de "DSpace (dspace-cris-7 branch)", por 4Science, 2023 (https://github.com/4Science/DSpace/tree/dspace-cris-7). La tabla describe algunos esquemas contenidos en el repositorio de DSpace, ubicados en el directorio /dspace/config/registries.

Dublin Core Metadata Initiative es, según Das y Sutradhar (2018), un pequeño conjunto de términos o campos de metadatos que pueden usarse para describir recursos. Iliadis, et al. (2022) relatan que el proyecto Schema.org surgió en el año 2011 de una colaboración entre Google, Microsoft, Yahoo y Yandex, para crear un esquema estructurado y compartido de metadatos para ser usados en los motores de búsqueda. Otros esquemas mencionados en la tabla 8 fueron creados por y para DSpace CRIS específicamente. Tal es el caso de los esquemas: DSpace Cris OrgUnit Types y DSpace Internal Types.

El diccionario de metadatos diseñado se presenta en un formato tabular organizado en un archivo de hoja de cálculo. Distribuye los metadatos bajo el criterio de entidades basadas en CERIF: investigador (persona), dependencia (organización), proyecto y publicación. El documento digital entregable considerado como diccionario de metadatos, se adjunta en la carpeta presentada junto al informe. La figura 7 presenta una vista parcial

de la forma en que el diccionario de metadatos presenta a los elementos que describen a una dependencia (organización). La tabla 8 define el contenido que presenta el diccionario de metadatos en cada columna.

Figura 7. Vista parcial del diccionario de metadatos.

Diccionario de metadatos						
Nombre del archivo:	Metadatos de la dependencia.				**Fecha de creación:**	Octubre, 2023
Descripción:	Se consideran los metadatos relacionados con la entidad dependencia (Organization Unit en CERIF), que utiliza DSpace CRIS. En el futuro, podrán agregarse los que se consideren necesarios.					
					Origen (RIMS)	
Nombre del metadato	**Definición**	**Campo en DSpace**	**Estándar**	**Tipo de metadato**	**Tabla**	**Columna**
Nombre	Nombre de la unidad académica o dependencia.	dc.title	Dublin Core (DCMI Metadata Terms)	Descriptivo	Dependencia	dep_nombre
Acrónimo	Acrónimo de la dependencia.	oairecerif.acronym	DSpace Cris OrgUnit Types	Descriptivo	Dependencia	dep_acronimo
Dirección	Dirección física de las instalaciones centrales.	organization.address. addressLocality	Schema.org	Descriptivo	Dependencia	dep_dirección

Nota. Datos en las columnas 3 y 4 adaptados de "DSpace (dspace-cris-7 branch)", por 4Science, 2023 (https://github.com/4Science/DSpace/tree/dspace-cris-7).

Tabla 9. Descripción de las columnas del diccionario de metadatos.

Columna	Descripción
Nombre del metadato	Proporciona un nombre significativo para identificar al metadato.
Definición	Ofrece una descripción concisa del metadato.
Campo en DSpace CRIS	Nombre que usa DSpace CRIS para el campo del metadato. Es un identificador compuesto por el prefijo del esquema, la parte básica del campo y calificadores, separados por punto.
Estándar	Nombre del esquema o estándar al que pertenece el campo. Guarda relación con el prefijo del elemento.
Tipo de metadato	Tipo del metadato según el ámbito semántico que describe. Podría ser: administrativo, descriptivo o estructural.
Origen (RIMS) Tabla	Tabla del modelo relacional del RIMS en la que se encontrará la columna que proveerá el valor para el campo de metadato en DSpace CRIS. Esto responde a la lógica de que el RIMS permitirá reunir y administrar datos que luego transmitirá al CRIS para dar visibilidad a los mismos.
Origen (RIMS) Columna	Columna específica de la tabla referenciada en la columna correspondiente, en la que el RIMS almacena el valor que se asignará al campo del metadato en cuestión.

10.3 Resultados objetivo 3

Con el propósito de conocer el comportamiento que registran los sistemas informáticos relacionados con la investigación realizada en USAC, se entabló comunicación con el área de informática y cómputo de la DIGI. El encargado del área identificó cuatro elementos informáticos que mantienen: convocatorias para proyectos de investigación, Registro Único de Investigadores, biblioteca virtual y el portal oficial. Las convocatorias para proyectos de investigación y el Registro Único de Investigadores utilizan la herramienta online para creación de formularios denominada LimeSurvey. Los resultados de estos formularios se almacenan y procesan en el software MS Excel.

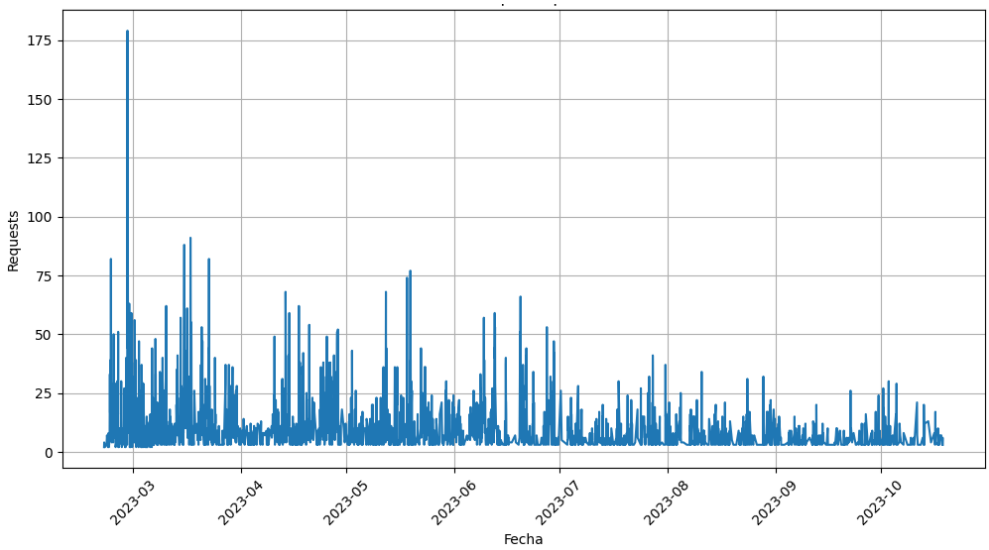
El área de informática y cómputo de la DIGI proveyó un reporte de accesos por fecha y hora al portal web. El reporte se obtuvo en los formatos digitales PDF y CSV. El documento CSV se utilizó como fuente de datos históricos de tráfico y visitas a sistemas informáticos que guardan relación con el CRIS en proceso de diseño. El reporte contenía 5000 registros estructurados en siete columnas: fecha y hora, ruta de página, número de vistas, número de usuarios, vistas por usuario, tiempo de interacción y número de eventos. Algunas de las páginas que presentaban accesos en el reporte fueron: convocatoria 2023, resultados de convocatoria diferenciada, programas de investigación, formatos 2023, fondo de investigación, revistas científicas, manuales administrativos, entre otras.

El objetivo del análisis del reporte de accesos por fecha y hora al portal web, era determinar un valor de peticiones por segundo en un contexto de sistemas informáticos de investigación de la universidad. A partir de ese valor, se aplicaron unas fórmulas como posible estrategia para estimar las características que se propondrán para el servidor del sistema informático diseñado.

Considerando el objetivo del análisis, planteado en el párrafo anterior, se consideró únicamente la fecha, hora y el total de vistas correspondientes. Se asumió la equivalencia entre vista o visita a una página del portal con un *request* o solicitud. La equivalencia reside

en el hecho de que para visualizar o visitar una página web, el navegador hace una solicitud o *request* al servidor que aloja el portal. Al aplicar el resumen que desestimó la ruta de página y se condensaron los registros por fecha y hora, consecuentemente, el reporte se redujo a 2306 entradas. Los datos corresponden al período comprendido del 27 de febrero al 18 de octubre de 2023. La figura 8 muestra la evolución diaria de la demanda (en *requests*) al portal de la DIGI en el período mencionado.

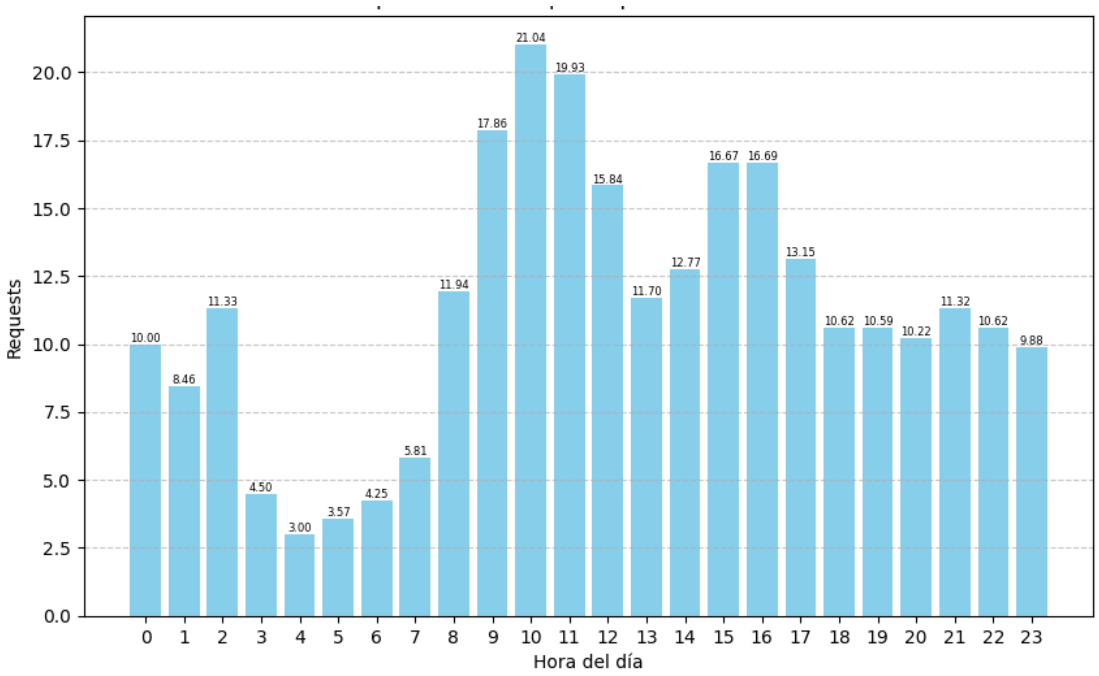
Figura 8. Solicitudes atendidas por el portal de la Web DIGI.



Fuente: elaboración propia, con datos proveídos por informática DIGI.

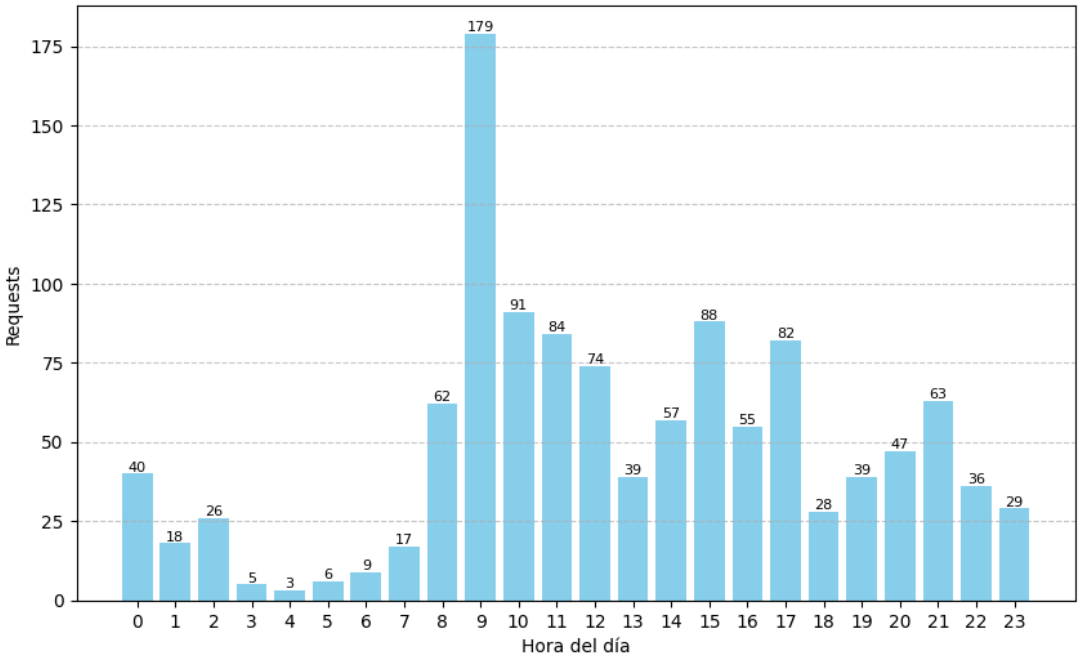
En la gráfica de la figura 8, se puede observar que el período que reporta mayor cantidad de solicitudes diarias abarca los meses de febrero a mayo. De manera que se concentró el análisis en ese rango de tiempo. Se procedió a encontrar el promedio y el valor máximo de *requests* por cada hora del día. Las figuras 9 y 10 muestran la media y el valor máximo de solicitudes registradas para cada hora del día. Para considerar el escenario de mayor demanda, se utilizó el valor máximo de requests en los cálculos detallados en los párrafos siguientes.

Figura 9. Valor promedio de request para cada hora del día.



Fuente: elaboración propia, con datos proveídos por informática DIGI.

Figura 10. Máximo valor de request para cada hora del día.



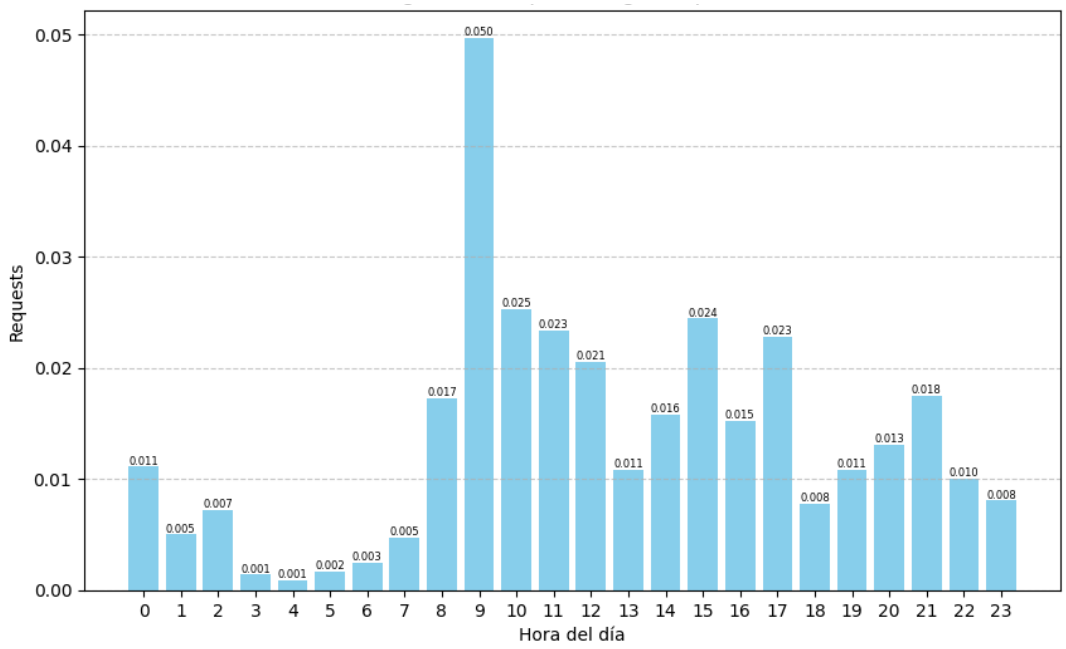
Fuente: elaboración propia, con datos proveídos por informática DIGI.

Se empleó la fórmula propuesta por Gowandy (2020), para el cálculo de *requests* por segundo RPS.

$$RPS = \frac{Total\ de\ peticiones}{Tiempo}$$

Dado que los valores de *requests* están dados en el marco de una hora, se aplicó la conversión a peticiones por segundo. La gráfica de la figura 4 muestra las medidas de requests por segundo RPS a partir del valor máximo de requests por cada hora del día. El apéndice 2 muestra los datos que respaldan la gráfica de la figura 11. El registro con mayor cantidad de peticiones en una hora, corresponde a las 09:00 hrs. del día 27 de febrero. Se utilizará la medida aproximada de 0.05 requests/segundo en las fórmulas propuestas para determinar el número de núcleos de procesador y memoria total requeridos.

Figura 11. Máximo registro de requests/segundo para cada hora.



Fuente: elaboración propia, con datos proveídos por informática DIGI.

Gowandy (2020), expone dos fórmulas que relacionan la capacidad de un servidor para atender peticiones a partir de características básicas del hardware: núcleos de la unidad central de procesamiento (CPU) y memoria RAM. Las variables usadas son: RPS: requests por segundo, NN: número de núcleos, DP: media de la duración del proceso (en segundos),

$$NN = RPS * DP$$

Cálculo del número de núcleos (NN) a partir de las peticiones por segundo.

$$MT = RPS * DP * MUP$$

Cálculo de la capacidad de memoria, a partir de peticiones por segundo.

MT: memoria total y MUP: media de la memoria usada por cada proceso. Las fórmulas propuestas por Gowandy (2020) son:La duración del proceso DP, se fija a 0.2 segundos ya que es un tiempo de respuesta recomendable (Lokhtenko, 2022). Para lograr ese tiempo de respuesta, proyectando escenarios en los que el servidor reciba múltiples solicitudes de

manera simultánea, se asigna un valor de 500Mb para la memoria usada por cada proceso MUP.

Tabla 10. Resumen del análisis de la demanda del portal DIGI.

Parámetro	Valor
Requests por segundo -RPS-	0.05
Duración del proceso -DP-	0.2 segundos
Memoria usada por el proceso -MUP-	500Mb
Número de núcleos -NN-	0.01 ≈ 1 núcleo
Memoria total -MT-	5 Mb

Fuente: Elaboración propia.

Los procesos de la DIGI que giran en torno al portal, y en general en todos los sistemas de información que administra actualmente, representan un pequeño porcentaje de las actividades de gestión de los proyectos y la información que manejan. Principalmente reflejan consultas y descargas de información centradas en las convocatorias. Por tal razón, la medida de *requests* por segundo que se calculó, a partir de los datos, es casi despreciable. De manera que los resultados no pueden ser utilizados para la propuesta de ambiente tecnológico.

Registros con valores máximos de *requests* para cada hora. Las fechas no son consecutivas ya que se consideraron las entradas con mayor cantidad de peticiones de cada hora. De manera que, se puede considerar el ejemplo de que, para las 8:00 horas, el día 27 de febrero fue el que reportó más peticiones. Para efectos prácticos, la columna de fecha, solamente sirve de referencia.

Por los hallazgos indicados y derivado que los resultados antes descritos no pueden ser utilizados como base para la propuesta de ambiente tecnológico, se procedió a otras fuentes: siendo una de ellas la documentación de DSpace y la otra la experimentación propia usando contenedores

Requerimientos de Hardware para el sistema CRIS (de acuerdo a documentación de DSpace)

De acuerdo con la documentación de DSpace (*Performance Tuning DSpace - DSpace 7.X Documentation - LYRASIS Wiki*, s. f.) se puede mencionar los siguiente requerimientos mínimos de hardware para el sistema CRIS.

Tabla 11. Requerimientos mínimos DSPACE.

Recurso	Capacidad Mínima
Frontend(UI)/Node.js	2 GB de memoria
Backend(REST API) / JVM / Tomcat	1 GB de memoria
PostgresSQL Database	512 MB de memoria
Solr	512 MB de memoria

Fuente: Elaboración propia.

Entre los requerimientos previamente mencionados, se requiere una capacidad de 4 núcleos de procesador para los servicios necesarios del sistema, sin embargo es una capacidad que puede ser ampliada según como el sistema lo requiera.

Durante la investigación e instalación del sistema Dspace en nuestros equipos, obtuvimos la experiencia de trabajar con estos servicios activados en diferentes contenedores, lo cual nos permitió llevar a cabo un análisis del rendimiento y consumo de recursos que son necesarios para levantar el sistema antes mencionado.

De acuerdo con estas pruebas se desarrolló la guía prueba de recursos DSPACE CRIS, en la cual podemos encontrar los siguientes requerimientos mínimos y recomendados para el sistema.

Tabla 12. Requisitos sistema DSPACE CRIS.

Requisitos mínimos	Requisitos recomendados
8 GB de memoria RAM	16 GB de memoria RAM
100 GB de almacenamiento interno	1 TB de almacenamiento interno

Fuente: Elaboración propia.

De acuerdo al consumo se determina contar con 4 núcleos de procesador mínimos para el sistema DSPACE CRIS.

Tabla 13. Requisitos sistema RIMS.

Requisitos mínimos	Requisitos recomendados
4 GB de memoria RAM	16 GB de memoria RAM
50 GB de almacenamiento interno	500 GB de almacenamiento interno

Fuente: Elaboración propia.

De acuerdo al consumo se determina contar con 4 núcleos de procesador mínimos para el sistema RIMS.

Con base en los hallazgos en la documentación oficial y las pruebas de manera local, se propone tomar en cuentas los siguientes recursos de hardware:

Tabla 14. Requerimientos de Hardware Sistema DSPACE CRIS.

	Mínimo	Recomendado
RAM	8GB	16GB
Procesador (núcleos)	4 N	8N
Almacenamiento	100 GB	1TB
Base de Datos	1GB	10GB

Fuente: Elaboración propia.

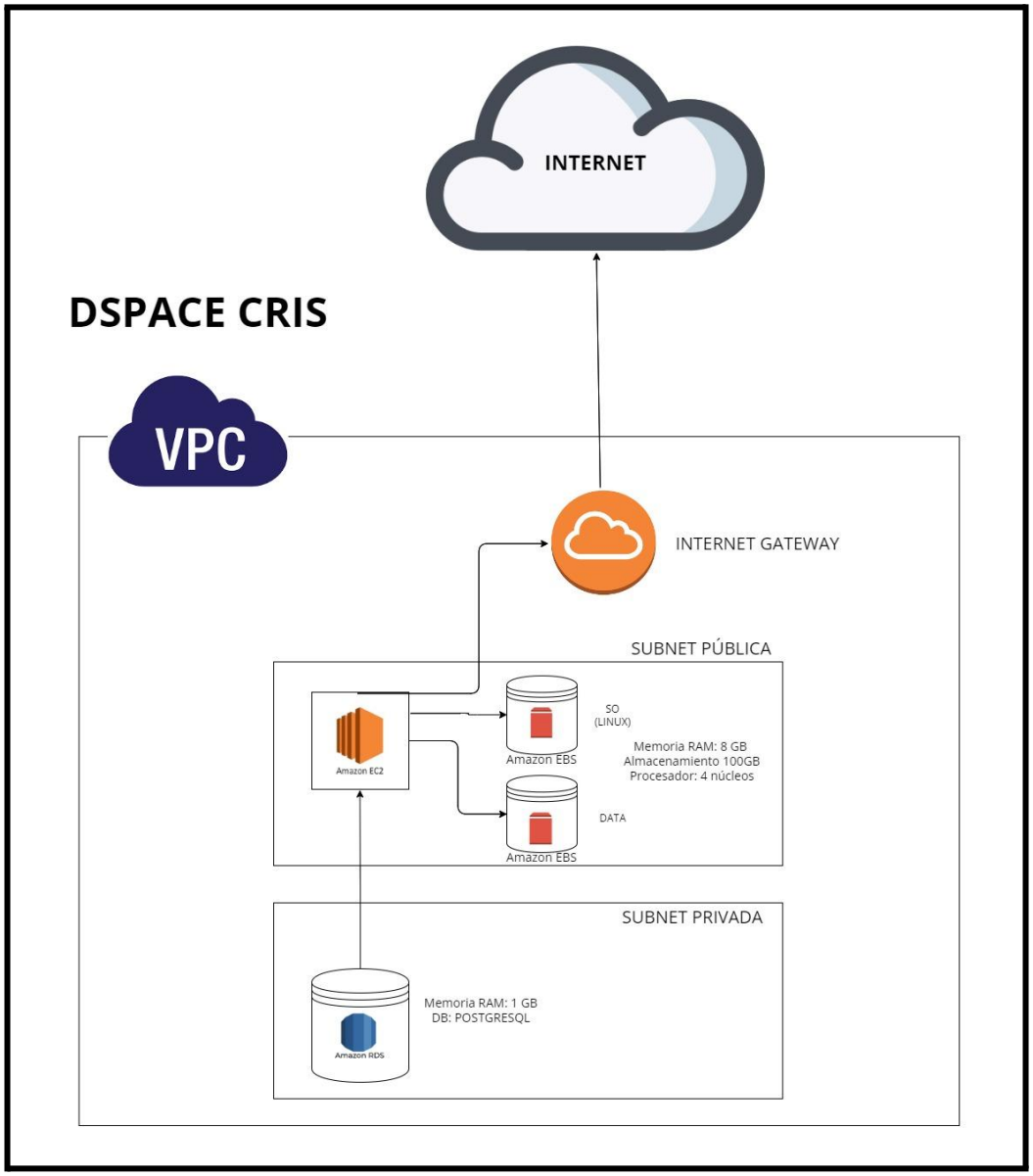
Tabla 15. Requerimientos de Hardware Sistema RIMS.

	Mínimo	Recomendado
RAM	4GB	16GB
Procesador (núcleos)	4 N	8N
Almacenamiento	50GB	1TB
Base de Datos	1GB	10GB

Fuente: Elaboración propia.

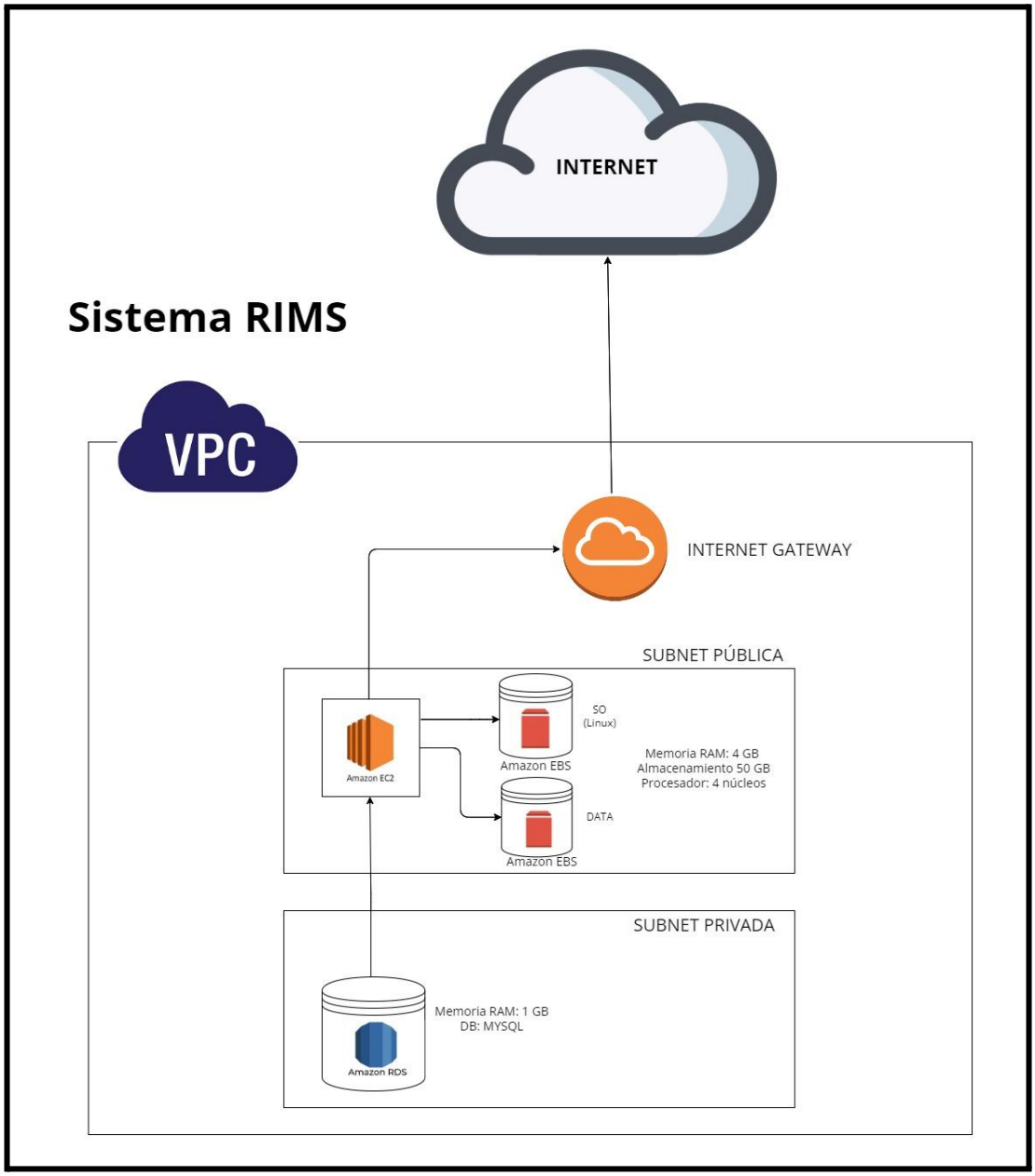
Al analizar los requerimientos de hardware propuestos, se procedió a investigar en la colección de servicios de computación en la nube ofrecidas a través de Internet por Amazon y se propone el siguiente escenario:

Figura 12. Arquitectura AWS DSPACE CRIS.



Fuente: Elaboración propia.

Figura 13. Arquitectura AWS RIMS.



Fuente: Elaboración propia.

Tomando en cuenta cada uno de los elementos y servicios de aws propuestos para los escenarios anteriores, se presenta la siguiente estimación de costos:

Tabla 16. Cotización de Costos.

Servicio	Pago mensual	Total de los primeros 12 meses	Moneda	Capacidad
Amazon EC2	25.57	306.84	USD	
Amazon RDS for MySQL	31.13	373.56	USD	30 GB
Amazon Elastic Block Store (EBS)	4.48	53.76	USD	100 GB
Amazon EC2	25.57	306.84	USD	
Amazon Elastic Block Store (EBS)	2.76	33.12	USD	50 GB
Amazon RDS for PostgreSQL	31.47	377.64	USD	30 GB
Total	120.98	1451.76	USD	

Fuente: Elaboración propia.

Para los servicios de AWS, requeridos para cada uno de los sistemas se han seleccionado las siguientes instancias:

Instancias Familia T3a

Las instancias Amazon EC2 T3a son el tipo de instancia ampliable de uso general de próxima generación que proporciona un nivel básico de rendimiento de la CPU con la capacidad de ampliar el uso de la CPU en cualquier momento durante el tiempo que sea necesario.

Las instancias T3a ofrecen un equilibrio de recursos informáticos, de memoria y de red y están diseñadas para aplicaciones con un uso moderado de CPU que experimentan picos temporales de uso. Las instancias T3a ofrecen hasta un 10 % de ahorro de costos en comparación con tipos de instancias comparables (Amazon Web Services, s. f.-b).

EC2

EC2, Nube Informática Elástica (Elastic Compute Cloud), es un servicio de AWS que proporciona capacidad de cómputo escalable en la nube.

Se ha seleccionado una instancia de tipo t3a.xlarge, que cuenta con las siguientes características:

Tabla 17. Características Amazon EC2.

Instancia	CPU Virtual	Memoria (GiB)
t3a.xlarge	4	16

Fuente: Elaboración propia.

El sistema operativo seleccionado para esta configuración es Linux. Inicialmente, se sugiere optar por el plan de pago por demanda, utilizando sólo 170 horas al mes. Este plan tiene un costo mensual de $25.57.

Posteriormente, se recomienda considerar la opción de pago que permite hacer una reserva por 1 año y luego por 3 años. Esta estrategia tiene como objetivo permitir un análisis más preciso del uso real de los recursos de los sistemas y, al mismo tiempo, obtener un descuento económico significativo.

RDS para MySQL

RDS, que significa Servicio de Base de Datos Relacional (Relational Database Service). "MySQL es la base de datos relacional de código abierto más popular del mundo y Amazon RDS hace que resulte más sencillo configurar, utilizar y escalar implementaciones de MySQL en la nube. Con Amazon RDS, se puede implementar servidores de MySQL escalables en unos minutos con una capacidad de equipo rentable y redimensionable" (Amazon Web Services, s. f.-d).

Para este servicio, se ha seleccionado el modelo de tipo db.t3.small, que cuenta con las siguientes características:

Tabla 18. Características RDS para MySQL.

Modelo	CPU Virtual	Memoria (GiB)	Almacenamiento (GB)	Opción de implementación
db.t3.small	2	2	30	Single-AZ

Fuente: Elaboración propia.

Inicialmente, se sugiere optar por el plan de pago por demanda, con un volumen de almacenamiento (SSD de uso general (gp2)), utilizando sólo 170 horas al mes. Este plan tiene un costo mensual de $31.13.

RDS para PostgreSQL

"Amazon RDS facilita las tareas de configuración, utilización y escalado de implementaciones de PostgreSQL en la nube. Con Amazon RDS, se pueden hacer implementaciones de PostgreSQL escalables en unos minutos con una capacidad de hardware rentable y redimensionable" (Amazon Web Services, s. f.-c).

Para este servicio al igual que el de MySQL, se ha seleccionado el modelo de tipo db.t3.small, que cuenta con las siguientes características:

Tabla 19. Características RDS para PostgreSQL.

Modelo	CPU Virtual	Memoria (GiB)	Almacenamiento (GB)	Opción de implementación
db.t3.small	2	2	30	Single-AZ

Fuente: Elaboración propia.

Inicialmente, se sugiere optar por el plan de pago por demanda, con un volumen de almacenamiento (SSD de uso general (gp2)), utilizando sólo 170 horas al mes. Este plan tiene un costo mensual de $31.47.

EBS

Elastic Block Store, es un servicio de AWS que proporciona almacenamiento en bloque para su uso con las instancias de Amazon EC2.

Se ha seleccionado este servicio, que cuenta con las siguientes características:

Tabla 20. Características Amazon EBS.

Volúmenes	Almacenamiento (GB)	Frecuencia de Instantáneas
1	50	2 veces al día

Fuente: Elaboración propia.

Inicialmente, se sugiere optar por el plan de pago por demanda, utilizando sólo 170 horas al mes y un SSD de propósito general. Este plan tiene un costo mensual de $2.76.

Recomendación de DLP

DLP, que significa Prevención de Pérdida de Datos (Data Loss Prevention). En el contexto de AWS, implementar DLP implica utilizar una combinación de servicios y configuraciones para garantizar que los datos críticos no se divulguen de manera no autorizada.

A continuación, se presentan algunas recomendaciones que podrían ser útiles para este propósito:

EC2

Aquí hay algunas sugerencias para garantizar la seguridad en EC2:

- **Seguridad en la Instancia EC2:** Utilizar grupos de seguridad para controlar el tráfico de red hacia y desde las instancias EC2. Configurar reglas específicas para permitir solo el tráfico necesario. Por ejemplo, solo se permiten conexiones SSH desde direcciones IP autorizadas y limitar el tráfico de aplicaciones a puertos específicos.

- **Almacenamiento de Datos en la Instancia:** Si la instancia EC2 maneja datos sensibles, encriptar el almacenamiento local utilizando herramientas como LUKS (Linux Unified Key Setup) para Linux.
- **Autenticación y Autorización:** Configurar adecuadamente las credenciales de acceso a la instancia EC2. Gestionar permisos y garantizar que solo usuarios autorizados tengan acceso a la instancia y a los recursos que contiene.

EBS

Aquí hay algunas sugerencias para garantizar la seguridad en EBS:

- **Encriptación de Datos:** Se recomienda usar el servicio KMS para manejo de las claves de cifrado.
- **AWS Key Management Service (AWS KMS):** Permite crear, administrar y controlar claves criptográficas en sus aplicaciones y servicios de AWS (Amazon Web Services, s. f.).
- **Instantáneas Seguras:** Si se crean instantáneas de los volúmenes EBS para realizar copias de seguridad, asegurarse de que también estén encriptadas. Esto proporciona una capa adicional de seguridad para los datos almacenados.

RDS

Es importante para asegurarse de que la comunicación entre la aplicación y la base de datos esté encriptada mediante SSL/TLS.

- **Configuración de Acceso:** Utilizar grupos de seguridad de RDS para controlar quién tiene acceso a la base de datos. Configurar las reglas de firewall para permitir conexiones sólo desde ubicaciones específicas, como las instancias EC2 que ejecutan la aplicación.
- **Gestión de Credenciales:** Implementar buenas prácticas para la gestión de credenciales, como la rotación regular de contraseñas y la restricción de permisos de usuario según el principio de privilegios mínimos necesarios.

- **Políticas de DLP:** Establecer políticas de DLP que definan qué tipos de datos son sensibles y cómo deben ser tratados. Configurar reglas para bloquear o alertar sobre la transferencia no autorizada de datos confidenciales.

Requisitos y Características Adicionales a Considerar en el Ambiente Tecnológico para la Implementación del Sistema de Información de Investigación

La implementación de un Sistema de Información de Investigación Actual requiere una consideración cuidadosa de los requisitos y características del ambiente tecnológico en el que se implementará. Además, hay características adicionales que pueden mejorar la eficiencia y la funcionalidad del sistema. En la siguiente sección, se darán a conocer estos requisitos y características en detalle, proporcionando una orientación completa para la implementación exitosa del sistema.

Mecanismos esenciales a considerar que garantizan la seguridad de la API diseñada

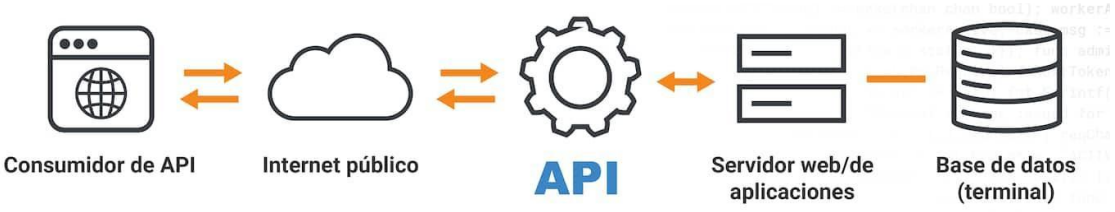
La seguridad es un aspecto fundamental en el diseño de las interfaces de programación de aplicaciones (API). Para garantizar la seguridad de la API diseñada, es crucial implementar una serie de mecanismos que protejan la API contra amenazas y vulnerabilidades. En el siguiente apartado, exploraremos en detalle que mecanismos se pueden implementar para garantizar la seguridad de la API diseñada.

¿Qué es la seguridad de las API web?

La protección de las interfaces de programación de aplicaciones (API) en la web se encarga de la transmisión segura de datos a través de las API que están vinculadas a Internet. OAuth (Autorización Abierta) es un estándar abierto que se utiliza para la delegación de acceso. Este estándar permite a los usuarios proporcionar acceso a recursos web a terceros, eliminando la necesidad de compartir contraseñas.

Utilizan Notación de Objetos de JavaScript (JSON), un formato de archivo que simplifica la transferencia de datos a través de los navegadores web. Gracias al uso de HTTP y JSON, estas API no requieren almacenar ni reorganizar los datos (RedHat, 2019).

Figura 14. Cómo funciona una API web.



Fuente: (Akamai, s. f.).

¿Por qué son tan cruciales la autenticación y la autorización para la protección de la API?

La autenticación y la autorización son componentes fundamentales para la seguridad de las API.

La autenticación es el proceso de identificar si las credenciales pasadas juntos con la solicitud son válidas o no, aquí las credenciales se pueden pasar como ID de usuario y contraseña con un token asignado para la sesión de usuario. La autenticación asegura que las peticiones a la API se originen de una fuente confiable.

La autorización es el proceso de identificar si la solicitud recibida puede acceder al punto final o método solicitado.

En la canalización del procesamiento de solicitudes, la autenticación viene primero y la autorización después. La autorización se produce solo después de la autenticación exitosa de la solicitud.

Juntos, estos procesos forman la primera línea de defensa en la protección de las API, asegurando que solo los usuarios o servicios legítimos y autorizados puedan acceder a la API.

Configurar Recursos

La configuración de recursos es un componente esencial en la seguridad de las API. Los recursos que deben ser configurados y considerados para asegurar el correcto funcionamiento de nuestra API incluyen:

- **Ancho de Banda:** Determina la cantidad de datos que se pueden transferir simultáneamente, lo cual es crucial para manejar las solicitudes de la API de manera eficiente.
- **Almacenamiento:** Es importante para el procesamiento de los datos de la API.
- **Memoria:** Es fundamental para la ejecución de la API.
- **Procesamiento:** Se refiere a la capacidad de la CPU para ejecutar instrucciones, lo cual es esencial para el procesamiento de las solicitudes de la API.

Tener en cuenta estos aspectos es vital para garantizar que las API puedan manejar eficazmente las cargas de trabajo y responder a las solicitudes de manera segura y eficiente.

Riesgos

La seguridad de las API es un aspecto crucial en el desarrollo y mantenimiento de aplicaciones web y móviles. Sin embargo, existen varios riesgos asociados con su uso. A continuación, se detallan los riesgos más comunes en las API:

- **Aprovechamientos de vulnerabilidades:** Ocurre cuando un agresor envía datos meticulosamente diseñados a un objetivo, aprovechando un fallo en la estructura del objetivo, como la inyección SQL, la desconfiguración de la seguridad, entre otros.
- **Ataques basados en la autenticación:** Los clientes necesitan autenticarse antes de poder realizar solicitudes de API para que el servidor de la API no acepte solicitudes de fuentes desconocidas o ilegítimas. Hay varias formas de hacerlo, pero cada una de ellas está sujeta a un compromiso. Por ejemplo, un atacante podría obtener las credenciales de un cliente legítimo, robar una clave de la API o interceptar y utilizar un token de autenticación (CLOUDFLARE, s. f.).

- **Errores de autorización:** La autorización establece el grado de acceso que cada usuario posee. Si la autorización no se maneja de manera adecuada, un cliente de la API podría tener acceso a información que no debería estar a su alcance, incrementando la probabilidad de una filtración de datos.
- **Ataques DoS y DDoS:** demasiadas solicitudes dirigidas a una API pueden ralentizar o detener el servicio para otros clientes. Algunos atacantes envían una cantidad enorme de solicitudes a una API a propósito en un ataque de denegación de servicio (DoS) o de denegación de servicio distribuido (DDoS) (CLOUDFLARE, s. f.).

Las estrategias de seguridad de la API pueden contribuir a disminuir estos y otros riesgos.

Prácticas recomendadas más comunes en relación con la seguridad de las API

Existen varias prácticas recomendadas que se pueden implementar para garantizar la seguridad de las API. Estas son algunas de las prácticas que se tienen que implementar para ayudar a proteger las API contra diversas amenazas y garantizar su funcionamiento seguro y eficiente:

- **Autenticación y autorización:** Según (Akamai, s. f.) las API "deben implementar mecanismos sólidos, como OAuth o JWT, para garantizar que solo los usuarios o servicios legítimos y autorizados puedan acceder a la API".
- **Utilizar tokens**: Configurar "identidades confiables y controlar el acceso a los servicios y a los recursos utilizando los tokens asignados a dichas identidades" (RedHat, 2019).
- **Utilizar métodos de cifrado y firmas**: Todos los datos transmitidos a través de las API deben cifrarse mediante protocolos seguros como TLS (es un estándar que mantiene privada la conexión a Internet y verifica que los datos enviados entre dos sistemas) y HTTPS para protegerse frente a ataques de espionaje y de máquinas intermediarias. "Solicitar el uso de firmas para asegurar que solamente los usuarios adecuados descifren y modifiquen sus datos" (RedHat, 2019).

- **Identificar las vulnerabilidades**: En 2019, RedHat señaló que es necesario "mantener actualizados los elementos de API, controladores, redes y el sistema operativo. Mantenerse al tanto de cómo funciona todo en conjunto, e identificar las debilidades que se podrían utilizar para entrar en nuestras API. Utilizar analizadores de protocolos para detectar los problemas de seguridad y rastrear las pérdidas de datos".

- **Utilizar cupos y límites**: Para (RedHat, 2019) hay que "establecer un cupo en la frecuencia con la que se puede recurrir a la API, y dar seguimiento al historial de uso. Encontrar más solicitudes a una API puede indicar un abuso. También podría ser un error de programación, como una solicitud a la API en un bucle sin fin. Establecer reglas de limitación para proteger las API de ataques de denegación de servicio y picos de uso" como las siguientes:

  - Número máximo de operaciones a ejecutar
  - Número máximo de procesos.
  - Memoria máxima asignable.
  - Tamaño de archivos de carga.

- **Utilizar una puerta de enlace de API**: Las puertas de enlace de API funcionan como el principal punto de control para el tráfico de las API. Una buena puerta de enlace permitirá autenticar el tráfico, así como controlar y analizar cómo se utilizan las API (RedHat, 2019).

- **La limitación de velocidad:** Para (CLOUDFLARE, s. f.) es necesario "poner un límite a la frecuencia con la que alguien puede repetir una acción dentro de un plazo determinado. Si un cliente de la API supera el número de solicitudes permitidas, la limitación de velocidad descartará o bloqueará sus solicitudes durante un periodo de tiempo".

- **La mitigación de DDoS:** Ayuda a detener los ataques DoS y DDoS. En un ataque DDoS, un atacante intenta sobrecargar una API con muchas solicitudes en un

período corto de tiempo. Con frecuencia, estas solicitudes proceden de varias fuentes diferentes (CLOUDFLARE, s. f.).

Servicios y herramientas para la gestión y la seguridad de las API

La gestión y seguridad de las API son aspectos fundamentales en el desarrollo y mantenimiento de aplicaciones web y móviles. Los servicios de gestión de API permiten a los desarrolladores controlar y supervisar cómo se utilizan sus APIs. Por otro lado, la seguridad de las API se centra en proteger las APIs de amenazas y vulnerabilidades. Entre los servicios más confiables y populares se encuentran:

Red Hat

Ofrece a su empresa soluciones de gestión de las API modulares, ligeras e integrales que utilizan el open source, se basan en los estándares abiertos y están disponibles tanto en las instalaciones como en la nube. "Gracias a esto, su equipo puede conectar todos los elementos, como las aplicaciones con los datos o los sistemas heredados con los nuevos, mientras usted sigue creciendo.

Red Hat, recomienda su solución galardonada Red Hat 3scale API Management, la cual incluye:

- Un administrador de API que gestiona las API, las aplicaciones y las funciones de los desarrolladores.
- Un administrador de tráfico (puerta de enlace de API) que aplica las políticas del administrador de API.
- Un centro proveedor de identidad (IDP) que admite una amplia variedad de protocolos de autenticación.

En la puerta de enlace de API, Red Hat 3scale API Management decodifica los tokens con registro de hora que caducan; comprueba que la identificación del cliente sea válida y confirma la firma mediante una clave pública" (RedHat, 2019).

Akamai

Ofrece una seguridad de API completa con total visibilidad, detección de endpoints y detección de firmas y comportamientos, así como políticas de respuesta automatizadas y en tiempo real.

La solución estrella de WAAP de Akamai, App & API Protector, salvaguarda los sitios web, las aplicaciones y las API al bloquear el tráfico entrante perjudicial en tiempo real. Dentro del tráfico permitido por App & API Protector y en cualquier parte de la empresa, API Security identifica todas las API, emplea análisis de comportamiento para detectar actividades inusuales y responde automáticamente a amenazas y abusos.

Además, Akamai Identity Cloud puede fortalecer los protocolos de autenticación de usuarios. Su equipo de expertos en seguridad, desde servicios gestionados hasta especialistas en la búsqueda de amenazas, también puede proporcionar asesoramiento experto sobre las mejores prácticas para la seguridad de API, ayudando a su empresa a establecer un entorno de API seguro (Akamai, s. f.).

Microsoft Azure

Este servicio utiliza la herramienta API Management (APIM ahora) que tiene como objetivo proporcionar un enfoque centralizado para la gestión de servicios, de manera que sea transparente para los mismos.

Técnicamente, APIM implementa el patrón de API Gateway, exponiendo una ruta central única donde se pueden registrar subrutas, cada una conectada a una API diferente. APIM recibe la solicitud y la redirige a la API de destino. Se presenta como un servicio con baja latencia que puede manejar millones de solicitudes diarias sin afectar sus tiempos de respuesta.

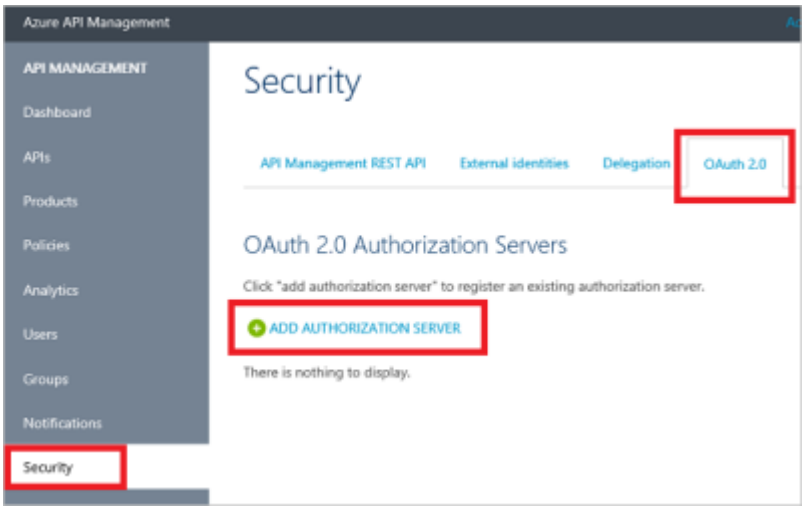
En su nivel básico, APIM ofrece un portal donde se administran las APIs, su publicación como producto, sus versiones, y permite definir su público y las reglas de acceso, como la necesidad de registro y su aprobación. También proporciona a los clientes

un portal con la documentación de la API, un cliente de acceso para pruebas, blogs, informes de problemas y su gestión.

Además, cada solicitud pasa por un sistema de 'reglas' que permiten enriquecer la información, bloquear la solicitud y mucho más. A nivel de seguridad, lo que se expone públicamente no son los servicios, sino la propia APIM. Esto simplifica la seguridad, ya que los servicios están en su red o en su red privada en Azure y no hay forma de acceder a ellos directamente desde la web.

APIM no se limita simplemente a transmitir la solicitud, implementa un sistema de autenticación, basado en Azure Ad y con la posibilidad de conectar con servidores OAUTH. Puede insertar, modificar o eliminar encabezados, reclamaciones. Todas las solicitudes son autenticadas. También es posible limitar el número de solicitudes, o el ancho de banda utilizado, para limitar a los clientes demasiado 'agresivos'.

Figura 15. Security (OAuth 2.0).

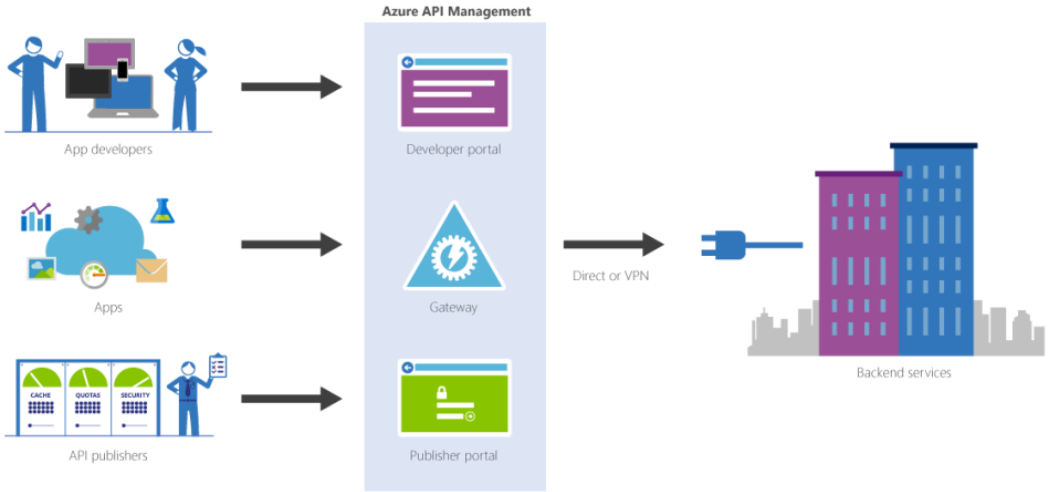


Fuente: (ITblogsogeti, 2018).

A nivel de rendimiento, puede definir políticas de caché, utiliza una caché distribuida de alto rendimiento, y permite definir operaciones sujetas a caché y sus variables, si es privada, pública, qué parámetros o encabezados utilizar para distinguir las respuestas a devolver.

Además, ofrece estadísticas de uso de cada API y la ubicación geográfica de las llamadas, análisis de datos, posibilidad de notificaciones y otras utilidades (ITblogsogeti, 2018).

Figura 16. Azure API Management.

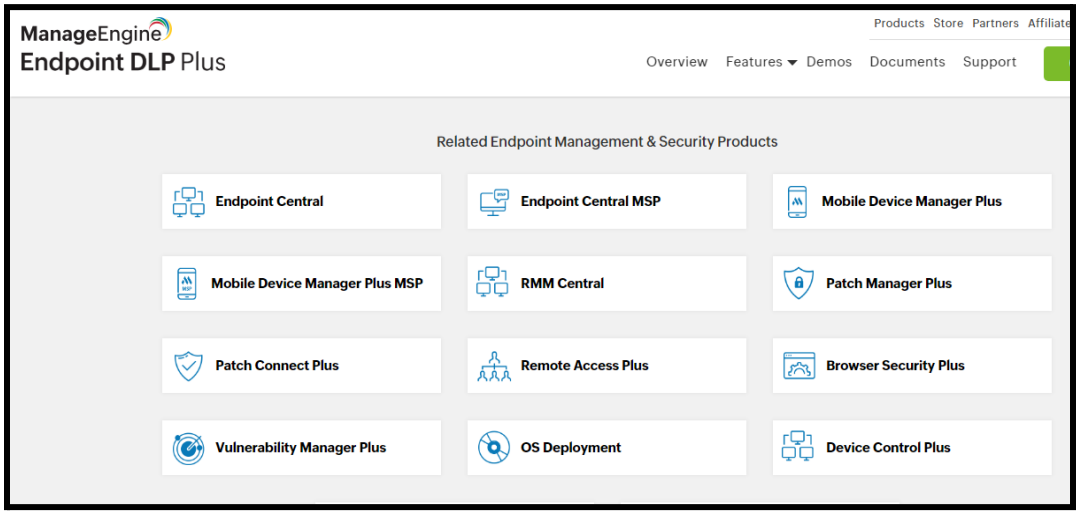


Fuente: (ITblogsogeti, 2018).

Endpoint DLP

De acuerdo con ManageEngine (s. f.) Endpoint DLP Plus protege los datos confidenciales de los dispositivos de endpoint gestionados contra la divulgación y el robo de datos mediante estrategias avanzadas de prevención de pérdida de datos, que incluyen la detección y la clasificación de los datos, así como la definición de reglas para el uso autorizado y la transmisión segura.

Figura 17. Endpoint DLP Plus.



Fuente:(ManageEngine, s. f.).

SSO (Single Sign On) LastPass

Según el sitio oficial de LastPass (LastPass, s. f.), se detallan las siguientes características:

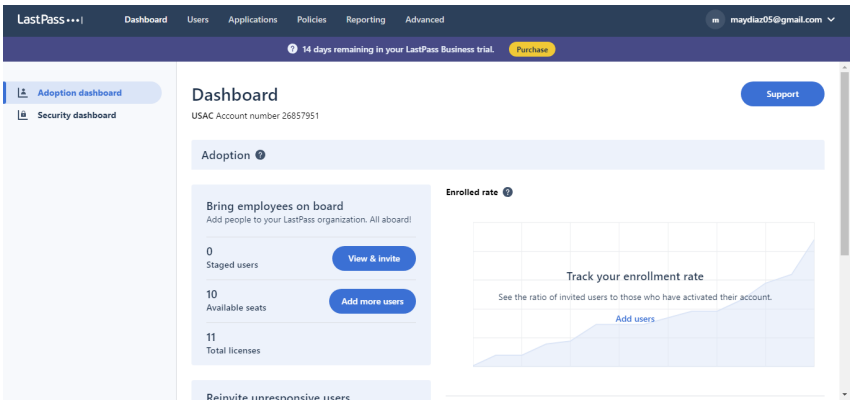
- Ofrece a cada usuario su propia bóveda cifrada para cuentas, contraseñas, notas y demás.
- Combina la gestión de contraseñas con el inicio de sesión único y la autenticación multifactor para una protección total de la identidad.
- Optimiza la implementación sin descuidar la seguridad con más de 100 políticas de autenticación y acceso personalizables.
- Gestiona los usuarios y las políticas de seguridad desde un panel de administración unificado y visible

Según Ramírez (2016) "Una ventaja importante de LastPass es que ofrece la posibilidad de generar contraseñas de un solo uso. Si se usa un computador público que no es de confianza por ejemplo donde se sospeche que tiene un KEYLOGGER instalado, y se requiere acceder a los datos en LastPass",

Entre otras ventajas conocidas podemos encontrar las siguientes mencionadas por Carlos (2019):

- Dedicada al entorno empresarial.
- Ofrece un buen sistema para la gestión de equipos de trabajo permitiendo asignar grupos de contraseñas en función de equipos o incluso de proyectos.
- Permite un control muy detallado sobre la asignación de usuarios a los secretos almacenados.

Figura 18. SSO Lastpass.

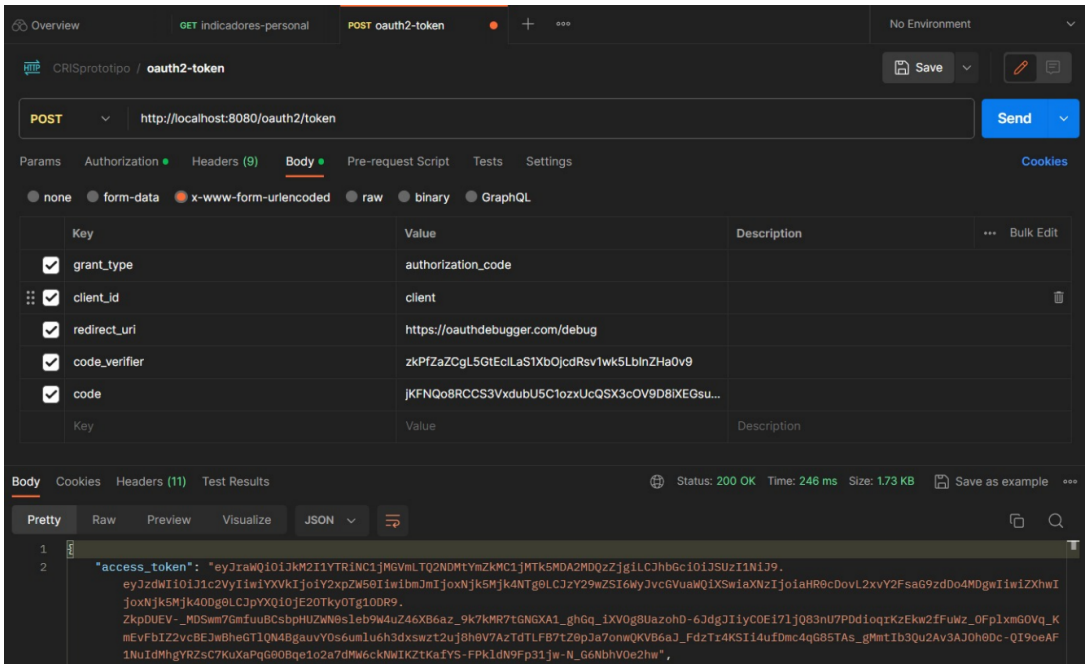


Fuente: (LastPass, s. f.).

Autorización con OAuth

OAuth 2.0 es el protocolo de autorización estándar de la industria. OAuth 2.0 se centra en la simplicidad del desarrollador del cliente y al mismo tiempo proporciona flujos de autorización específicos para aplicaciones web, aplicaciones de escritorio, teléfonos móviles y dispositivos de sala de estar. Esta especificación y sus extensiones se están desarrollando dentro del Grupo de Trabajo OAuth del IETF (*OAuth 2.0 — OAuth*, s. f.-b).

Figura 19. Consumo de datos del Prototipo OAuth 2.0.



Fuente: Elaboración propia.

Figura 20. Endpoints de OAuth 2.0 en Spring Boot.


```
{
  "issuer": "http://localhost:8080",
  "authorization_endpoint": "http://localhost:8080/oauth2/authorize",
  "token_endpoint": "http://localhost:8080/oauth2/token",
  "token_endpoint_auth_methods_supported": [
    "client_secret_basic",
    "client_secret_post",
    "client_secret_jwt",
    "private_key_jwt"
  ],
  "jwks_uri": "http://localhost:8080/oauth2/jwks",
  "response_types_supported": [
    "code"
  ],
  "grant_types_supported": [
    "authorization_code",
    "client_credentials",
    "refresh_token"
  ],
  "revocation_endpoint": "http://localhost:8080/oauth2/revoke",
  "revocation_endpoint_auth_methods_supported": [
    "client_secret_basic",
    "client_secret_post",
    "client_secret_jwt",
    "private_key_jwt"
  ],
  "introspection_endpoint": "http://localhost:8080/oauth2/introspect",
  "introspection_endpoint_auth_methods_supported": [
    "client_secret_basic",
    "client_secret_post",
    "client_secret_jwt",
    "private_key_jwt"
  ],
  "code_challenge_methods_supported": [
    "S256"
  ]
}
```


Fuente: Elaboración propia.

Seguridad de la API de DSpace-CRIS

La seguridad en la comunicación y las peticiones a los endpoints en DSpace-CRIS se manejan a través de varios mecanismos:

1. **Autenticación:** DSpace-CRIS utiliza un sistema de autenticación para verificar la identidad de los usuarios antes de permitirles acceder a ciertos endpoints.
2. **Autorización:** Una vez que un usuario está autenticado, DSpace-CRIS utiliza un sistema de autorización para determinar qué endpoints puede acceder el usuario.

3. **Control de acceso:** DSpace-CRIS tiene una bandera de estado para cada instancia de entidad. Cuando la bandera está configurada como "Público", la página de detalles es visible para cualquiera y se enlazará cuando sea apropiado. El registro se incluye en los resultados de la búsqueda pública.
4. **Seguridad de los datos:** DSpace-CRIS protege los datos almacenados en su sistema a través de varios mecanismos, incluyendo el cifrado (Donohue, 2018).

La implementación específica de estos mecanismos puede variar dependiendo de la configuración del sistema DSpace-CRIS.

Cada uno de los elementos anteriormente presentados, se han tomado en cuenta para el ambiente tecnológico en la Nube, para la implementación del sistema DSPACE CRIS y para el desarrollo del sistema RIMS, que son necesarios para la creación, desarrollo, mantenimiento y seguridad del sistema en cuestión.

10.4 Resultados objetivo 4

Diseño arquitectónico del Sistema de Información de Investigación Actual para USAC

Descripción general del diseño arquitectónico del sistema

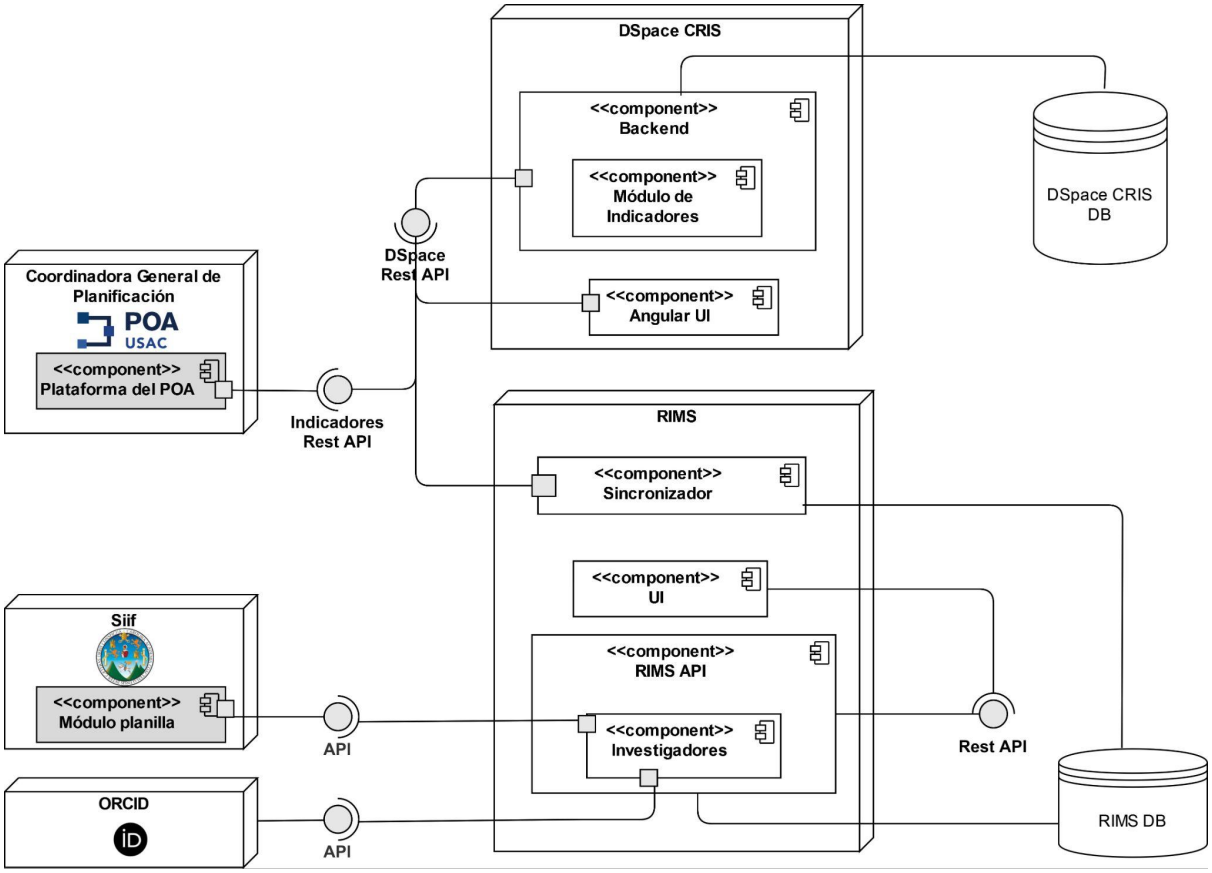
Las figuras siguientes muestran las dos opciones que se proponen para el despliegue de sistemas y servicios. Ambos diagramas contemplan los elementos a implementar: DSpace CRIS, frontend y backend, módulos del RIMS: sincronizador, investigadores, indicadores y la interfaz y las interfaces de comunicación con entidades externas como: SIIF, ORCID y la plataforma del POA de la Coordinadora General de Planificación.

Diseño de arquitectura general del sistema (opción 1. Módulo de indicadores dentro de DSPACE CRIS)

El diagrama presenta la alternativa de despliegue considerando que los indicadores se calcularán y presentarán en DSpace CRIS. Evaluando de abajo hacia arriba, el diagrama indica que la recolección y administración de información relacionada a las investigaciones se realizará en el RIMS. Lo anterior incluye la obtención de datos pertinentes del SIIF y de ORCID. El sincronizador del RIMS se conectará con la API Rest de DSpace CRIS para enviar los datos resumidos (agregados, agrupados) y en el formato adecuado para DSpace CRIS. El módulo de indicadores, que formará parte del backend de DSpace CRIS, se

encargará de administrar el cálculo y presentación de los mismos. Además, se proveerá una API de indicadores, que aproveche la API Rest de DSpace, para ofrecer los valores ya calculados de los mismos.

Figura 21. Diseño de arquitectura general del sistema opción 1.



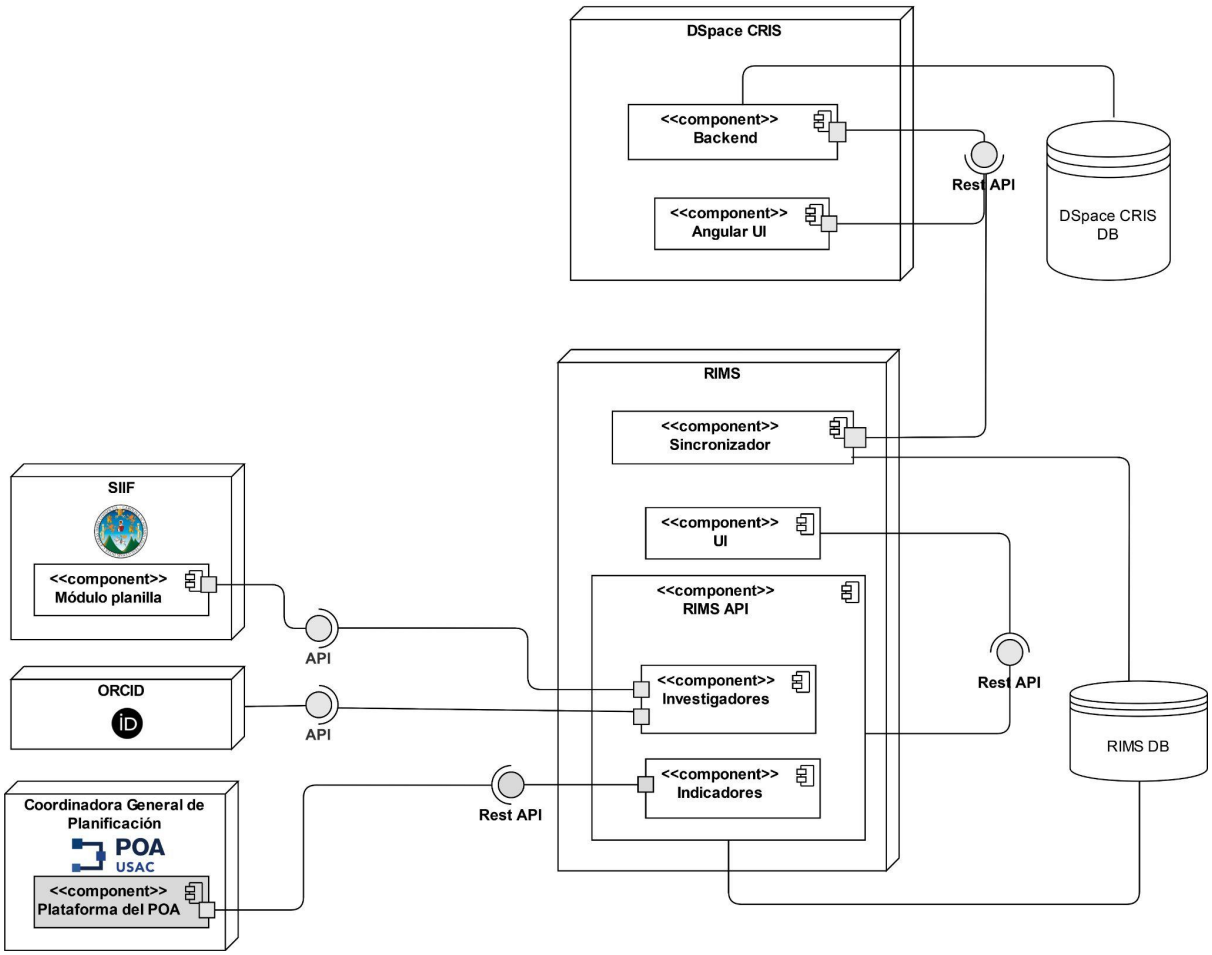
Fuente: Elaboración propia.

Diseño de arquitectura general del sistema (opción 2. Módulo de indicadores dentro de RIMS)

El diagrama muestra la opción de despliegue asumiendo que los indicadores se calcularán y presentarán en el RIMS. Recorriendo la figura de abajo hacia arriba, se interpreta que la recolección y administración de información relacionada a las investigaciones (proyectos, investigadores, redes de investigación, etc) se realizará en el RIMS. Lo anterior incluye la obtención de datos pertinentes del SIIF y de ORCID. El módulo de indicadores formará parte de la API del RIMS y se encargará de administrar el cálculo y presentación de los mismos. Además, se proveerá una API de indicadores, para

ofrecer a terceros los valores ya calculados de los mismos. El sincronizador del RIMS se conectará con la API Resta de DSpace CRIS para enviar los datos resumidos (agregados, agrupados) y en el formato adecuado para DSpace CRIS. De manera que, bajo este esquema, el CRIS daría visibilidad a proyectos, investigadores, redes de investigación, entre otros; pero no mostraría indicadores.

Figura 22. Diseño de arquitectura general del sistema opción 2.



Fuente: Elaboración propia.

DSPace 7

La elección de la API REST de DSpace 7 como la alternativa preferida para el acceso y modificación de datos se basa en una serie de consideraciones estratégicas y

técnicas que respaldan su idoneidad para nuestro contexto y objetivos específicos, como lo fueron la realización de pruebas del sistema en un entorno local con ayuda de la documentación oficial de 4SCIENCE.

La API REST de DSpace 7 ofrece una interfaz que garantiza la interoperabilidad y la flexibilidad en el manejo de datos. Esto se traduce en la capacidad de recuperar y manipular registros relacionados con nuestras entidades de interés las cuales deben estar apegadas al estándar CERIF, de manera eficiente y sin la necesidad de una manipulación directa utilizando sentencias DML (Lenguaje de Manipulación de Datos). Esto resulta en una menor complejidad en el desarrollo y mantenimiento de nuestro sistema, lo que ahorra recursos significativos en el tiempo y el costo.

También ofrece una robusta capa de seguridad, lo que garantiza que el acceso y la modificación de datos estén protegidos de manera adecuada. Esto es esencial para cumplir con los requisitos de privacidad de los datos y asegurar que solo usuarios autorizados puedan interactuar con la información.

SISTEMAS A CONSIDERAR PARA LA INTEROPERABILIDAD

- **RIMS** (Research Information Management System): Se justifica por la necesidad de administrar y gestionar eficientemente la investigación y los proyectos de investigación en el contexto de la implementación del sistema DSpace-CRIS.
- **SIIF:** Se identifica la necesidad de utilizar los datos personales de los investigadores que se encuentran registrados en el SIIF y de esta forma facilitar el ingreso de información minimizando la redundancia de trabajo.
- **ORCID:** Se ha reconocido la importancia de acceder a la información personal contenida en los perfiles de los investigadores debidamente registrados en el sistema ORCID, así como a los proyectos en los que han participado. Esta necesidad surge con el propósito de consolidar y centralizar estos datos en el sistema DSPACE-CRIS, con el fin de destacar y dar visibilidad a sus contribuciones y trabajos.

Informe final de proyecto de investigación. Año 2023

Información proporcionada por los stakeholders

- Estructura interna del Sistema Integrado de Información Financiera SIIF.
- Estructura interna de las unidades de investigación dentro de la Universidad San Carlos de Guatemala.
- Retroalimentación por parte de la Universidad UNA de Costa Rica.
- Retroalimentación por parte de la DIGI.
- Funcionamiento y gestión del Registro Universitario de Investigadores RUI.

Necesidades específicas de consumo y/o manipulación de datos entre los sistema

- **ORCID API:** Obtener el historial académico, temas de interés (*keywords*) y publicaciones (*works*) de los investigadores. La solicitud se realizará usando el ORCID ID del investigador.
- **DSPACE-CRIS**: Importación y exportación de ítems (Proyectos, investigaciones, investigadores, etc) por lotes dentro de Dspace-Cris en formato SAF (Simple Archive Format).
- **ORCID:** Servicio de importación y exportación de datos y proyectos entre los sistemas ORCID y DSPACE-CRIS.
- **RIMS:**
    - Obtener datos personales del investigador y sus contratos docentes (cuando aplique) registrados en el SIIF. También se obtendrán datos de las dependencias o unidades académicas.
    - Se usará el número de registro de personal o el Código Único de Identificación (CUI) como parámetro de búsqueda.

Alternativas para acceder y modificar la base de datos, considerando las características específicas del modelo relacional de la base de datos de DSpace CRIS.

Para garantizar una integración efectiva y eficiente entre el sistema RIMS y DSpace CRIS, es esencial que ambos sistemas comparten tecnologías comunes tanto en la base de datos como en el frontend y la API.

Ambos sistemas, RIMS y DSpace CRIS, utilizarán PostgreSQL para sus bases de datos. PostgreSQL es conocido por su robustez, escalabilidad y cumplimiento con los estándares SQL, lo que facilita su integración con otros sistemas. Comprender las relaciones entre las tablas es crucial para realizar consultas efectivas a la base de datos y garantizar que los datos se puedan recuperar y modificar de manera eficiente.

En cuanto al frontend, el nuevo sistema RIMS se desarrollará en Angular, al igual que DSpace CRIS. La consistencia en la tecnología del frontend facilitará aún más la integración entre los dos sistemas.

Es esencial crear la API del sistema RIMS utilizando la misma tecnología que DSpace CRIS, en este caso, Spring Boot. Spring Boot es un marco de desarrollo que agiliza tanto la configuración como la implementación de aplicaciones Spring. Esto facilita la creación de servicios web de manera eficaz y veloz. La elección de la misma tecnología para las APIs de ambos sistemas garantiza una comunicación eficaz y cohesionada entre ellos.

Caracterización de las alternativas consideradas para el acceso y modificación de la base de datos

PostgreSQL para la Base de Datos:

Según Phaujdar, A. sus ventajas son:

- Robustez y Confiabilidad: PostgreSQL es conocido por su sólida arquitectura y su capacidad para manejar grandes volúmenes de datos, lo que lo hace ideal para sistemas de gestión de información crítica como RIMS y DSpace CRIS.
- Cumplimiento con Estándares SQL: PostgreSQL cumple con los estándares SQL, lo que simplifica la interoperabilidad con otras bases de datos y sistemas.
- Escalabilidad: La escalabilidad horizontal y vertical permite adaptarse a las necesidades de crecimiento de la organización.

De acuerdo Phaujdar, A. sus desventajas son las siguientes:

- Mayor Curva de Aprendizaje: Para algunos usuarios, PostgreSQL puede requerir un período de aprendizaje más prolongado en comparación con bases de datos más simples.
- Consumo de Recursos: Al ser tan versátil, PostgreSQL puede consumir más recursos de hardware que soluciones más ligeras.

Angular para el Frontend:

De acuerdo con Delgado, L. y Powell, Z. sus ventajas son:

- Rendimiento y Velocidad: Angular es conocido por su alto rendimiento en aplicaciones web, lo que garantiza una experiencia de usuario fluida en sistemas como RIMS y DSpace CRIS.
- Consistencia con DSpace CRIS: Utilizar la misma tecnología en el frontend que DSpace CRIS facilita la integración y mantiene una experiencia de usuario uniforme en ambas aplicaciones.
- Amplia Comunidad y Soporte: Angular cuenta con una gran comunidad de desarrolladores y abundante documentación, lo que simplifica la resolución de problemas y actualizaciones.

Según Delgado, L. sus desventajas son las siguientes:

- Curva de Aprendizaje: Para los desarrolladores nuevos, Angular puede tener una curva de aprendizaje empinada, lo que puede retrasar el desarrollo inicial.
- Carga Inicial Más Grande: Las aplicaciones Angular pueden tener una carga inicial más pesada en comparación con tecnologías más livianas, lo que puede afectar el tiempo de carga.

Spring Boot para la API:

De acuerdo con Idbrahim, M. sus ventajas son:

- Rapidez en el Desarrollo: Spring Boot simplifica la configuración y el desarrollo de servicios web, lo que permite una implementación más rápida de la API de RIMS.
- Amplia Comunidad y Ecosistema: Spring Boot cuenta con una comunidad activa y una variedad de módulos que facilitan el desarrollo y la expansión de funcionalidades.
- Seguridad Integrada: Spring Boot ofrece características de seguridad robustas y se integra bien con soluciones de autenticación y autorización.

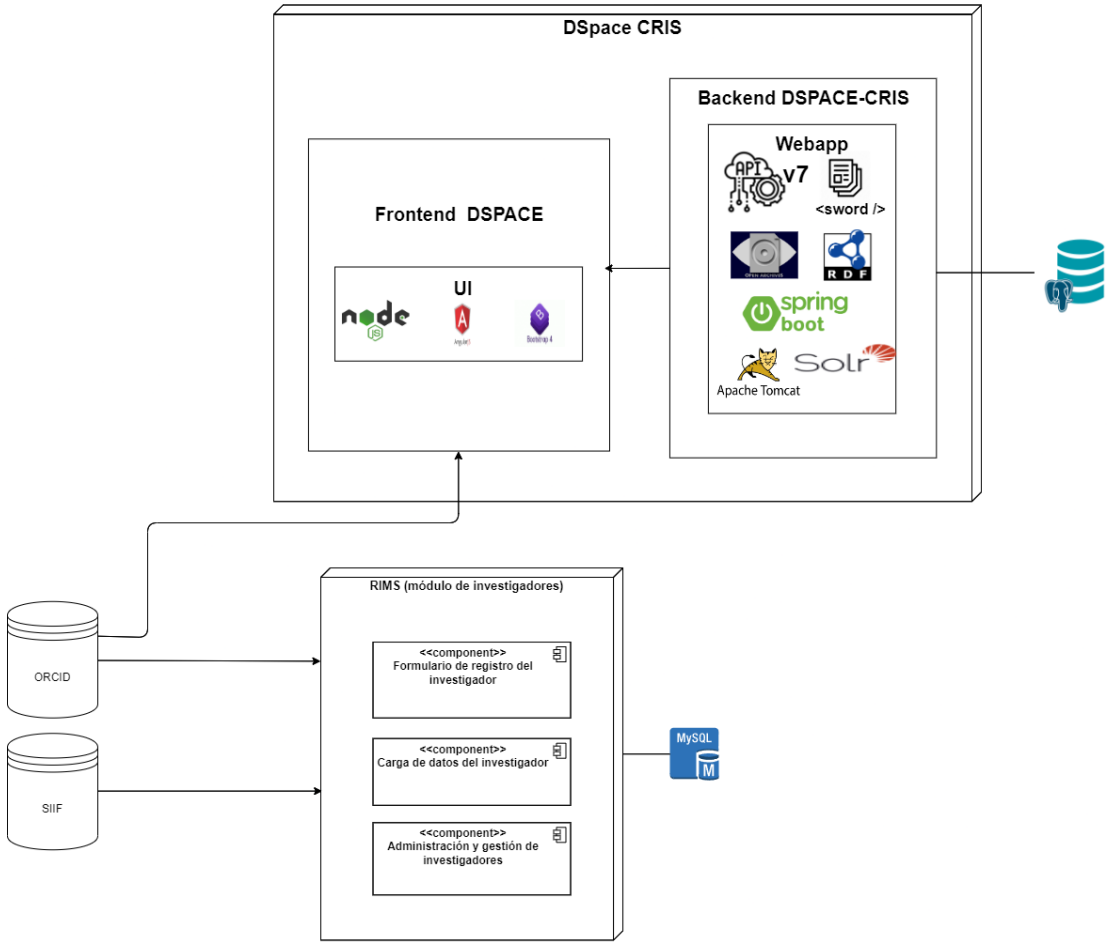
Segun Idbrahim, M. sus desventajas son:

- Flexibilidad Limitada: Aunque la simplicidad de Spring Boot es una ventaja, puede tener ciertas limitaciones para aplicaciones altamente personalizadas que requieren configuraciones muy específicas.
- Mayor Consumo de Recursos en Comparación con Tecnologías Más Livianas: Para aplicaciones extremadamente simples, Spring Boot puede ser considerado excesivo en términos de recursos.

Diagrama de tecnologías utilizadas en DSpace CRIS y RIMS (módulo de investigadores)

Estas son las tecnologías utilizadas por el sistema DSpace CRIS en su arquitectura interna tanto de backend como de frontend. Se propone utilizar las mismas tecnologías para el sistema RIMS.

Figura 23. Diagrama de tecnologías utilizadas en DSpace CRIS y RIMS.

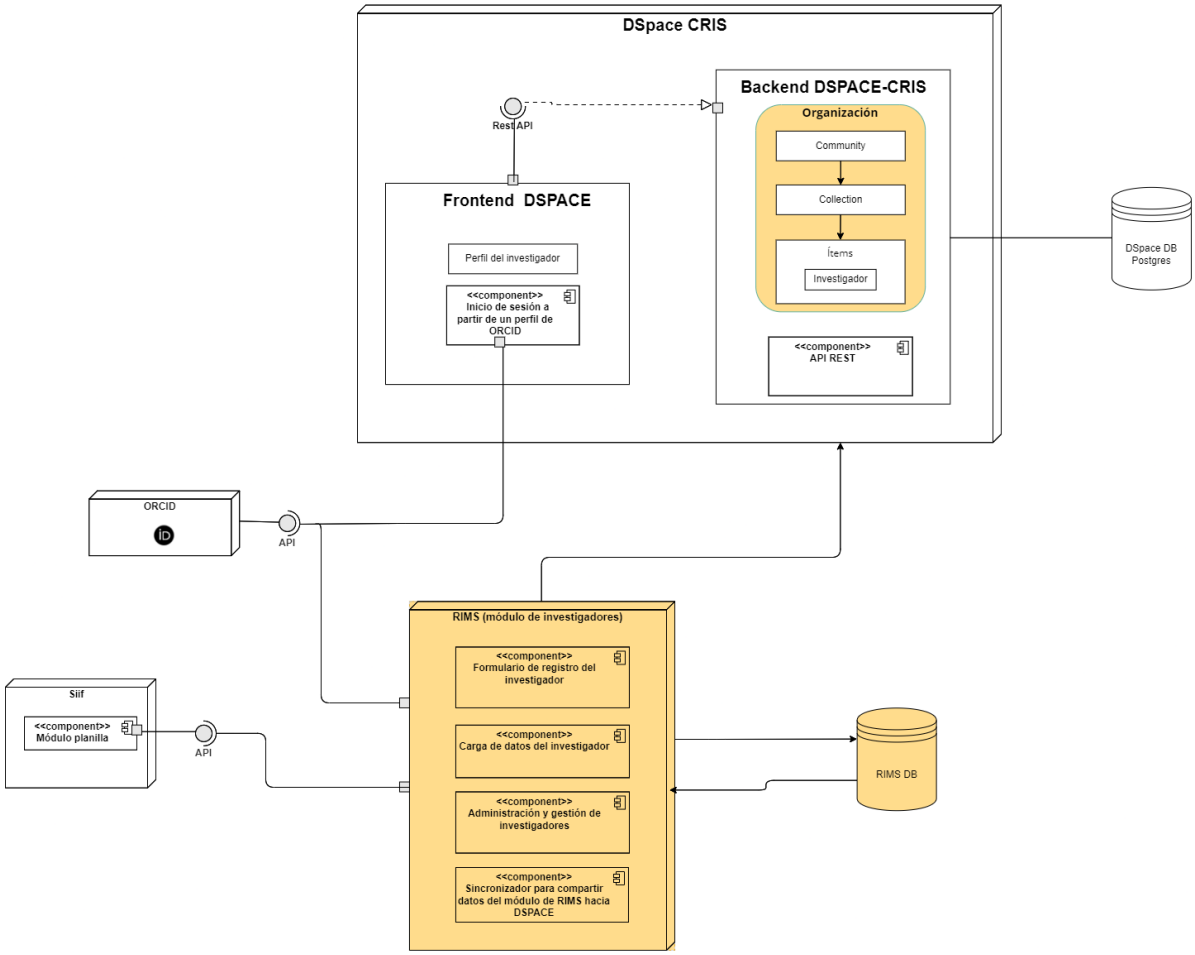


Fuente: Elaboración propia.

Diagrama de organización interna de DSpace CRIS y RIMS (módulo de investigadores)

El siguiente diagrama muestra la estructura de cómo se organizan las comunidades y colecciones enfocadas a los investigadores, la visibilidad del perfil del investigador la posibilidad de inicio de sesión en DSpace CRIS a partir de un usuario de ORCID y la interconexión con el sistema RIMS específicamente en el módulo de investigadores.

Figura 24. Diagrama de organización interna de DSpace CRIS y RIMS.

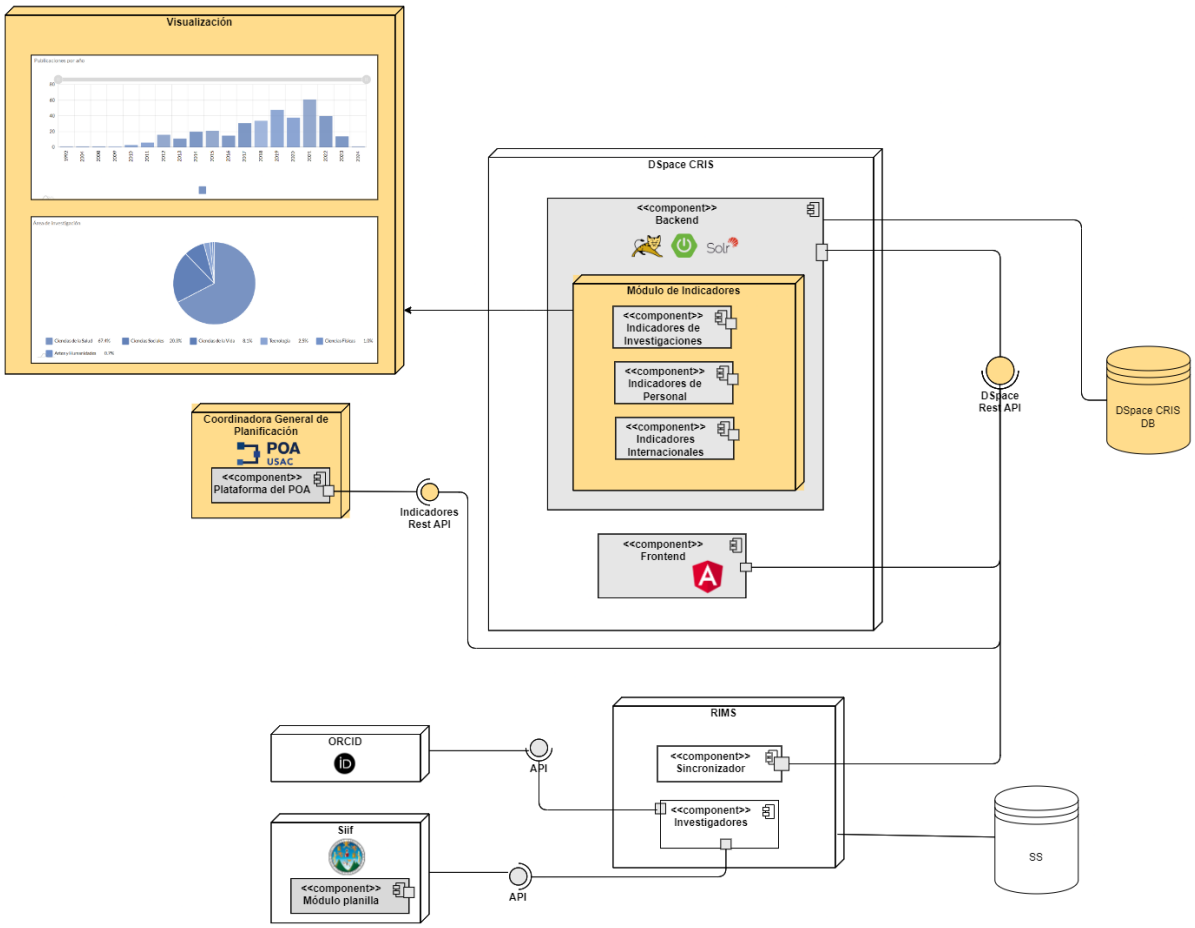


Fuente: Elaboración propia.

Arquitectura de Módulo de Indicadores Sistema CRIS

El diagrama muestra la arquitectura del módulo de indicadores, como parte del sistema DSPACE CRIS y la interoperabilidad con los diferentes sistemas que alimentan y modifican los indicadores, así como la visualización de los mismos.

Figura 25. Arquitectura de Módulo de Indicadores Sistema CRIS.



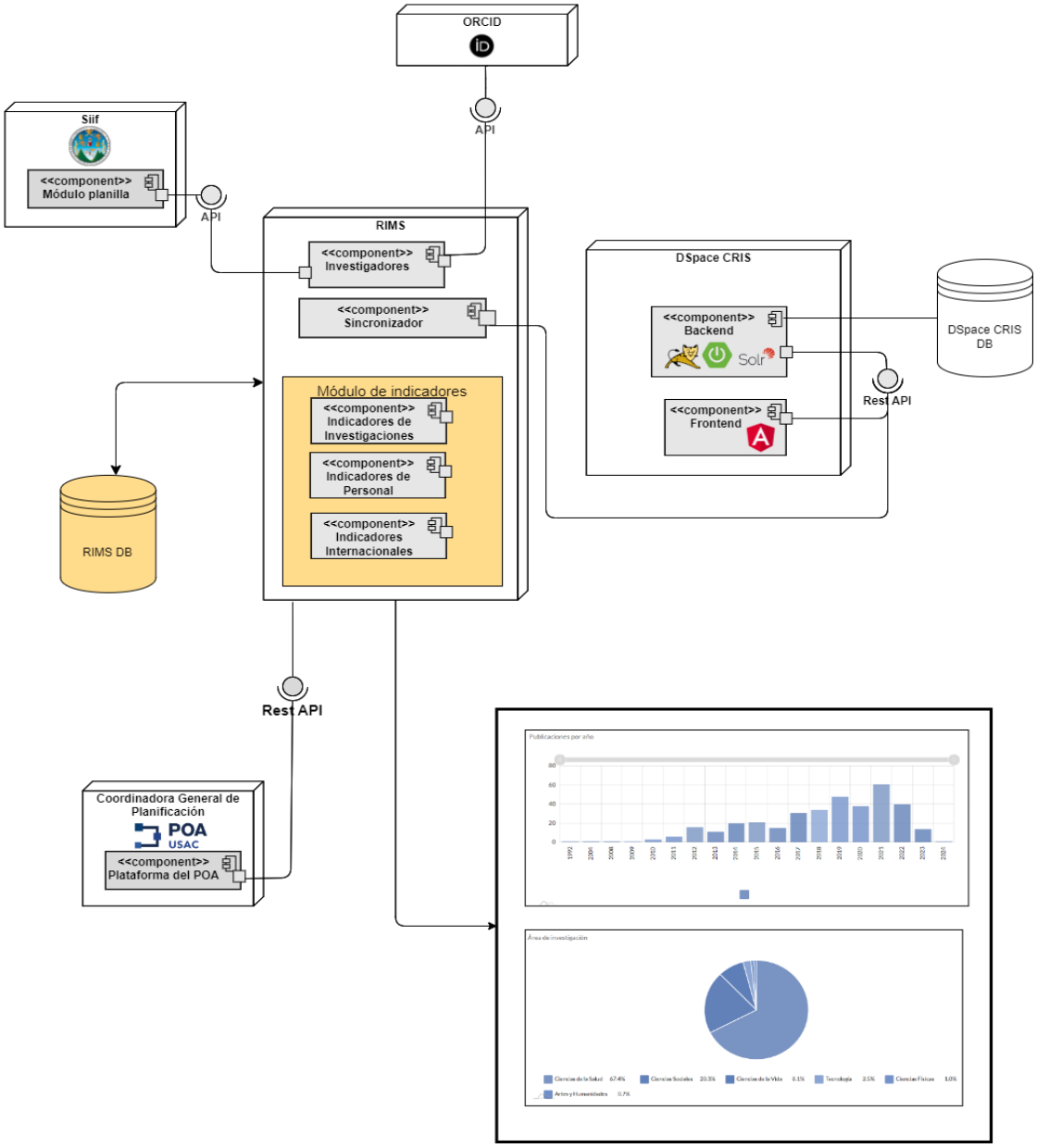
Fuente: Elaboración propia.

Informe final de proyecto de investigación. Año 2023

Arquitectura de Módulo de Indicadores Sistema RIMS

El diagrama muestra la arquitectura del módulo de indicadores,como parte de un módulo propio desarrollado dentro del sistema RIMS y la interoperabilidad con los diferentes sistemas que alimentan y modifican los indicadores, así como la visualización de los mismos.

Figura 26. Arquitectura de Módulo de Indicadores Sistema RIMS.



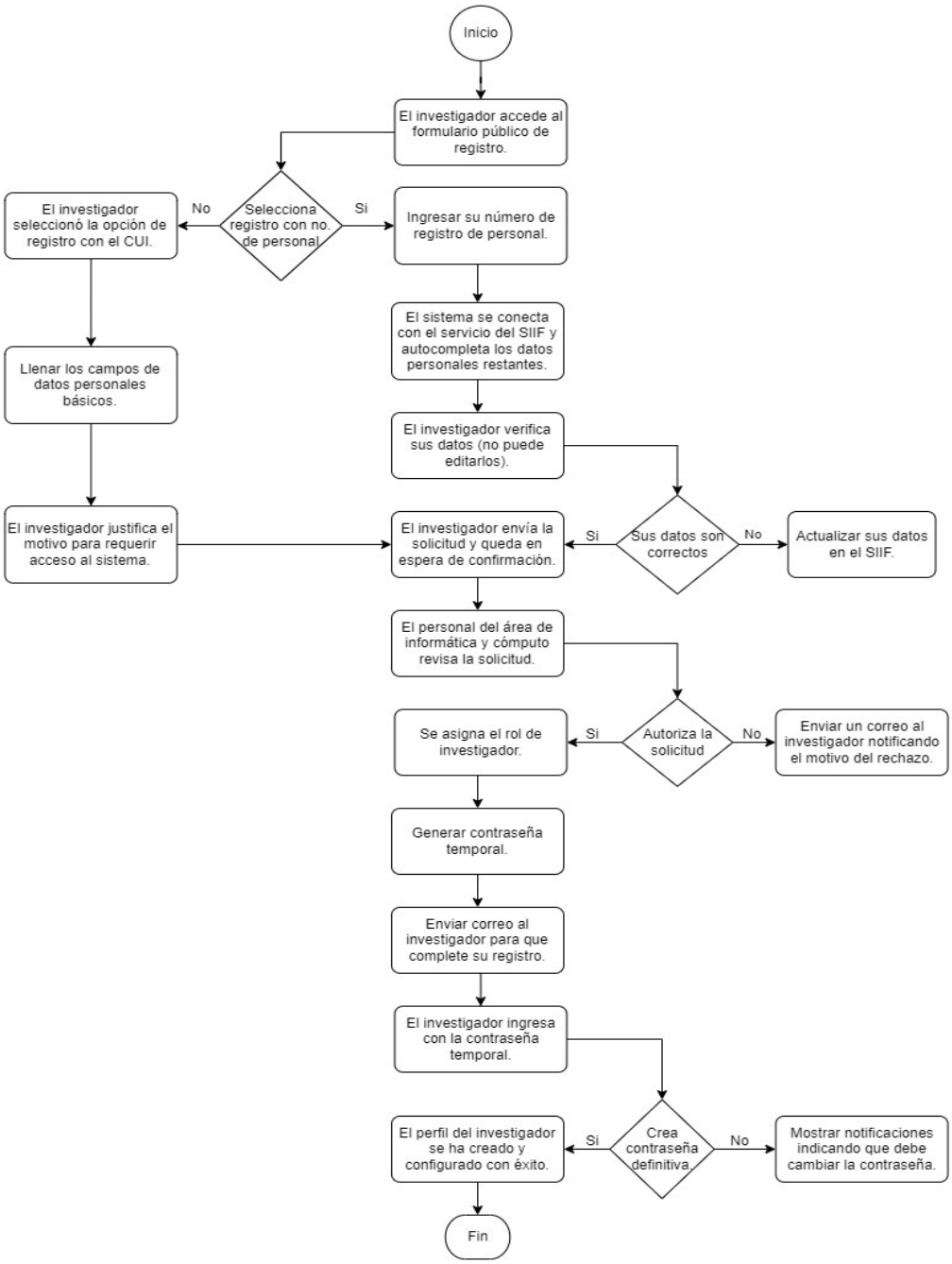
Fuente: Elaboración propia.

Informe final de proyecto de investigación. Año 2023

Registro de investigadores en el sistema RIMS

El siguiente diagrama indica el proceso que debe realizarse para poder registrar a los nuevos usuarios en el sistema ya sea por medio de su CUI o su número de registro personal.

Figura 27. Registro de investigadores en el sistema RIMS.

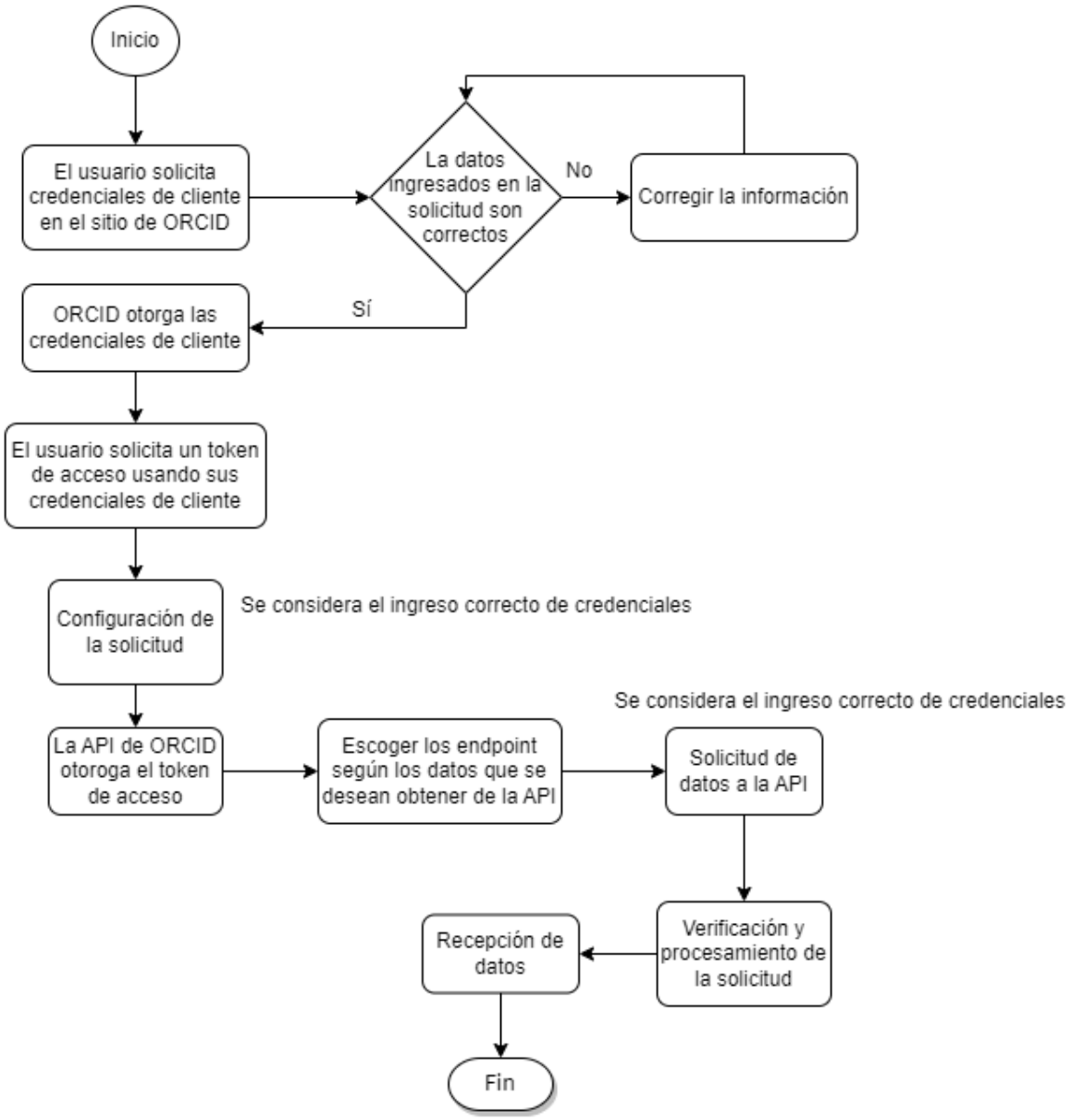


Fuente: Elaboración propia.

Consumo de datos del servicio ORCID

El siguiente diagrama muestra todo el proceso necesario para consultar datos de un perfil de investigador de ORCID haciendo uso de su API definida para ello.

Figura 28. Consumo de datos del servicio ORCID.



Fuente: Elaboración propia.

Definición de los Endpoints que se deberán proveer a terceros con los datos calculados de los indicadores de investigación

GET /indicadores: Este endpoint se utilizaría para recuperar una lista de todos los indicadores disponibles en la base de datos.

GET /indicadores/historial: Este endpoint se utilizaría para consultar el registro de revisiones o cambios realizados en un indicador en particular.

GET /indicadores/investigaciones: Este endpoint se utilizaría para recuperar una lista de todas las investigaciones registradas con apoyo de financiamiento de la DIGI.

GET /indicadores/investigaciones/tipos-investigacion: Este endpoint se utilizaría para obtener los indicadores de tipos de investigación disponibles.

GET /indicadores/investigaciones/areas-conocimiento: Este endpoint se utilizaría para obtener los indicadores de investigaciones según las áreas de conocimiento.

GET /indicadores/investigaciones/objetivos-socioeconómicos: Este endpoint se utilizaría para obtener los indicadores de investigaciones según objetivos socioeconómicos de las investigaciones.

GET /indicadores/investigaciones/programas-investigacion: Este endpoint se utilizaría para obtener los indicadores de investigaciones según programas de investigación.

GET /indicadores/investigaciones/dependencias-universitarias: Este endpoint se utilizaría para obtener los indicadores de investigaciones según dependencias universitarias avaladoras.

GET /indicadores/investigaciones/grupos-dependencias-universitarias: Este endpoint se utilizaría para obtener los indicadores de investigaciones según grupos de dependencias universitarias.

GET / indicadores /personal: Este endpoint se utilizaría para obtener los indicadores sobre el total del personal involucrado en los proyectos de investigación.

GET /indicadores/personal/distribucion-sexo: Este endpoint se utilizaría para obtener los indicadores sobre el personal según su clasificación por sexo.

GET /indicadores/personal/grado-academico: Este endpoint se utilizaría para obtener los indicadores sobre el personal según el grado académico más alto.

GET /indicadores/personal/area-conocimiento: Este endpoint se utilizaría para obtener los indicadores sobre el personal según el área del conocimiento.

GET /indicadores/personal/ocupacion: Este endpoint se utilizaría para obtener los indicadores sobre el personal según la ocupación o cargo dentro de la investigación.

GET /indicadores/personal/remuneracion: Este endpoint se utilizaría para obtener los indicadores sobre el personal según su clasificación por remuneración.

GET /indicadores/{anio}/personal: Este endpoint se utilizaría para obtener el total del personal en un año específico.

GET /indicadores/{anio}/investigaciones: Este endpoint se utilizaría para obtener el total de investigaciones en un año específico.

GET /indicadores/internacionales: Este endpoint se utilizaría para obtener los indicadores a nivel internacional disponibles en la base de datos.

GET /indicadores?filtro: Este endpoint se utilizaría para permitir la búsqueda y filtrado de indicadores según diferentes criterios.

Estados de las Peticiones de los EndPoints

201 – Creado

200 – OK

202 – Aceptado y en cola para procesamiento

204 – Sin contenido

304 – No modificado

400 Petición Incorrecta

401 – No autorizado

403 – Prohibido

404 No encontrado

405 – Método no permitido

406 – No aceptable (utilizado con tipos de contenido)

415 – Tipo de medio no compatible

429 – Dos muchas solicitudes

Correspondencia entre los datos que se obtendrán de los servicios de sistemas de terceros y los atributos de las entidades definidas en el modelo relacional del proyecto

El Sistema de Información de Investigación Actual que desarrollará la DIGI aumentará su eficiencia al aprovechar los datos que almacena el SIIF. Para los módulos de investigadores e indicadores, correspondientes a la primera fase, se han detectado requerimientos de datos sobre dependencias, empleados y contratos docentes de los empleados. La siguiente tabla especifica cuatro puntos de acceso o *endpoints* que se requieren para realizar operaciones de obtención de datos.

Tabla 21. Endpoints para obtención de datos del SIIF.

No.	Descripción	Método	*Endpoint*	Respuesta

1	Devuelve información básica de las dependencias o unidades académicas de la universidad.	GET	/dependencias	[
				...
				{
				"nombre": String,
				"acronimo": String,
				"direccion": String,
				"correo": String
				}
				…
				]
				
				Ejemplo:
				[...{
				"nombre": "Centro Universitario de Oriente",
				"acronimo": "CUNORI",
				"direccion": "Finca el Zapotillo, Zona 5, Chiquimula",
				"correo": "info@cunori.edu.gt"
				}...]

2	Recurso que se usará para obtener datos personales de un empleado identificado por su registro de personal.	GET	/empleado/{Reg. Per.} ***Parámetros:*** Reg.Per.: registro personal del empleado. ***Ejemplo:*** GET /empleado/14859	{ "cui": Number, "colegio_profesional": String, "colegiado": Number, "email_personal": String, "email_institucional":String, "primer_nombre": String, "segundo_nombre": String, "tercer_nombre": String, "primer_apellido": String, "segundo_apellido": String, "apellido_casada": String, "sexo": String, "nacionalidad": String,

No.	Descripción	Método	*Endpoint*	Respuesta
				"fecha_nacimiento": String, "telefono": String, } La fecha de nacimiento cumplirá el formato [dd-mm-YYYY].

3	Endpoint que devolverá datos personales de un empleado, identificado por su CUI. Es similar en cuanto a la estructura de la respuesta con el endpoint especificado en el inciso 2.	GET	/empleado/{CUI}
			
			Parámetros:
			
			CUI: Código Único de Identificación del empleado.
			
			Ejemplo:
			
			GET /empleado/2326866772001
				{
				"registro_personal": Number,
				"colegio_profesional": String,
				"colegiado": Number,
				"email_personal": String,
				"email_institucional":String,
				"primer_nombre": String,
				"segundo_nombre": String,
				"tercer_nombre": String,
				"primer_apellido": String,
				"segundo_apellido": String,
				"apellido_casada": String,
				"sexo": String,
				"nacionalidad": String,
				"fecha_nacimiento": String,
				"telefono": String,
				}
				
				La fecha de nacimiento cumplirá el formato [dd-mm-YYYY].

4	Proporciona información sobre los contratos de docencia del empleado identificado con el registro personal	GET	/empleado/{Reg. Per.}/docencia ***Parámetros:*** Reg.Per.: registro personal del	{ "titularidad": String, "dependencia": String, "hrs_contratadas": Number, "fecha_inicio": String, "fecha_fin": String | null, "estado": String,

No.	Descripción	Método	*Endpoint*	Respuesta
	consignado en la petición.		empleado.	}
			Ejemplo: GET /empleado/89485 /docencia	Las fechas de inicio y fin del contrato cumplirán el formato [dd-mm-YYYY].

Fuente: Elaboración propia.

11 Conclusiones

- Se creó un diseño de base de datos adaptada al estándar del modelo de datos CERIF para almacenar información sobre investigadores y elementos necesarios para construir indicadores. Este enfoque proporciona una estructura sólida y estandarizada que se ajusta a las particularidades de los procesos de investigación en la Universidad de San Carlos de Guatemala. Este diseño facilita la gestión eficiente de la información y contiene los elementos claves para la generación de indicadores significativos.

- Se definió un diccionario de datos para organizar la información manejada por el sistema. Esto permite una comprensión más profunda de los datos almacenados. La claridad en la estructura de datos contribuirá a una implementación más precisa y posteriormente a toma de decisiones informadas y eficaces. DSpace CRIS incorpora y emplea eficientemente metadatos para favorecer la interoperabilidad. Se identificaron los metadatos definidos por DSpace CRIS, correspondientes a las entidades abordadas. Para esos campos, se elaboró un esquema organizado, en el cual se definen y asocian con las tablas y las columnas de la base de datos que constituirán el origen de los valores.

- Se propuso un ambiente tecnológico que cuenta con los requisitos y características adecuadas para la implementación del ecosistema diseñado, garantizando la compatibilidad y adecuación de la infraestructura tecnológica para asegurar un rendimiento óptimo, una experiencia de usuario eficiente y respaldando así la funcionalidad integral del sistema.

- Se realizó un diseño de los mecanismos de integración de los sistemas nuevos propuestos con fuentes y sistemas preexistentes. Permitiendo que estos componentes de software puedan intercambiar información y operar de manera conjunta de manera efectiva. Además se contempló que contengan funcionalidades internas muy interrelacionadas y centradas en un propósito específico generando con esto alta cohesión. Así mismo los módulos operan de manera independiente esto contribuye a que todo el ecosistema sea más flexible y fácil de mantener.

12 Recomendaciones

- Se recomienda una implementación cuidadosa del diseño de la base de datos y del sistema, con un enfoque en pruebas exhaustivas para garantizar su funcionamiento sin problemas. Además, se sugiere establecer un plan de mantenimiento proactivo para adaptarse a las evoluciones en los procesos de investigación y tecnologías emergentes.

- Es esencial proporcionar capacitación adecuada a los usuarios del sistema para maximizar su adopción. Se recomienda desarrollar programas de formación que aborden tanto los aspectos técnicos como los funcionales, asegurando así que los usuarios puedan aprovechar plenamente las capacidades del sistema.

- Se aconseja implementar un sistema de monitoreo continuo para evaluar el rendimiento del sistema, identificar posibles mejoras y ajustar la plataforma según sea necesario. La retroalimentación constante de los usuarios también debe ser considerada para asegurar la evolución del sistema de acuerdo con las necesidades cambiantes.

- Fomentar la colaboración entre las instancias involucradas es clave. Se sugiere mantener supervisión sobre la implementación y evolución del sistema, garantizando la alineación continua con los objetivos estratégicos de la Dirección General de Investigación.

13 Referencias

4Science. (2023). DSpace (dspace-cris-7 branch). GitHub. https://github.com/4Science/DSpace/tree/dspace-cris-7

Aguilar, J. E. (2021). Estadística descriptiva, regresión y probabilidad con aplicaciones. Bogotá:Ediciones de la U.

Akamai. (s. f.). *¿Qué es un proveedor de seguridad de API?* https://www.akamai.com/es/glossary/what-is-an-api-security-provider

Amazon Web Services. (s. f.-a). *AWS Key Management Service*. Amazon Web Services, Inc. https://aws.amazon.com/es/kms/

Amazon Web Services. (s. f.-b). *Computación: Tipos de instancias de Amazon EC2 – AWS*. Amazon Web Services, Inc. https://aws.amazon.com/es/ec2/instance-types/

Amazon Web Services. (s. f.-c). *Hosted PostgreSQL, Amazon RDS for PostgreSQL, AWS*. Amazon Web Services, Inc. https://aws.amazon.com/es/rds/postgresql/

Amazon Web Services. (s. f.-d). *MySQL en host - Amazon RDS para MySQL - AWS*. Amazon Web Services, Inc. https://aws.amazon.com/es/rds/mysql/

Andrade, A. V. B., Miguel, N., Reyes, F. G., Albán, R. C., & Andrés, B. (2022). Visibility of Scientific Production and Digital Identity of Researchers through Digital Technologies. *Education Sciences*, *12*(12), 926. https://doi.org/10.3390/educsci12120926

Asserson, A., & Jeffery, K. G. (2010). CRIS and Institutional Repositories. *Data Science Journal*, *9*, CRIS14-CRIS23. https://doi.org/10.2481/dsj.cris3

Azeroual, O., Ershadi, M. J., Azizi, A., Banihashemi, M., & Abadi, R. F. (2021). Data Quality Strategy Selection in CRIS: Using a Hybrid Method of SWOT and BWM. *Informatica*, *45*(1). https://doi.org/10.31449/inf.v45i1.2995

Azeroual, O., & Schöpfel, J. (2019). Quality Issues of CRIS Data: An Exploratory Investigation with Universities from Twelve Countries. *Publications*, *7*(1), 14. https://doi.org/10.3390/publications7010014

Beigel, F. (2021). LOS SISTEMAS CRIS, SU POTENCIALIDAD PARA VISIBILIZAR DIVERSAS FORMAS DE PRODUCCIÓN E IMPULSAR NUEVAS. . . *ResearchGate*. https://www.researchgate.net/publication/359985658_LOS_SISTEMAS_CRIS_SU_POTENCIALIDAD_PARA_VISIBILIZAR_DIVERSAS_FORMAS_DE_PRODUCCION_E_IMPULSAR_NUEVAS_MODALIDADES_DE_EVALUACION

Bryant, R., Clements, A., Feltes, C., Groenewegen, D., Huggard, S., Mercer, H., Missingham, R., Oxnam, M., Rauh, A. E., & Wright, J. (2017). Research information management: defining RIM and the library's role. *OCLC Research*. https://doi.org/10.25333/c3nk88

Carlos, A. R. (2019). Gestor de secretos Open-Source para equipos de trabajo. http://hdl.handle.net/2183/24127

CLOUDFLARE. (s. f.). *¿Cuáles son los riesgos de seguridad más habituales de las API?* https://www.cloudflare.com/es-es/learning/security/api/what-is-api-security/

Congreso de la República de Guatemala. Decreto Número 57-2008. (22 de octubre de 2008). Ley de Acceso a la Información Pública. https://www.contraloria.gob.gt/wp-content/

uploads/2018/02/10-LEY-DE-ACCESO-A-LA-INFORMACION-PU-BLICA-DECRETO-57-2008.pdf

Das, A., & Sutradhar, B. (2018). Harvesting of Additional Metadata Schema into DSpace through OAI-PMH: Issues and Challenges. SRELS Journal of Information Management, 55(1), 1-7.

De Albuquerque, P., Tettamanti, S. & Lira, A. (2021). Curso de capacitación informática para UNPSJB. http://sedici.unlp.edu.ar/handle/10915/126108

De Castro, P. (2018). The role of Current Research Information Systems (CRIS) in supporting Open Science implementation : the case of Strathclyde. *ITLib journal*. https://doi.org/10.25610/itlib-2018-0003

De Castro, P., Dvořák, J., & Simons, E. (2019). OpenAIRE compatibility for CRIS systems: recent progress. *Procedia Computer Science*, *146*, 182-189. https://doi.org/10.1016/j.procs.2019.01.093

De-Castro, P. (2022). Revisitando el panorama CRIS en España. *Anuario Think EPI*. https://doi.org/10.3145/thinkepi.2022.e16a36

Delgado, L. (2020, 29 diciembre). *Angular vs React: Cuál Elegir Para tu Aplicación*. freeCodeCamp.org. https://www.freecodecamp.org/espanol/news/angular-vs-react-cual-elegir-para-su-aplicacion/

Díaz, L. (2014, 1 julio). *Integración del Sistema de Gestión de la Investigación (CRIS) con un Repositorio Institucional. El modelo de la Universidad Carlos III de Madrid*. http://hdl.handle.net/10016/19464

Dijk, E., & Van Meel, M. (2010). Using a CRIS for e-Infrastructure: e-Infrastructure for Scholarly Publications. *Data Science Journal*, *9*, CRIS59-CRIS65. https://doi.org/10.2481/dsj.cris8

Directorate-General for Research and Innovation (European Commission). (2016). *Open Innovation, Open Science, Open to the World: A Vision for Europe*. https://op.europa.eu/en/publication-detail/-/publication/3213b335-1cbc-11e6-ba9a-01aa75ed71a1

Donohue, T. (2018, 23 octubre). *Authentication and Authorization*. DSpace 7.x Documentation - LYRASIS Wiki. https://wiki.lyrasis.org/display/DSDOC7x/Authentication+and+Authorization

Donohue, T. (2018). Metadata and Bitstream Format Registries. DSpace 7.x Documentation. https://wiki.lyrasis.org/display/DSDOC7x/Metadata+and+Bitstream+Format+Registries

Donohue, T. (2023). Metadata Recommendations. DSpace 7.x Documentation. https://wiki.lyrasis.org/display/DSDOC7x/Metadata+Recommendations

Gowandy, R. (2020). How to calculate server max requests per second - Geek culture - medium. Medium. https://medium.com/geekculture/how-to-calculate-server-max-requests-per-second-38a39bb96a85

Hernández, R. (2022, 13 abril). *The Hércules Project: An Update*. EuroCRIS. http://hdl.handle.net/11366/1958

Idbrahim, M. (2023, 25 mayo). *Spring Boot vs Django: Choosing the Right Framework for Web Development*.

https://www.blog.brightcoding.dev/2023/05/25/spring-boot-vs-django-choosing-the-right-framework-for-web-development/

Iliadis, A., Acker, A., Stevens, W., & Kavakli, S. (2022). One schema to rule them all: How Schema. org models the world of search. Journal of the Association for Information Science and Technology.

ITblogsogeti. (2018, 9 enero). *API Management para todos*. Artículos y tendencias sobre soluciones tecnológicas, IT:Blog. https://itblogsogeti.com/2018/01/09/api-management-para-todos/

Jetten, M., Simons, E., & Rijnders, J. (2019). The role of CRIS's in the research life cycle. A case study on implementing a FAIR RDM policy at Radboud University, the Netherlands. *Procedia Computer Science*, *146*, 156-165. https://doi.org/10.1016/j.procs.2019.01.090

Joint, N. (2008). Current research information systems, open access repositories and libraries. *Library Review*. https://doi.org/10.1108/00242530810899559

LastPass. (s. f.). App líder en bóveda y gestión de contraseñas con inicio de sesión único y soluciones MFA - LastPass. https://www.lastpass.com/es

ManageEngine. (s. f.). Software integrado para la prevención de pérdida de datos | Solución Data Loss Prevention (DLP Software) - ManageEngine EndPoint DLP Plus. https://www.manageengine.com/latam/endpoint-dlp/

Miller, S. (2022). Metadata for digital collections. American Library Association.

Ministerio de Finanzas Públicas. (2018). Manual de Clasificaciones Presupuestarias para el Sector Público de Guatemala. https://www.minfin.gob.gt/images/downloads/leyes_manuales/manuales_dtp/manual%20secpub%20220318.pdf

Nikolić, S., Konjovic, Z., Penca, V., & Surla, D. (2015). A CERIF Compatible CRIS-UNS Model Extension for Assessment of Conference Papers. *ResearchGate*. https://www.researchgate.net/publication/285582828_A_CERIF_Compatible_CRIS-UNS_Model_Extension_for_Assessment_of_Conference_Papers

Nurminen, M. (2014). *Preparing for CRIS: Challenges and Opportunities for Systems Integration at Finnish Universities*. http://urn.fi/URN:NBN:fi:jyu-201604072020

OAuth 2.0 — OAuth. (s. f.). https://oauth.net/2/

Paredes, N., Chaparro, E., & Morales, J. (2022). Sistema de gestión de la investigación de la Escuela Politécnica Nacional (EPN) Ecuador. Rehuso, 6(Especial), 151-160. https://doi.org/10.33936/rehuso.v6iespecial.4474

Phaujdar, A. (2023, 12 septiembre). *PostgreSQL vs MySQL: Descubre sus 12 diferencias críticas*. Kinsta®. https://kinsta.com/es/blog/postgresql-vs-mysql/

Powell, Z. (2023, 20 agosto). *Angular vs React: Una Comparación En Profundidad*. Kinsta®. https://kinsta.com/es/blog/angular-vs-react/

Ramirez Leguizamon, M. (2016). Análisis de Seguridad y Uso de Gestores de Contraseñas.

Ramírez, M., & Lugo, J. (2020). Revisión sistemática de métodos mixtos en el marco de la innovación educativa. Comunicar, 65, 9-20. https://doi.org/10.3916/C65-2020-01

RedHat. (2019, 8 enero). *La seguridad de las API*. Recuperado 12 de octubre de 2023, de https://www.redhat.com/es/topics/security/api-security

Rodríguez, A. (2015, 3 julio). *Sistemas de Gestión de la Investigación: aproximación a los CRIS Institucionales*. Universidad de Salamanca. http://hdl.handle.net/10366/129659

Ruiz, M. F., Díaz, B. H., y Noriega, M. T. (2022). Gestión e investigación en ingeniería: revisión sistemática de literatura para Iberoamérica. Revista Venezolana de Gerencia, 27(98), 597-618. https://doi.org/10.52080/ rvgluz.27.98.14

Secretaría de Planifiación y Programación de la Presidencia. (2016). Objetivos de Desarrollo Sostenible, metas priorizadas Guatemala. https://www.pnd.gt/Documentos/ODS_ Metas_priorizadas.pdf

Sosa López, R. N., & García Pérez, S. (2019). *Indicadores del sistema de investigación de la Universidad de San Carlos de Guatemala* (2015.ª-2018.ª ed.). Dirección General de Investigación. ttps://portal.digi.usac.edu.gt/wp-content/uploads/2023/01/Indicadores 2015-2018 .pdf

Sosa López, R. N., García Pérez, S., Nave Herrera, F., & Arroyo Catalán, G. (2016). *Indicadores del Sistema de Investigación de la Universidad de San Carlos de Guatemala* (2012.ª-2014.ª ed.). Dirección General de Investigación. https://portal.digi.usac.edu.gt/wp-content/uploads/2023/01/indicadores2012-2014.pd f

Sosa, R. (2023). Parámetros a requerir en las investigaciones financiadas por la DIGI y en las investigaciones realizadas en las unidades de investigación con financiamiento propio o con entidades externas (nacional o internacional). Dirección General de Investigación, Universidad de San Carlos de Guatemala.

Surla, D., Ivanović, D., & Konjović, Z. (2013). Development of the software system CRIS UNS. *International Symposium on Intelligent Systems and Informatics*. https://doi.org/10.1109/sisy.2013.6662552

Vásquez, R. (2019, 6 mayo). *Implementación de un Modelo para el desarrollo de un Sistema de Gestión de la Investigación (CRIS) como estrategia de poblamiento de un Repositorio Institucional. Plan de investigación*. Universidad de Salamanca. http://hdl.handle.net/10366/139579

Vásquez, R., Amaya, C., Hernández, E. & Martínez, J. (2017). *Desarrollo de un sistema de gestión de investigación (CRIS) para la consolidación del Repositorio Institucional NINIVE de la Universidad Autónoma de San Luis Potosí*. http://sedici.unlp.edu.ar/handle/10915/151012

Wood, M. (2023). Metadata Recommendations. DSpace 7.x Documentation. https://wiki.lyrasis.org/pages/diffpagesbyversion.action?pageId=104566744&selectedPageVersions=4&selectedPageVersions=3

Wojcik, R., Bachmann, F., Bass, L., Clements, P., Merson, P., Nord, R., & Wood, W. (2006). *Attribute-Driven Design (ADD), Version 2.0 (CMU/SEI-2006-TR-023)*. Software Engineering Institute, Carnegie Mellon University. http://resources.sei.cmu.edu/library/asset-view.cfm?AssetID=8147

14 Apéndice

ÍNDICE

Apéndice 1. Fotografías 113
Apéndice 2. Valores máximos de request para cada hora 123
Apéndice 3. Diagramas Complementarios 124
Apéndice 4. Códigos de muestra 129
Apéndice 5. Formularios 137
Apéndice 6. Guías de Instalación 167

Apéndice 1. Fotografías

Fotografía 1

Créditos: M.Sc. Bárbara Arguello
Comunicación, divulgación y vinculación
Dirección General de Investigación DIGI-USAC

Descripción: Revisión de avances del proyecto

Lugar: CUNORI salón K-5

Informe final de proyecto de investigación. Año 2023

Fotografia 2

Créditos: M.Sc. Bárbara Arguello
Comunicación, divulgación y vinculación
Dirección General de Investigación DIGI-USAC

Descripción: Revisión de avances del proyecto

Lugar: CUNORI salón K-5

Informe final de proyecto de investigación. Año 2023

Fotografía 3

Créditos: M.A. Inga. Samy Eunice Pinto Castañeda
Coordinadora del proyecto

Descripción: Presentación Vicerrectora de la UNA de Costa Rica

Lugar: Campus Central USAC edificio S-12

Informe final de proyecto de investigación. Año 2023

Fotografía 4

Créditos: M.A. Inga. Samy Eunice Pinto Castañeda
Coordinadora del proyecto

Descripción: Presentación demo DSpace

Lugar: Campus Central USAC edificio S-12

Informe final de proyecto de investigación. Año 2023

Fotografía 5

Créditos: M.A. Inga. Samy Eunice Pinto Castañeda
Coordinadora del proyecto

Descripción: Miembros del equipo trabajando

Lugar: CUNORI salón K-5

Informe final de proyecto de investigación. Año 2023

Fotografía 6

Créditos: M.A. Inga. Samy Eunice Pinto Castañeda
Coordinadora del proyecto

Descripción: Miembros del equipo trabajando

Lugar: CUNORI salón K-5

Informe final de proyecto de investigación. Año 2023

Fotografía 7

Créditos: M.Sc. Bárbara Arguello
Comunicación, divulgación y vinculación
Dirección General de Investigación DIGI-USAC

Descripción: Integrantes del equipo de trabajo

Lugar: CUNORI salón K-5

Informe final de proyecto de investigación. Año 2023

Fotografía 8

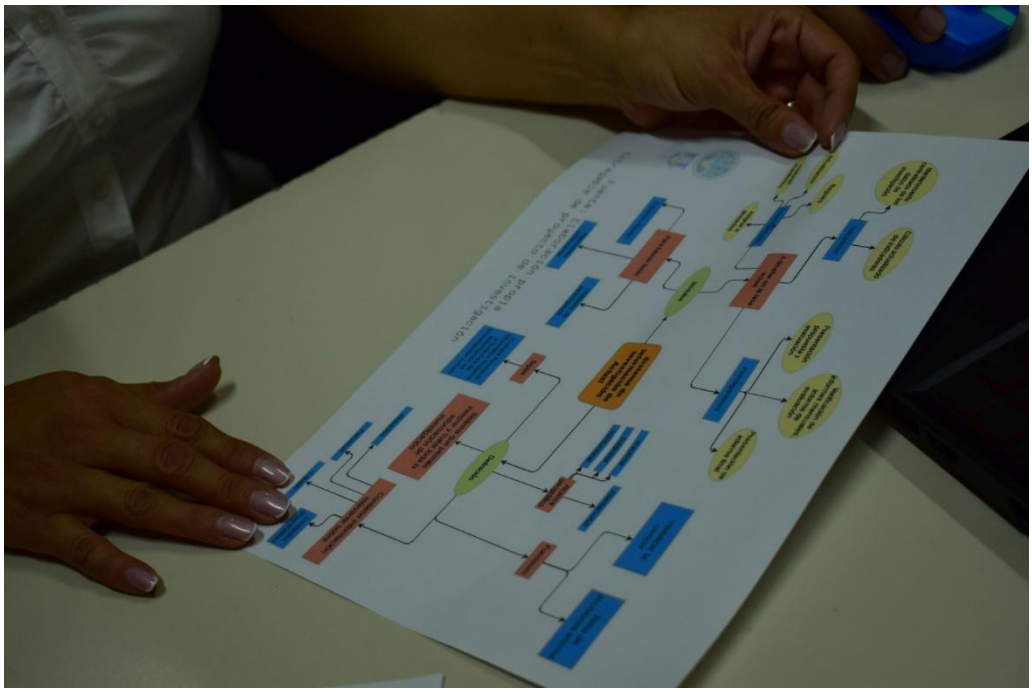
Créditos: M.Sc. Bárbara Arguello
Comunicación, divulgación y vinculación
Dirección General de Investigación DIGI-USAC

Descripción: Integrantes del equipo de trabajo

Lugar: CUNORI salón K-5

Informe final de proyecto de investigación. Año 2023

Fotografía 9



Créditos: M.Sc. Bárbara Arguello
Comunicación, divulgación y vinculación
Dirección General de Investigación DIGI-USAC

Descripción: Mapa mental general del sistema CRIS

Lugar: CUNORI salón K-5

Informe final de proyecto de investigación. Año 2023

Fotografía 10

Créditos: Jeyson Federico Martínez López
Auxiliar de investigación 1

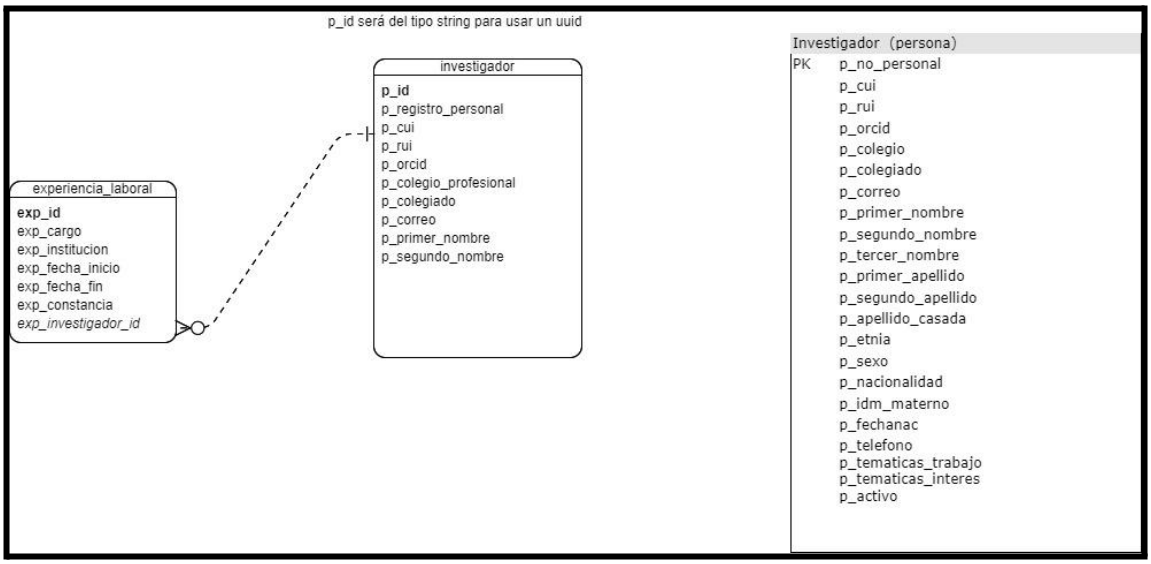
Descripción: Propuesta de personalización de DSpace CRIS

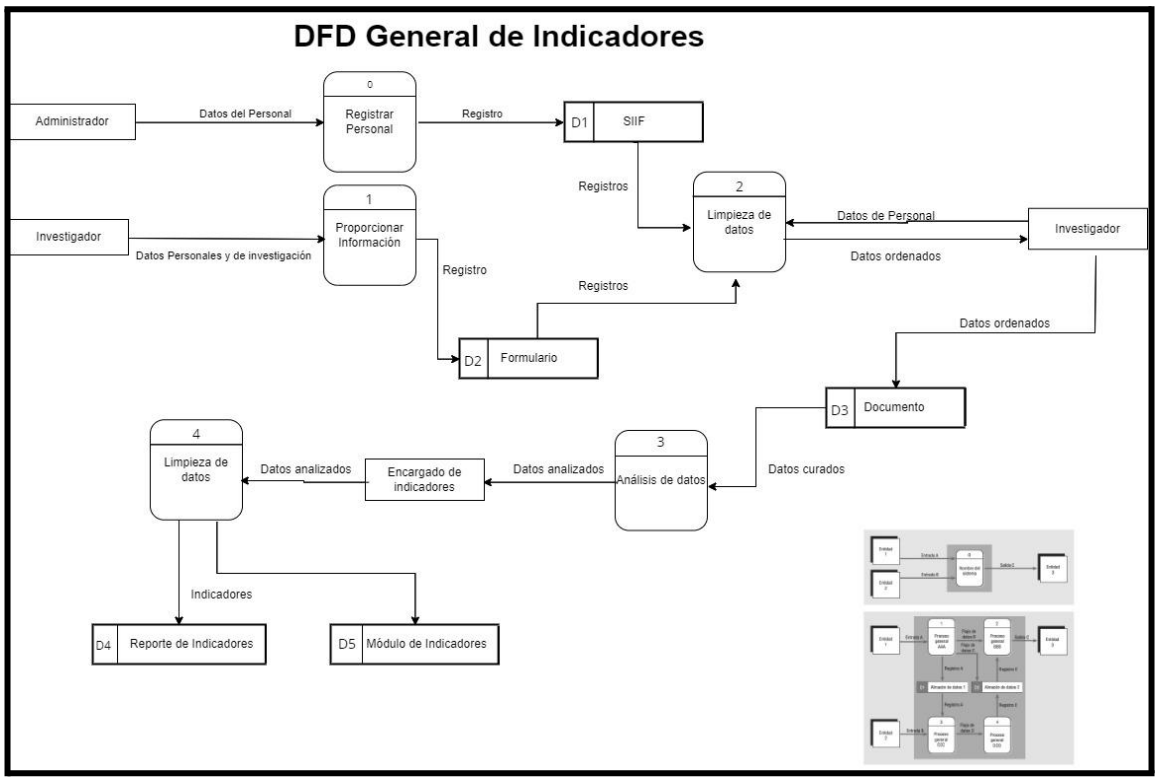
Apéndice 2. Valores máximos de request para cada hora.

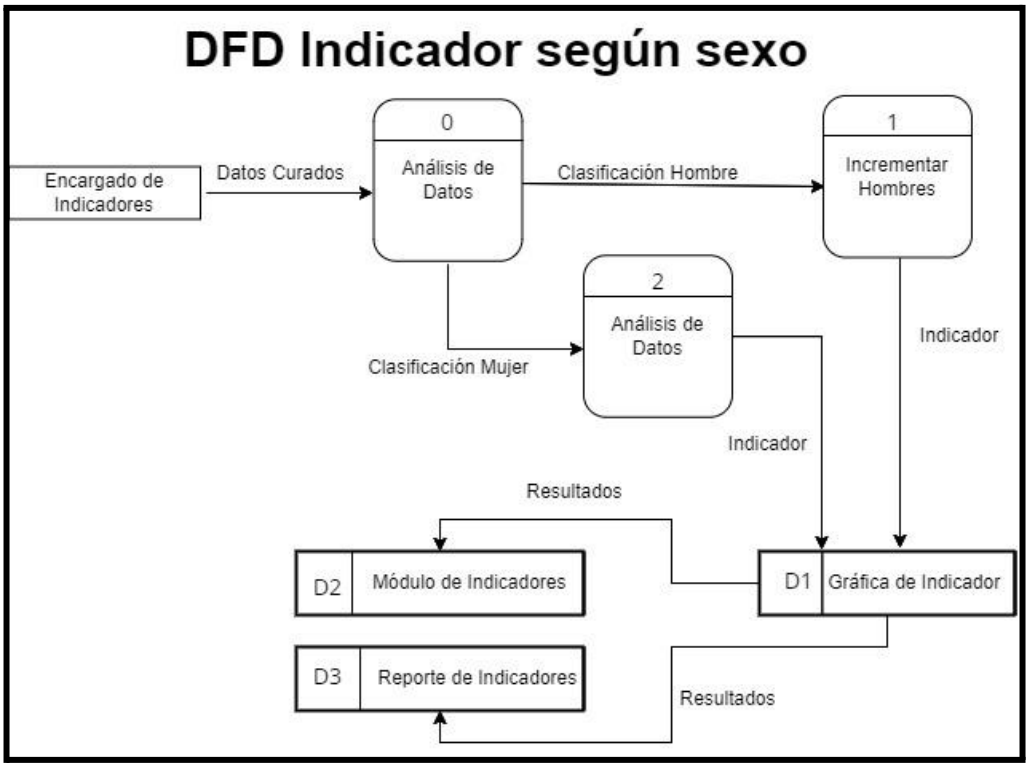
Fecha	Hora	Requests	Requests por segundo
2023-02-27	00:00:00	40	0.0111
2023-05-23	01:00:00	18	0.005
2023-03-15	02:00:00	26	0.0072
2023-03-07	03:00:00	5	0.0014
2023-03-08	04:00:00	3	0.0008
2023-05-19	05:00:00	6	0.0017
2023-05-11	06:00:00	9	0.0025
2023-03-17	07:00:00	17	0.0047
2023-02-27	08:00:00	62	0.0172
2023-02-27	09:00:00	179	0.0497
2023-03-17	10:00:00	91	0.0253
2023-02-27	11:00:00	84	0.0233
2023-02-27	12:00:00	74	0.0206
2023-05-12	13:00:00	39	0.0108
2023-03-14	14:00:00	57	0.0158
2023-03-15	15:00:00	88	0.0244
2023-03-17	16:00:00	55	0.0153
2023-03-22	17:00:00	82	0.0228
2023-03-14	18:00:00	28	0.0078
2023-03-01	19:00:00	39	0.0108
2023-03-20	20:00:00	47	0.0131
2023-02-27	21:00:00	63	0.0175
2023-05-14	22:00:00	36	0.0100
2023-02-22	23:00:00	29	0.0081

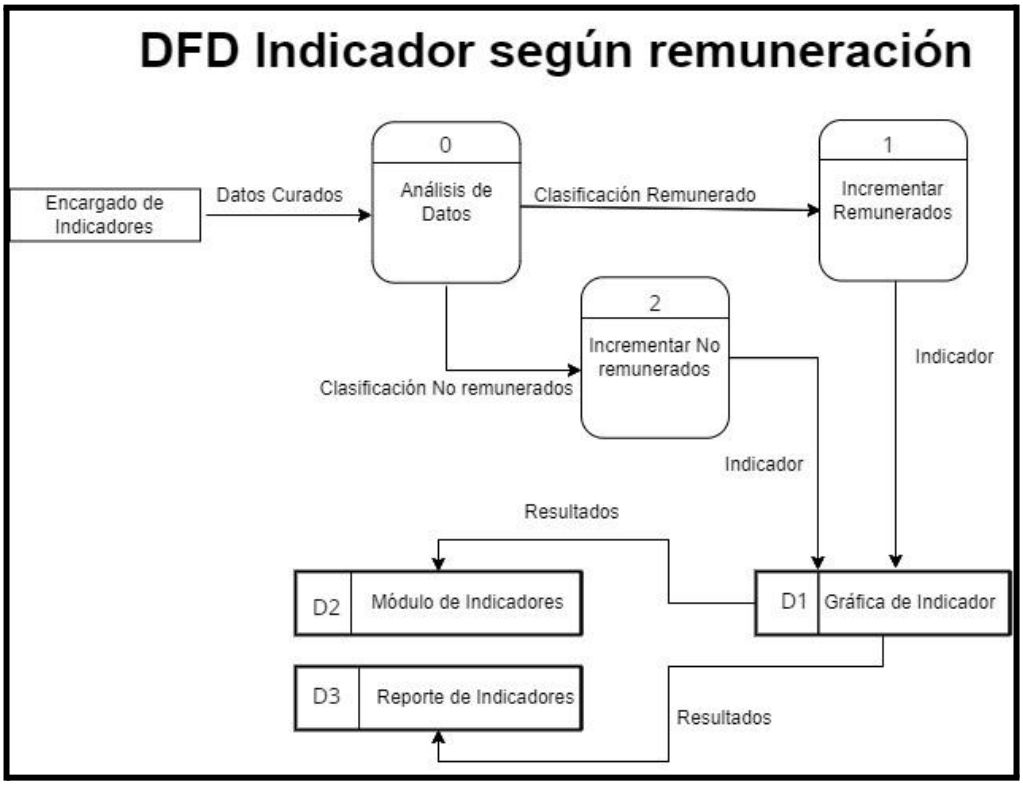
Fuente: Informática DIGI.

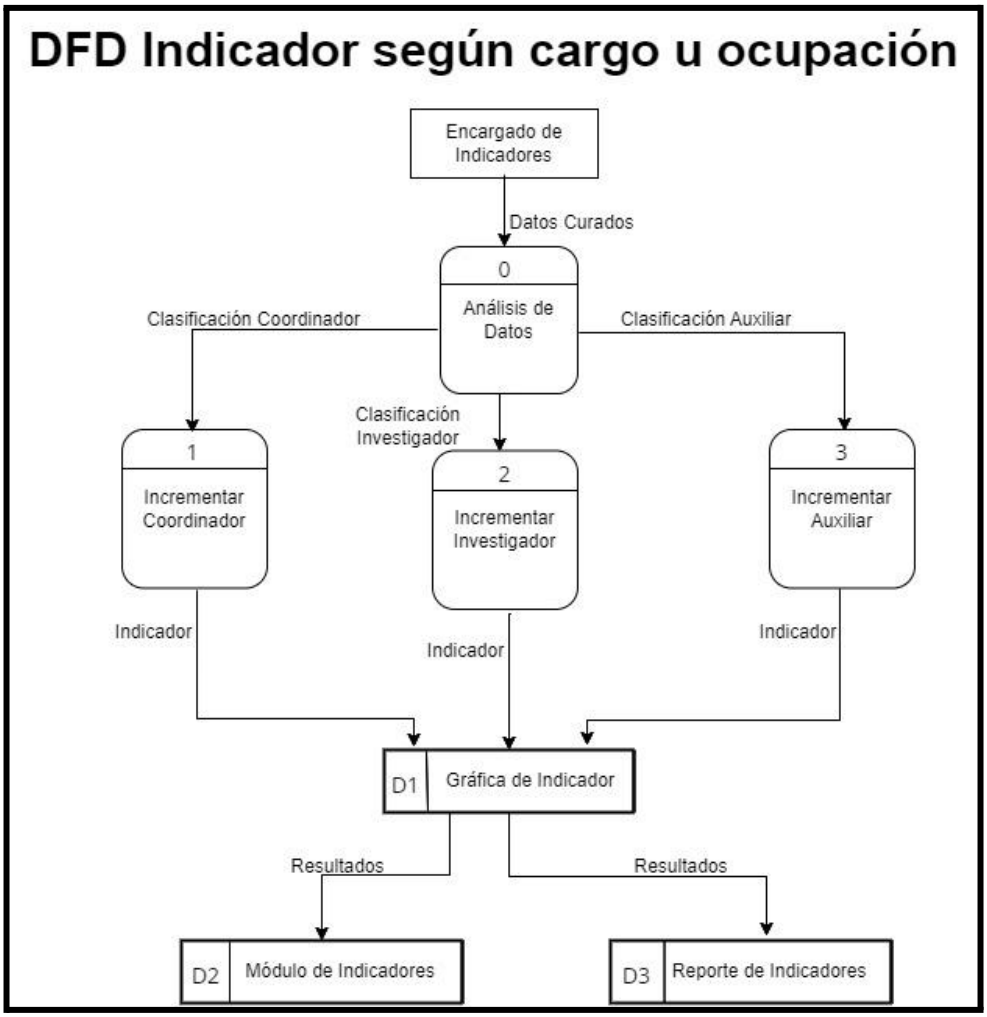
Apéndice 3. Diagramas Complementarios



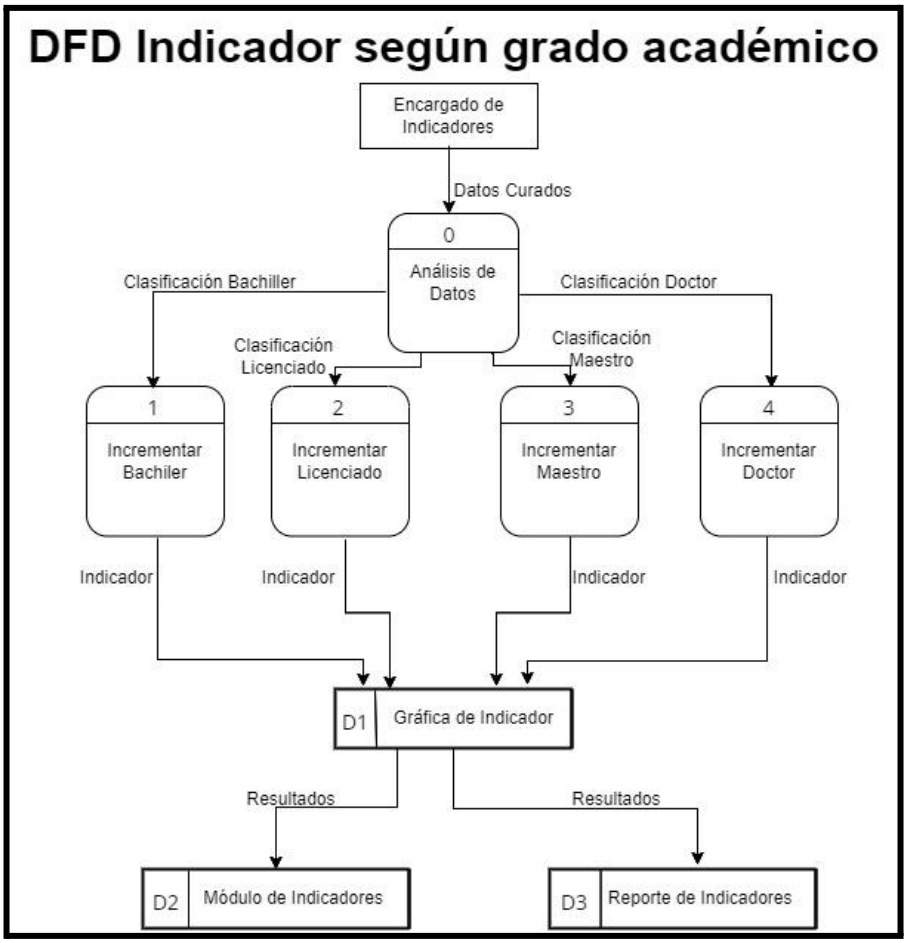


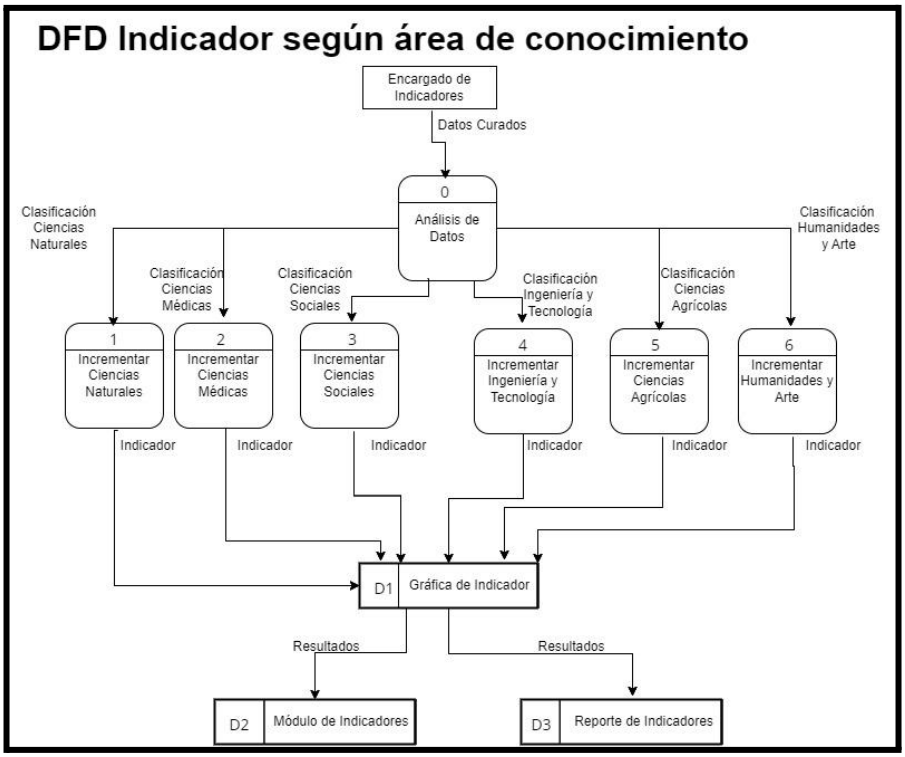

DFD Indicador según sexo
Encargado de Indicadores
Datos Curados
0
Análisis de Datos
Clasificación Hombre
1
Incrementar Hombres
Clasificación Mujer
2
Análisis de Datos
Indicador
Indicador
Resultados
D2
Módulo de Indicadores
D1
Gráfica de Indicador
D3
Reporte de Indicadores
Resultados

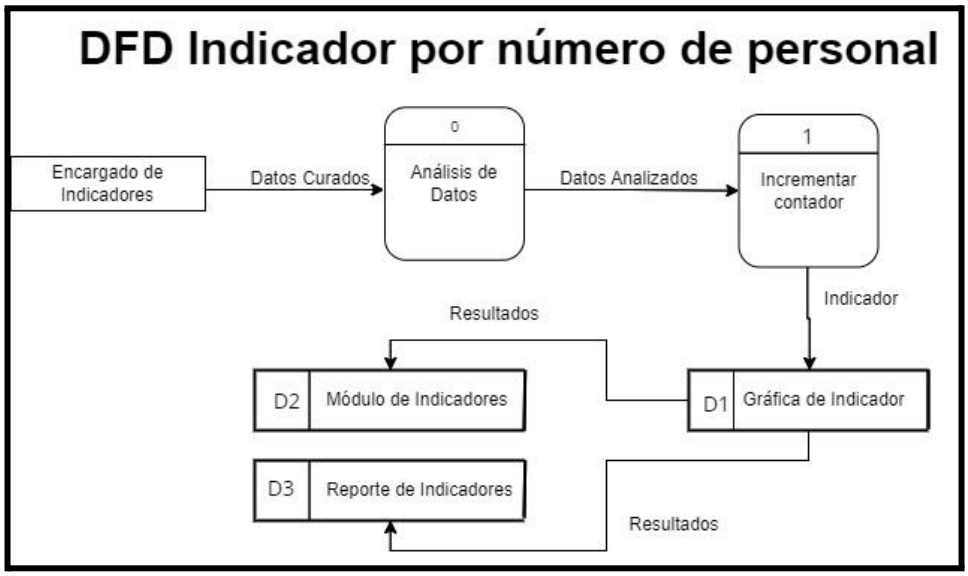

DFD Indicador según remuneración
Encargado de Indicadores
Datos Curados
0
Análisis de Datos
Clasificación Remunerado
1
Incrementar Remunerados
Clasificación No remunerados
2
Incrementar No remunerados
Indicador
Indicador
Resultados
D2
Módulo de Indicadores
D1
Gráfica de Indicador
D3
Reporte de Indicadores
Resultados


DFD Indicador según cargo u ocupación
Encargado de Indicadores
Datos Curados
0
Análisis de Datos
Clasificación Coordinador
Clasificación Auxiliar
Clasificación Investigador
1
Incrementar Coordinador
2
Incrementar Investigador
3
Incrementar Auxiliar
Indicador
Indicador
Indicador
D1
Gráfica de Indicador
Resultados
Resultados
D2
Módulo de Indicadores
D3
Reporte de Indicadores

Informe final de proyecto de investigación. Año 2023




DFD Indicador según área de conocimiento
Encargado de Indicadores
Datos Curados
0
Análisis de Datos
Clasificación Ciencias Naturales
Clasificación Ciencias Médicas
Clasificación Ciencias Sociales
Clasificación Ingeniería y Tecnología
Clasificación Ciencias Agrícolas
Clasificación Humanidades y Arte
1 Incrementar Ciencias Naturales
2 Incrementar Ciencias Médicas
3 Incrementar Ciencias Sociales
4 Incrementar Ingeniería y Tecnología
5 Incrementar Ciencias Agrícolas
6 Incrementar Humanidades y Arte
Indicador
D1 Gráfica de Indicador
Resultados
D2 Módulo de Indicadores
D3 Reporte de Indicadores


DFD Indicador por número de personal
Encargado de Indicadores
Datos Curados
0
Análisis de Datos
Datos Analizados
1
Incrementar contador
Indicador
Resultados
D2 Módulo de Indicadores
D1 Gráfica de Indicador
D3 Reporte de Indicadores
Resultados

Apéndice 4. Códigos de Muestra

Código Componente de Investigadores RIMS

```
<mat-sidenav-container class="sidenav-container">
  <mat-sidenav>
    <mat-toolbar>Menu</mat-toolbar>
    <mat-nav-list>
      <a mat-list-item [routerLink]="['./home']" routerLinkActive="active">Home</a>
      <a mat-list-item [routerLink]="['./documentos']" routerLinkActive="active">Documentos</a>

      <mat-expansion-panel class="small-expansion-panel">
        <mat-expansion-panel-header>
          Crear Documento
        </mat-expansion-panel-header>
        <a mat-list-item [routerLink]="['./atg']" routerLinkActive="active"
          matTooltip="Acta Aprobacion Trabajos Graducacion">Acta Aprobacion Trabajos
Graducacion</a>
        <a mat-list-item [routerLink]="['./ant']" routerLinkActive="active" matTooltip="Acta
Nombramiento Terna">Acta
          Nombramiento Terna</a>
        <a mat-list-item [routerLink]="['./documentos', 'subdivision3']" routerLinkActive="active"
          matTooltip="Carta Asesor Trabajo Graduacion">Carta Asesor Trabajo Graduacion</a>
        <a mat-list-item [routerLink]="['./documentos', 'subdivision4']" routerLinkActive="active"
          matTooltip="Circular">Circular</a>
        <a mat-list-item [routerLink]="['./documentos', 'subdivision5']" routerLinkActive="active"
          matTooltip="Correspondencia">Correspondencia</a>
        <a mat-list-item [routerLink]="['./documentos', 'subdivision5']" routerLinkActive="active"
          matTooltip="Finiquito EPS">Finiquito EPS</a>
        <a mat-list-item [routerLink]="['./documentos', 'subdivision5']" routerLinkActive="active"
          matTooltip="Memorandum">Memorandum</a>
        <a mat-list-item [routerLink]="['./documentos', 'subdivision5']" routerLinkActive="active"
          matTooltip="Minuta">Minuta</a>
        <a mat-list-item [routerLink]="['./documentos', 'subdivision5']" routerLinkActive="active"
          matTooltip="Providencia">Providencia</a>
        <a mat-list-item [routerLink]="['./documentos', 'subdivision5']" routerLinkActive="active"
          matTooltip="Solvencia Carrera">Solvencia Carrera</a>
        <a mat-list-item [routerLink]="['./documentos', 'subdivision5']" routerLinkActive="active"
```

```
      matTooltip="Solicitud EPS">Solicitud EPS</a>
    <a mat-list-item [routerLink]="['./documentos', 'subdivision5']" routerLinkActive="active"
      matTooltip="Finiquito EPS (Sin punteo)">Finiquito EPS (Sin punteo)</a>
    <a mat-list-item [routerLink]="['./documentos', 'subdivision5']" routerLinkActive="active"
      matTooltip="Resolución EPS">Resolución EPS</a>
    <a mat-list-item [routerLink]="['./documentos', 'subdivision5']" routerLinkActive="active"
      matTooltip="Nombramiento Terna">Nombramiento Terna</a>
    <a mat-list-item [routerLink]="['./documentos', 'subdivision5']" routerLinkActive="active"
      matTooltip="Solicitud Asesor EPS">Solicitud Asesor EPS</a>
    <a mat-list-item [routerLink]="['./documentos', 'subdivision5']" routerLinkActive="active"
      matTooltip="Acta Examen Privado">Acta Examen Privado</a>
    <a mat-list-item [routerLink]="['./documentos', 'subdivision5']" routerLinkActive="active"
      matTooltip="Carta Solicitud Trabajos Graduación">Carta Solicitud Trabajos Graduación</a>
  </mat-expansion-panel>

  <a mat-list-item [routerLink]="['./personas']" routerLinkActive="active">Personal</a>

  <a mat-list-item [routerLink]="['./perfil']" routerLinkActive="active">Perfil</a>

 </mat-nav-list>
</mat-sidenav>
<mat-sidenav-content>
 <mat-toolbar color="primary">
  <button type="button" aria-label="Toggle sidenav" mat-icon-button>
    <mat-icon aria-label="Side nav toggle icon">menu</mat-icon>
  </button>


  <div class="user-info">
    <span class="user-name"> Kevin Lajpop </span>
    <mat-icon class="user-icon">person</mat-icon>

  </div>


 </mat-toolbar>
 <router-outlet></router-outlet>
</mat-sidenav-content>
```

```
</mat-sidenav-container>
```

Código de Componente de Usuario RIMS

```
<form [formGroup]="addressForm" novalidate (ngSubmit)="onSubmit()">
  <mat-card class="shipping-card">
    <mat-card-header>
      <mat-card-title>Shipping Information</mat-card-title>
    </mat-card-header>
    <mat-card-content>
      <div class="row">
        <div class="col">
          <mat-form-field class="full-width">
            <input matInput placeholder="Company" formControlName="company">
          </mat-form-field>
        </div>
      </div>
      <div class="row">
        <div class="col">
          <mat-form-field class="full-width">
            <input matInput placeholder="First name" formControlName="firstName">
            <mat-error *ngIf="addressForm.controls['firstName'].hasError('required')">
              First name is <strong>required</strong>
            </mat-error>
          </mat-form-field>
        </div>
        <div class="col">
          <mat-form-field class="full-width">
            <input matInput placeholder="Last name" formControlName="lastName">
            <mat-error *ngIf="addressForm.controls['lastName'].hasError('required')">
              Last name is <strong>required</strong>
            </mat-error>
          </mat-form-field>
        </div>
      </div>
      <div class="row">
```

```
  <div class="col">
    <mat-form-field class="full-width">
      <textarea matInput placeholder="Address" formControlName="address"></textarea>
      <mat-error *ngIf="addressForm.controls['address'].hasError('required')">
        Address is <strong>required</strong>
      </mat-error>
    </mat-form-field>
  </div>
</div>
<div class="row" *ngIf="!hasUnitNumber">
  <div class="col">
    <button mat-button type="button" (click)="hasUnitNumber = !hasUnitNumber">
      + Add C/O, Apt, Suite, Unit
    </button>
  </div>
</div>
<div class="row" *ngIf="hasUnitNumber">
  <div class="col">
    <mat-form-field class="full-width">
      <textarea matInput placeholder="Address 2" formControlName="address2"></textarea>
    </mat-form-field>
  </div>
</div>
<div class="row">
  <div class="col">
    <mat-form-field class="full-width">
      <input matInput placeholder="City" formControlName="city">
      <mat-error *ngIf="addressForm.controls['city'].hasError('required')">
        City is <strong>required</strong>
      </mat-error>
    </mat-form-field>
  </div>
  <div class="col">
    <mat-form-field class="full-width">
      <mat-select placeholder="State" formControlName="state">
        <mat-option *ngFor="let state of states" [value]="state.abbreviation">
          {{ state.name }}
        </mat-option>
```

```
          </mat-select>
          <mat-error *ngIf="addressForm.controls['state'].hasError('required')">
            State is <strong>required</strong>
          </mat-error>
        </mat-form-field>
      </div>
    </div>
    <div class="row">
      <div class="col">
        <mat-form-field class="full-width">
          <input matInput #postalCode maxlength="5" placeholder="Postal Code" type="number"
formControlName="postalCode">
          <mat-hint align="end">{{postalCode.value.length}} / 5</mat-hint>
        </mat-form-field>
      </div>
    </div>
    <div class="row">
      <div class="col">
        <mat-radio-group formControlName="shipping">
          <mat-radio-button value="free">Free Shipping</mat-radio-button>
          <mat-radio-button value="priority">Priority Shipping</mat-radio-button>
          <mat-radio-button value="nextday">Next Day Shipping</mat-radio-button>
        </mat-radio-group>
      </div>
    </div>
  </mat-card-content>
  <mat-card-actions>
    <button mat-raised-button color="primary" type="submit">Submit</button>
  </mat-card-actions>
 </mat-card>
</form>
```

```
import { Component, inject } from '@angular/core';
```

```
import { FormBuilder, Validators } from '@angular/forms';


@Component({
  selector: 'app-user',
  templateUrl: './user.component.html',
  styleUrls: ['./user.component.scss']
})
export class UserComponent {
  private fb = inject(FormBuilder);
  addressForm = this.fb.group({
    company: null,
    firstName: [null, Validators.required],
    lastName: [null, Validators.required],
    address: [null, Validators.required],
    address2: null,
    city: [null, Validators.required],
    state: [null, Validators.required],
    postalCode: [null, Validators.compose([
      Validators.required, Validators.minLength(5), Validators.maxLength(5)])
    ],
    shipping: ['free', Validators.required]
  });

  hasUnitNumber = false;

  states = [
    {name: 'Alabama', abbreviation: 'AL'},
    {name: 'Alaska', abbreviation: 'AK'},
    {name: 'American Samoa', abbreviation: 'AS'},
    {name: 'Arizona', abbreviation: 'AZ'},
    {name: 'Arkansas', abbreviation: 'AR'},
    {name: 'California', abbreviation: 'CA'},
    {name: 'Colorado', abbreviation: 'CO'},
    {name: 'Connecticut', abbreviation: 'CT'},
    {name: 'Delaware', abbreviation: 'DE'},
    {name: 'District Of Columbia', abbreviation: 'DC'},
    {name: 'Federated States Of Micronesia', abbreviation: 'FM'},
```

{name: 'Florida', abbreviation: 'FL'},
{name: 'Georgia', abbreviation: 'GA'},
{name: 'Guam', abbreviation: 'GU'},
{name: 'Hawaii', abbreviation: 'HI'},
{name: 'Idaho', abbreviation: 'ID'},
{name: 'Illinois', abbreviation: 'IL'},
{name: 'Indiana', abbreviation: 'IN'},
{name: 'Iowa', abbreviation: 'IA'},
{name: 'Kansas', abbreviation: 'KS'},
{name: 'Kentucky', abbreviation: 'KY'},
{name: 'Louisiana', abbreviation: 'LA'},
{name: 'Maine', abbreviation: 'ME'},
{name: 'Marshall Islands', abbreviation: 'MH'},
{name: 'Maryland', abbreviation: 'MD'},
{name: 'Massachusetts', abbreviation: 'MA'},
{name: 'Michigan', abbreviation: 'MI'},
{name: 'Minnesota', abbreviation: 'MN'},
{name: 'Mississippi', abbreviation: 'MS'},
{name: 'Missouri', abbreviation: 'MO'},
{name: 'Montana', abbreviation: 'MT'},
{name: 'Nebraska', abbreviation: 'NE'},
{name: 'Nevada', abbreviation: 'NV'},
{name: 'New Hampshire', abbreviation: 'NH'},
{name: 'New Jersey', abbreviation: 'NJ'},
{name: 'New Mexico', abbreviation: 'NM'},
{name: 'New York', abbreviation: 'NY'},
{name: 'North Carolina', abbreviation: 'NC'},
{name: 'North Dakota', abbreviation: 'ND'},
{name: 'Northern Mariana Islands', abbreviation: 'MP'},
{name: 'Ohio', abbreviation: 'OH'},
{name: 'Oklahoma', abbreviation: 'OK'},
{name: 'Oregon', abbreviation: 'OR'},
{name: 'Palau', abbreviation: 'PW'},
{name: 'Pennsylvania', abbreviation: 'PA'},
{name: 'Puerto Rico', abbreviation: 'PR'},
{name: 'Rhode Island', abbreviation: 'RI'},
{name: 'South Carolina', abbreviation: 'SC'},
{name: 'South Dakota', abbreviation: 'SD'},

```
    {name: 'Tennessee', abbreviation: 'TN'},
    {name: 'Texas', abbreviation: 'TX'},
    {name: 'Utah', abbreviation: 'UT'},
    {name: 'Vermont', abbreviation: 'VT'},
    {name: 'Virgin Islands', abbreviation: 'VI'},
    {name: 'Virginia', abbreviation: 'VA'},
    {name: 'Washington', abbreviation: 'WA'},
    {name: 'West Virginia', abbreviation: 'WV'},
    {name: 'Wisconsin', abbreviation: 'WI'},
    {name: 'Wyoming', abbreviation: 'WY'}
  ];

  onSubmit(): void {
    alert('Thanks!');
  }
}
```

Apéndice 5. Formularios

Formulario de Indicadores

DIRECCIÓN GENERAL DE INVESTIGACIÓN

INSTITUTO DE INVESTIGACIÓN DEL CENTRO UNIVERSITARIO DE ORIENTE

Sección 1 de 6

Proyecto: Diseño de un Sistema de Información de Investigación Actual (CRIS)

Estimado profesional:

La Dirección General de Investigación (DIGI), a través de la convocatoria diferenciada del año 2023, ha emprendido un proyecto para diseñar un Sistema de Información de Investigación Actual (CRIS). La investigación surgió por la identificación de la necesidad de facilitar el acceso a la información generada por las investigaciones de la Universidad de San Carlos de Guatemala. Se espera que el sistema favorecerá la visibilización de investigadores y resultados.

En el desarrollo metodológico de la investigación, se planteó realizar entrevistas a distintos actores clave en los procesos de investigación, con el objetivo de obtener, confirmar y definir requerimientos que orienten el diseño del sistema. Con esto se logrará considerar las condiciones y necesidades actuales del contexto de investigación de la USAC.

Usted ha sido seleccionado para participar en este proceso, aportando su visión, experiencia y conocimientos en la materia. Para ello, le agradecemos que conteste las preguntas de este cuestionario. **Todas las preguntas que se plantean giran en torno a la generación de indicadores de investigación.** Por favor considere extenderse abarcando todos los detalles que considere pertinentes.

Tiempo estimado para responder el formulario: 50 minutos.

Informe final de proyecto de investigación. Año 2023

Sección 2 de 6

Datos personales

Descripción (opcional)

Correo electrónico *

Texto de respuesta breve

Nombre y apellidos *

Texto de respuesta breve

Después de la sección 2 Ir a la siguiente sección

Sección 3 de 6

En el documento: "Indicadores del Sistema de Investigación de la Universidad de San Carlos de Guatemala, 2019-2021", se menciona que "para la recopilación de los datos se recurrió a encuestas en línea y a la información proporcionada por la Tesorería del Fondo de Investigación de la DIGI". Para nuestro mejor entendimiento de los flujos de información en la DIGI, por favor explíquenos lo requerido en las siguientes preguntas.

Descripción (opcional)

Informe final de proyecto de investigación. Año 2023

Tesorería del Fondo de Investigación de la DIGI

Descripción (opcional)

1. ¿Cuál es la función de la Tesorería del Fondo de Investigación de la DIGI? *

Texto de respuesta largo

2. En términos generales ¿cuáles fueron los datos que obtuvo de la Tesorería del Fondo de Investigación? *

Texto de respuesta largo

3. ¿Quiénes son los actores que proveen de esos datos a la instancia en mención? *

Texto de respuesta largo

4. ¿En qué momentos son ingresados y actualizados los datos que maneja la Tesorería del Fondo de Investigación de la DIGI? *

Texto de respuesta largo

5. ¿En qué formato le fueron otorgados los datos? *

Texto de respuesta largo

6. ¿Necesitó realizar limpieza y reorganización de los datos para poderlos utilizar en la generación de los indicadores (corregir errores, inconsistencias, duplicados, formato incorrecto)? *

- Sí
- No

7. Si su respuesta a la pregunta anterior fue "sí", por favor indique, dada su experiencia, ¿cuáles recomendaciones considera pertinentes para preparar los datos de manera efectiva?

Texto de respuesta largo

Después de la sección 3 Ir a la siguiente sección

Sección 4 de 6

Encuestas

Descripción (opcional)

8. De manera general, ¿cuáles fueron los datos que obtuvo de las encuestas? *

Texto de respuesta largo

9. Según su criterio, ¿cuál sería la manera más efectiva de mantener un registro actualizado de los datos referidos en su respuesta anterior? *

Texto de respuesta largo

10. ¿Cuáles serían los perfiles adecuados para registrar y actualizar los datos en cuestión? *

Texto de respuesta largo

11. ¿ En qué formato se encuentran almacenados los datos recolectados en las encuestas ? *

Texto de respuesta largo

12. ¿Considera que nuestro equipo puede tener acceso a esos datos con el propósito de generar un diseño de base de datos apropiado, que favorezca el cálculo de indicadores de investigación en el sistema CRIS? *

- Si
- No

13. Si su respuesta a la pregunta anterior fue "Sí", por favor indique ¿cómo debemos proceder para acceder a los datos referidos?

Texto de respuesta largo

14. Si su respuesta a la pregunta 12 fue "No", ¿qué alternativa propone para que nuestro equipo de investigación tenga un panorama completo de los datos recopilados en las encuestas y la forma en que se encuentran estructurados? (Manuales o documentos creados para describir con detalle el conjunto de datos)

Texto de respuesta largo

Después de la sección 4 Ir a la siguiente sección

Sección 5 de 6

Generalidades

Descripción (opcional)

15. ¿Es posible que una investigación se registre como orientada a la consecución de dos o más objetivos socioeconómicos? *

- Si
- No

15. ¿Es posible que una investigación se registre como orientada a la consecución de dos o más objetivos socioeconómicos? *

- Si
- No

16. En su experiencia, ¿encontró proyectos que pertenecen a más de una red de investigación? *

- Si
- No

Después de la sección 5 Ir a la siguiente sección

Sección 6 de 6

Fin del cuestionario

Agradecemos y valoramos su contribución y el tiempo que utilizó para contestar este cuestionario. Sus criterios y perspectiva serán útiles en el diseño del sistema y promoverán los avances en el eje de investigación de la USAC.

Cuando se analicen sus respuestas, probablemente necesitaremos profundizar en algunos detalles. De ser así, volveremos a contactarle para programar una réplica.

Nuevamente, ¡gracias!

Informe final de proyecto de investigación. Año 2023

Formulario Personal de Informática

DIRECCIÓN GENERAL DE INVESTIGACIÓN

INSTITUTO DE INVESTIGACIÓN DEL CENTRO UNIVERSITARIO DE ORIENTE

Sección 1 de 5

Proyecto: Diseño de un Sistema de Información de Investigación Actual (CRIS)

Estimado profesional:

La Dirección General de Investigación (DIGI), a través de la convocatoria diferenciada del año 2023, ha emprendido un proyecto para diseñar un Sistema de Información de Investigación Actual (CRIS). La investigación surgió por la identificación de la necesidad de facilitar el acceso a la información generada por las investigaciones de la Universidad de San Carlos de Guatemala. Se espera que el sistema favorecerá la visibilización de investigadores y resultados.

En el desarrollo metodológico de la investigación, se planteó realizar entrevistas a distintos actores clave en los procesos de investigación, con el objetivo de obtener, confirmar y definir requerimientos que orienten el diseño del sistema. Con esto se logrará considerar las condiciones y necesidades actuales del contexto de investigación de la USAC.

Usted ha sido seleccionado para participar en este proceso, aportando su visión, experiencia y conocimientos en la materia. Para ello, le agradecemos que conteste las preguntas de este cuestionario. **Todas las preguntas que se plantean giran en torno a la gestión de la información de investigadores e investigaciones y la generación de indicadores de investigación.** Por favor considere extenderse abarcando todos los detalles que considere pertinentes.

Tiempo estimado para responder el formulario: 50 minutos.

Después de la sección 1 Ir a la siguiente sección

Sección 2 de 5

Datos personales

Descripción (opcional)

Correo electrónico *

Texto de respuesta breve

Nombre y apellidos *

Texto de respuesta breve

Sección 3 de 5

Obtención de datos de otros sistemas informáticos

Descripción (opcional)

1. ¿Con cuáles sistemas informáticos existentes en la Universidad de San Carlos, considera que debe conectarse el CRIS para obtener datos? *

Texto de respuesta largo

2. ¿Cuáles de los sistemas mencionados en la respuesta anterior, cuentan actualmente con API´s para comunicar información a otros sistemas? *

Texto de respuesta largo

3. ¿Conoce de la existencia de procesos administrativos que se requieran para que los administradores de esos sistemas otorguen permisos para consumir sus servicios? *

- Sí
- No

4. Si su respuesta a la pregunta anterior fue "Sí", por favor describa los pasos de los procedimientos administrativos, requisitos y tiempo estimado de la gestión.

Texto de respuesta largo

5. ¿Cuántos de los sistemas mencionados cuentan con documentación de su API respectiva con la finalidad de facilitar el trabajo de integración? *

- Todos
- Algunos
- Ninguno

Si la respuesta anterior fue distinta a "Ninguno", por favor conteste los ítems 6-8

Descripción (opcional)

6. Indique el nombre de los sistemas o módulos que sí cuentan con documentación de su API

Texto de respuesta largo

7. ¿Considera que nuestro equipo de investigación podría tener acceso a la documentación mencionada?

- ◯ Sí
- ◯ No

8. Si su respuesta a la pregunta anterior fue "Sí", por favor explique las gestiones correspondientes que nuestro equipo de investigación debe realizar para acceder a la documentación.

Texto de respuesta largo

9. De manera general, ¿cuáles serán los datos que se pueden obtener de la conexión del sistema CRIS con el SIIF? *

Texto de respuesta largo

10. ¿Con qué periodicidad sugiere que debería actualizarse la base de datos del sistema CRIS con datos que provengan de otros módulos informáticos de la USAC? *

Texto de respuesta largo

Sección 4 de 5

Generalidades del diseño del sistema

Descripción (opcional)

11. ¿Cuáles son los requisitos de seguridad y privacidad que deben ser abordados en el sistema CRIS, tales como: el acceso seguro a la información, la protección de datos sensibles y el cumplimiento de las regulaciones pertinentes? *

Texto de respuesta largo

12. El protocolo de nuestra investigación propone usar datos sobre el tráfico y visitas a los sistemas informáticos relacionados con las investigaciones realizadas en la USAC. Lo anterior, con la finalidad de estimar la demanda que podría tener el CRIS e incluir en el diseño, las características recomendadas para la plataforma en la que sea desplegado. ¿Considera factible que obtengamos dichos datos? *

- Sí
- No

13. Si su respuesta a la pregunta anterior fue "Sí", por favor indique ¿cómo debemos proceder para obtener los datos en cuestión?

Texto de respuesta largo

14. ¿Cuáles medidas han resultado exitosas para proteger los datos sensibles de los investigadores e investigaciones, almacenados en los sistemas que administra actualmente la DIGI? *

Texto de respuesta largo

Después de la sección 4 Ir a la siguiente sección

Sección 5 de 5

Fin del cuestionario

Agradecemos y valoramos su contribución y el tiempo que utilizó para contestar este cuestionario. Sus criterios y perspectiva serán útiles en el diseño del sistema y promoverán los avances en el eje de investigación de la USAC.

Cuando se analicen sus respuestas, probablemente necesitaremos profundizar en algunos detalles. De ser así, volveremos a contactarle para programar una réplica.

Nuevamente, ¡gracias!

Informe final de proyecto de investigación. Año 2023

Formulario del Personal Administrativo

Usted ha sido seleccionado para participar en este proceso, aportando su visión, experiencia y conocimientos en la materia. Para ello, le agradecemos que conteste las preguntas de este cuestionario. **Todas las preguntas que se plantean giran en torno a la gestión de la información de investigadores e investigaciones y la generación de indicadores de investigación**. Por favor considere extenderse abarcando todos los detalles que considere pertinentes.

Después de la sección 1 Ir a la siguiente sección

Sección 2 de 8

Datos personales

Descripción (opcional)

Correo electrónico *

Texto de respuesta breve

Nombre y apellidos *

Texto de respuesta breve

Informe final de proyecto de investigación. Año 2023

Sección 3 de 8

Generalidades

Descripción (opcional)

1. Respecto a la visibilidad de los investigadores y las investigaciones ¿Cuáles son los principales desafíos específicos que enfrenta la Dirección General de Investigación de la USAC en la actualidad? *

Texto de respuesta largo

2. ¿Cómo se espera que el sistema CRIS contribuya a superar los desafíos indicados en la respuesta a la pregunta anterior? *

Texto de respuesta largo

3. Por favor, indique si la siguiente afirmación es correcta: *

Los tipos de investigación que se desarrollan en la DIGI son:

Los tipos de investigación que se desarrollan en la DIGI son:

Investigación básica.
Investigación aplicada.
Desarrollo experimental.

- ◯ Sí
- ◯ No

4. Por favor, confirme si el siguiente listado incluye las opciones de tema regional de investigación en las que se puede enmarcar un proyecto: *

1. Democracia, gobernabilidad y seguridad ciudadana.
2. Población, educación, inclusión social e interculturalidad.
3. Territorialidad, riesgo y desarrollo local.
4. Salud pública, enfermedades crónicas no transmisibles e infecciosas.
5. Tecnologías, innovación y energías renovables.
6. Desarrollo y seguridad alimentaria y nutricional.
7. Cambio climático y gestión ambiental.
8. Transformación digital, automatización, inteligencia artificial y ciencia de datos.
9. Crecimiento económico, productividad, competitividad y desarrollo sostenible.

- ◯ Sí
- ◯ No

5. Por favor, confirme si los elementos de la siguiente lista, son los objetivos socioeconómicos (SOE) que se asocian con las investigaciones en la actualidad: *

1. Exploración y explotación de la Tierra.
Medioambiente.
2. Exploración y explotación del espacio.
3. Transporte, telecomunicación y otras infraestructuras.
4. Energía.
5. Producción y tecnología industrial.
6. Sanidad.
7. Agricultura.
8. Educación.
9. Cultura, ocio, religión y medios de comunicación.
10. Sistemas, estructuras y procesos políticos y sociales.
11. Avance general del conocimiento: I+D financiada con los fondos generales de las universidades (FGU).
12. Avance general del conocimiento: I+D financiada por otras fuentes.
13. Defensa.

- ◯ Sí
- ◯ No

6. ¿Existe alguna relación de dependencia entre las categorías de clasificación mencionadas en las preguntas 3-5? (ejemplo: la investigación básica solamente se desarrolla en las áreas de ciencias naturales y humanidades) *

- ◯ Sí
- ◯ No

7. Si su respuesta a la pregunta anterior fue "sí", explique las dependencias existentes entre las categorías de clasificación:

Texto de respuesta largo

Sección 4 de 8

Elementos específicos del sistema

Visto desde la perspectiva del acceso, un sistema CRIS tiene dos secciones: pública y administrativa. Diferentes tipos de usuario pueden interactuar con el sistema CRIS. Algunos tienen acceso solamente a consultar información, otros pueden agregar y modificar datos.

El esquema que se muestra en la parte inferior representa un ejemplo de los posibles tipos de usuario, acciones y el flujo de los datos.

- Un **usuario visitante** es el que solamente podrá consultar la información que se ha hecho pública, por ejemplo: el perfil de los investigadores, resultados de los proyectos, indicadores, entre otros.
- La **sección pública** es la que permite al sistema CRIS, cumplir con el objetivo de potenciar la visibilidad del trabajo de investigación y de los investigadores.
- Los tipos de **usuarios con credenciales** podrán agregar, modificar o confirmar datos (sección administrativa). Cada tipo de usuario tendrá permisos para efectuar cambios en ciertos datos (según se defina en los requerimientos). Por ejemplo: los proyectos de investigación podrán ser creados o actualizados por los coordinadores de proyectos o personal administrativo de la DIGI, no así por los investigadores. Lo anterior es un ejemplo de lo que se necesita consensuar.
- El esquema muestra tres ejemplos de lo que podrían ser tipos de usuarios con credenciales: administrativo DIGI, coordinador e investigador. Se pueden *definir tantos tipos de usuario como se considere pertinente*.
- Los **usuarios con credenciales** pueden acceder a las dos secciones: pública (solo consulta) y administrativa (alimentar la base de datos).
- La **sección administrativa** ofrece distintos formularios para los módulos de investigadores, proyectos, unidades de investigación, entre otros. Los formularios podrán ser accedidos por los usuarios que cuenten con los permisos correspondientes.
- La **sección administrativa** efectúa cambios a la base de datos para mantenerla actualizada. Pero también, la base de datos puede ser alimentada con fuentes externas como ORCID y Google Académico.

8. ¿Cuáles deben ser los tipos de usuario (director, investigador, administrador) y responsabilidades específicas de los mismos en el sistema CRIS? *

Texto de respuesta largo

9. ¿Qué acciones o tareas podrán realizar los tipos de usuario definidos en la respuesta a la pregunta anterior? *

Texto de respuesta largo

10. ¿Cuáles usuarios deberán tener acceso a la sección administrativa del sistema? (formularios para actualizar perfiles de investigadores, investigaciones y configuraciones del sistema) *

Texto de respuesta largo

11. ¿Qué beneficios o ventajas se esperan obtener al integrar los identificadores ORCID en el nuevo sistema CRIS? *

Texto de respuesta largo

12. ¿Qué beneficios se esperan obtener al integrar el sistema CRIS con el perfil de los investigadores de Google Académico? *

Texto de respuesta largo

Después de la sección 4 Ir a la siguiente sección

Sección 5 de 8

Procesos de gestión de los datos de investigadores e investigaciones

Descripción (opcional)

13. ¿Cuál es el proceso para el registro y actualización de investigadores? *

Texto de respuesta largo

14. ¿Cuáles tipos de usuario podrán participar en el registro y actualización de investigadores? *

Texto de respuesta largo

15. ¿Cuáles son los datos personales de los investigadores que pueden hacerse públicos en el CRIS? *

Texto de respuesta largo

16. ¿Cuáles datos personales de los investigadores son requeridos para su registro en el sistema, pero no deben hacerse públicos? *

Texto de respuesta largo

17. ¿Cuáles son los procesos de gestión de las investigaciones actuales que se desean automatizar o mejorar con el sistema CRIS? *

Texto de respuesta largo

18. Por favor, describa de manera general y ordenada, las acciones y datos que se requieren para ejecutar los procesos enumerados en la respuesta de la pregunta anterior. *

Texto de respuesta largo

Informe final de proyecto de investigación. Año 2023

Sección 6 de 8

Cumplimiento de regulaciones

Se abordan requisitos, regulaciones y estándares a los que el sistema CRIS debe acoplarse.

19. ¿Cuáles son los requisitos de seguridad y privacidad que deben ser abordados en el sistema CRIS, como el acceso seguro a la información, la protección de datos sensibles y el cumplimiento de las regulaciones pertinentes? *

Texto de respuesta largo

20. ¿Cuáles son las regulaciones o estándares, tales como normativas éticas o de propiedad intelectual, a los que el sistema CRIS debe adherirse? *

Texto de respuesta largo

21. ¿Cómo propone administrar la información y elementos sujetos a propiedad intelectual y derechos de autor de las investigaciones registradas en el sistema? *

Texto de respuesta largo

Sección 7 de 8

Conclusión

Descripción (opcional)

22. Según su criterio, ¿Cuáles podrán ser las principales barreras o desafíos que se espera enfrentar durante el diseño del sistema CRIS? *

Texto de respuesta largo

23. ¿De qué manera la DIGI proporcionaría ayuda para superar los desafíos mencionados en la respuesta de la pregunta anterior? *

Texto de respuesta largo

Sección 8 de 8

Fin del cuestionario

Agradecemos y valoramos su contribución y el tiempo que utilizó para contestar este cuestionario. Sus criterios y perspectiva serán útiles en el diseño del sistema y promoverán los avances en el eje de investigación de la USAC.

Cuando se analicen sus respuestas, probablemente necesitaremos profundizar en algunos detalles. De ser así, volveremos a contactarle para programar una réplica.

Nuevamente, ¡gracias!

Informe final de proyecto de investigación. Año 2023

Formulario Orientado a regulaciones y normativas

DIRECCIÓN GENERAL DE INVESTIGACIÓN

INSTITUTO DE INVESTIGACIÓN DEL CENTRO UNIVERSITARIO DE ORIENTE

Sección 1 de 6

Proyecto: Diseño de un Sistema de Información de Investigación Actual (CRIS)

Estimado profesional:

La Dirección General de Investigación (DIGI), a través de la convocatoria diferenciada del año 2023, ha emprendido un proyecto para diseñar un Sistema de Información de Investigación Actual (CRIS). La investigación surgió por la identificación de la necesidad de facilitar el acceso a la información generada por las investigaciones de la Universidad de San Carlos de Guatemala. Se espera que el sistema favorecerá la visibilización de investigadores y resultados.

En el desarrollo metodológico de la investigación, se planteó realizar entrevistas a distintos actores clave en los procesos de investigación, con el objetivo de obtener, confirmar y definir requerimientos que orienten el diseño del sistema. Con esto se logrará considerar las condiciones y necesidades actuales del contexto de investigación de la USAC.

Usted ha sido seleccionado para participar en este proceso, aportando su visión, experiencia y conocimientos en la materia. Para ello, le agradecemos que conteste las preguntas de este cuestionario. **Todas las preguntas que se plantean giran en torno a la gestión de la información de investigadores e investigaciones y la generación de indicadores de investigación**. Por favor considere extenderse abarcando todos los detalles que estime pertinentes.

Tiempo estimado para responder el formulario: 20 minutos.

Después de la sección 1 Ir a la siguiente sección

Sección 2 de 6

Datos personales

Descripción (opcional)

Correo electrónico *

Texto de respuesta breve

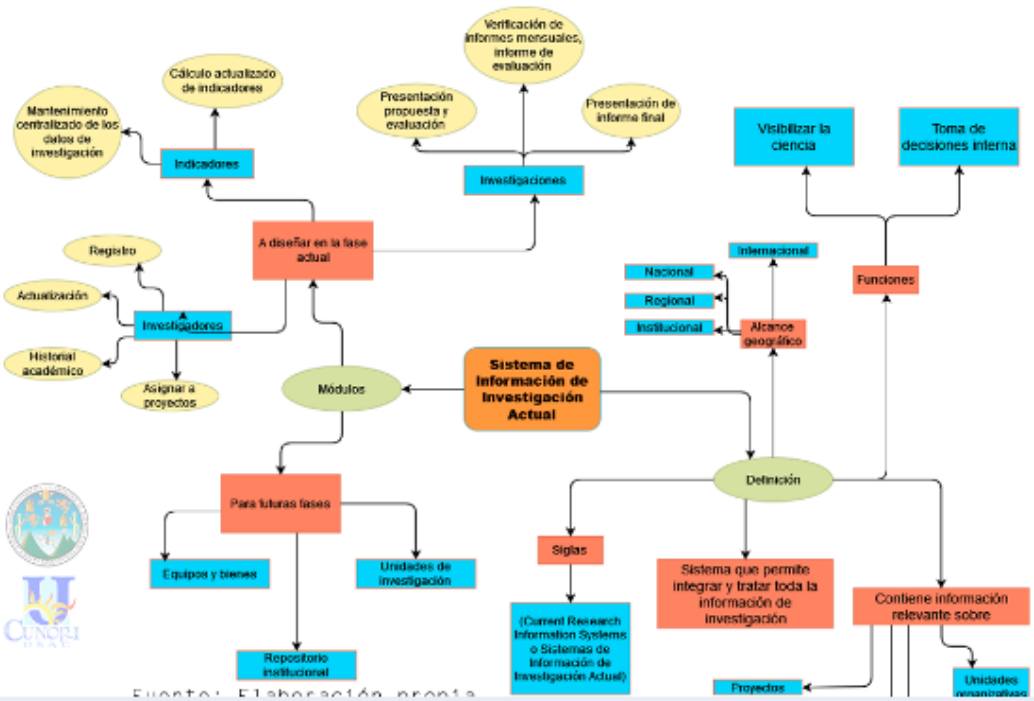
Informe final de proyecto de investigación. Año 2023

Sección 3 de 6

Definición de un sistema CRIS

Descripción (opcional)

Definición y módulos de un sistema CRIS



Fuente: Elaboración propia

Sección 4 de 6

Elementos específicos un sistema CRIS

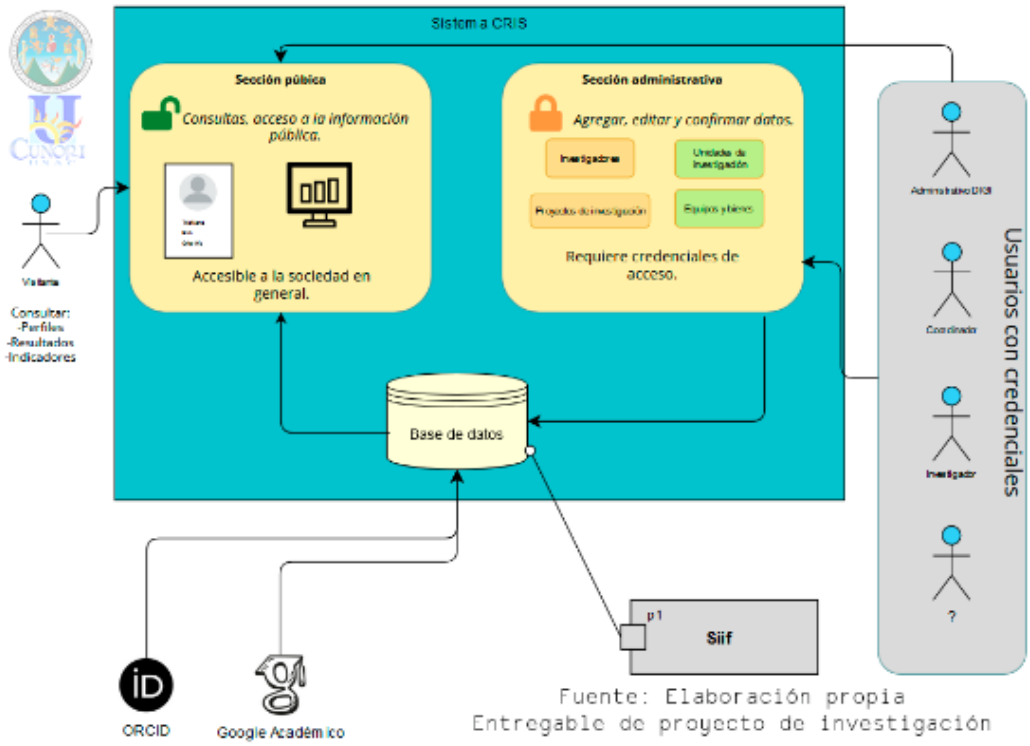
Visto desde la perspectiva del acceso, un sistema CRIS tiene dos secciones: pública y administrativa. Diferentes tipos de usuario pueden interactuar con el sistema CRIS. Algunos tienen acceso solamente a consultar información, otros pueden agregar y modificar datos.

El esquema que se muestra en la parte inferior representa un ejemplo de los posibles tipos de usuario, acciones y el flujo de los datos.

- Un **usuario visitante** es el que solamente podrá consultar la información que se ha hecho pública, por ejemplo: el perfil de los investigadores, resultados de los proyectos, indicadores, entre otros.
- La **sección pública** es la que permite al sistema CRIS, cumplir con el objetivo de potenciar la visibilidad del trabajo de investigación y de los investigadores.
- Los tipos de **usuarios con credenciales** podrán agregar, modificar o confirmar datos (sección administrativa). Cada tipo de usuario tendrá permisos para efectuar cambios en ciertos datos (según se defina en los requerimientos). Por ejemplo: los proyectos de investigación podrán ser creados o actualizados por los coordinadores de proyectos o personal administrativo de la DIGI, no así por los investigadores. Lo anterior es un ejemplo de lo que se necesita consensuar.
- El esquema muestra tres ejemplos de lo que podrían ser tipos de usuarios con credenciales: administrativo DIGI, coordinador e investigador. Se pueden *definir tantos tipos de usuario como se considere pertinente*.
- Los **usuarios con credenciales** pueden acceder a las dos secciones: pública (solo consulta) y administrativa (alimentar la base de datos).
- La **sección administrativa** ofrece distintos formularios para los módulos de investigadores, proyectos, unidades de investigación, entre otros. Los formularios podrán ser accedidos por los usuarios que cuenten con los permisos correspondientes.
- La **sección administrativa** efectúa cambios a la base de datos para mantenerla actualizada. Pero también, la base de datos puede ser alimentada con fuentes externas como ORCID y Google Académico.

Informe final de proyecto de investigación. Año 2023

Esquema ejemplo de usuarios y acciones en un sistema CRIS



Fuente: Elaboración propia
Entregable de proyecto de investigación

Sección 5 de 6

Cumplimiento de regulaciones

Se abordan regulaciones y estándares a los que el sistema CRIS debe acoplarse.

1. ¿Cuáles son las regulaciones o estándares, tales como normativas éticas o de propiedad intelectual, a los que el sistema CRIS debe adherirse? *

Texto de respuesta largo

2. ¿Cómo propone administrar la información y elementos sujetos a propiedad intelectual y derechos de autor de las investigaciones registradas en el sistema? *

Texto de respuesta largo

Sección 6 de 6

Fin del cuestionario

Agradecemos y valoramos su contribución y el tiempo que utilizó para contestar este cuestionario. Sus criterios y perspectiva serán útiles en el diseño del sistema y promoverán los avances en el eje de investigación de la USAC.

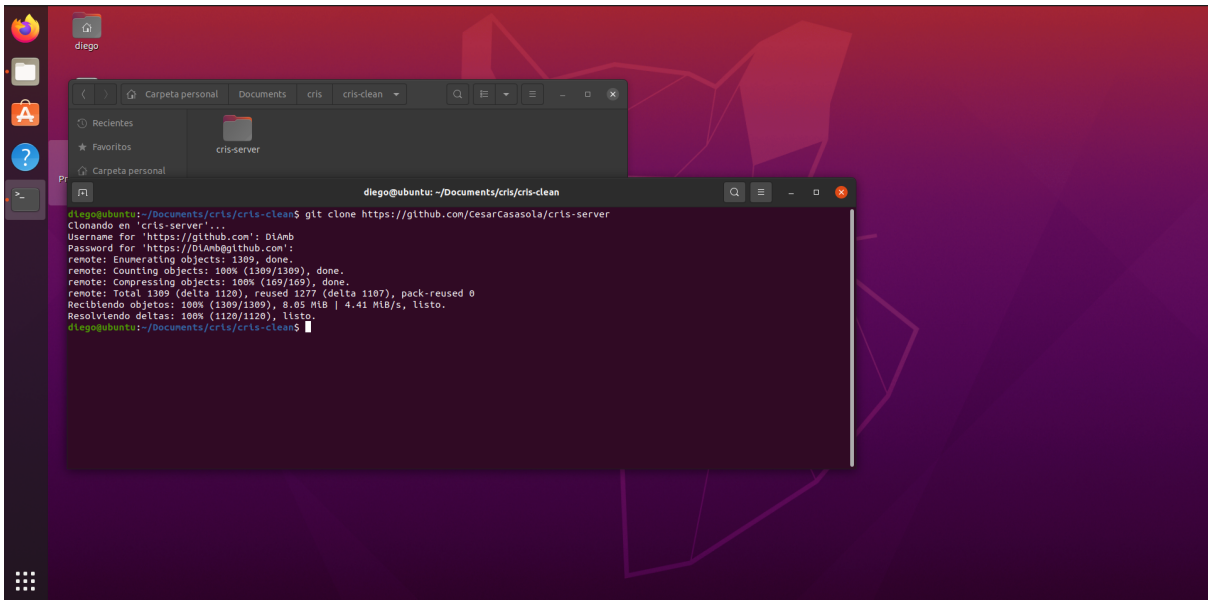
Cuando se analicen sus respuestas, probablemente necesitaremos profundizar en algunos detalles. De ser así, volveremos a contactarle para programar una réplica.

Nuevamente, ¡gracias!

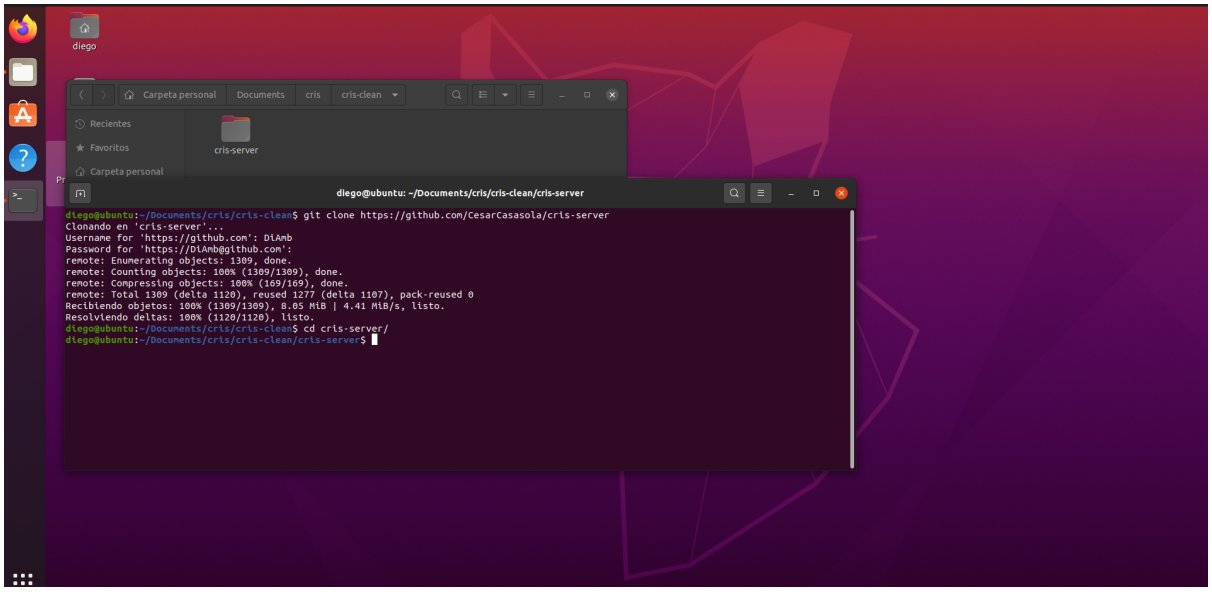
Apéndice 6. Guías de Instalación

Guia de instalación para Dspace-Cris 7

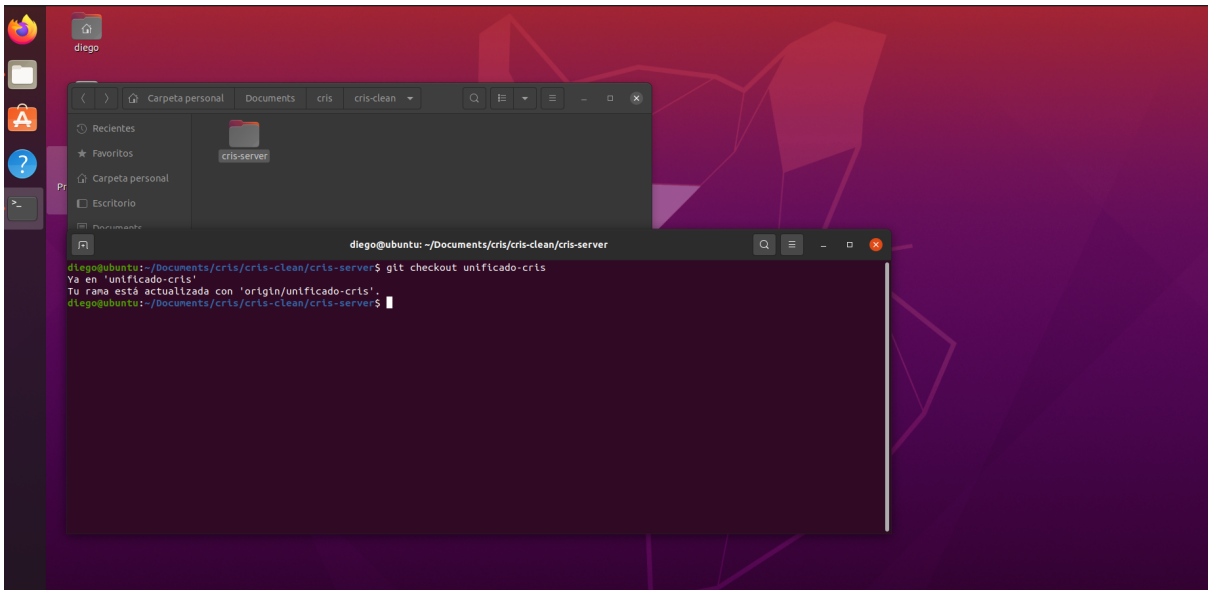
1. Asegúrate de tener Docker así como Docker-Compose instalado en tu sistema.
2. Clona el repositorio de trabajo, con la siguiente instrucción:
   git clone https://github.com/CesarCasasola/cris-server



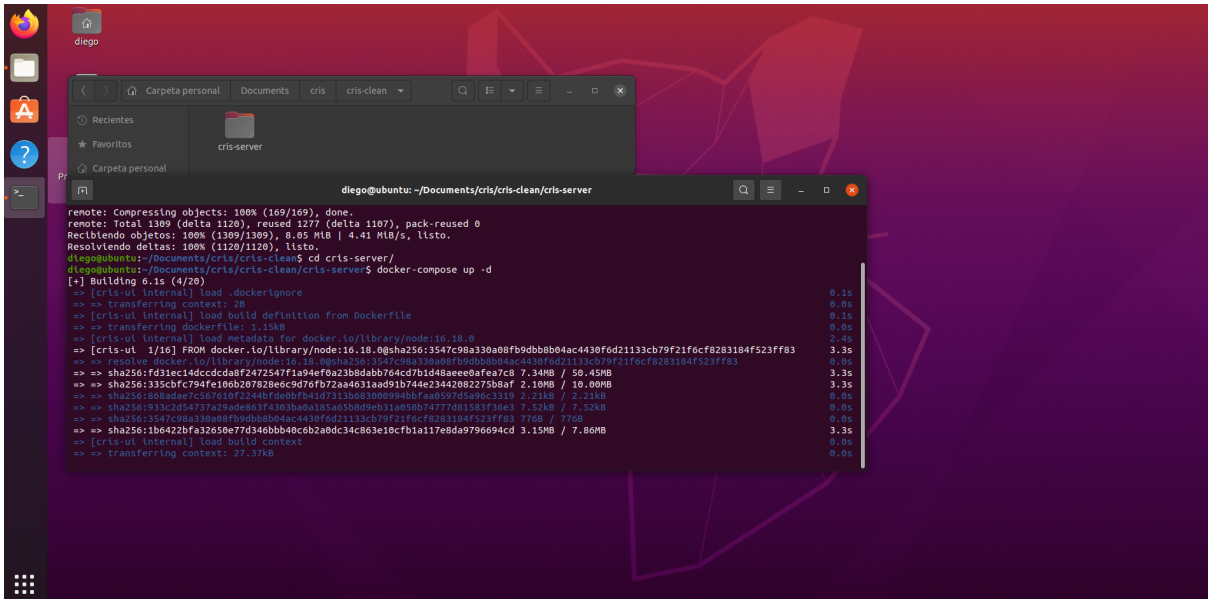
3. Ingresa dentro del proyecto:
   cd cris-server

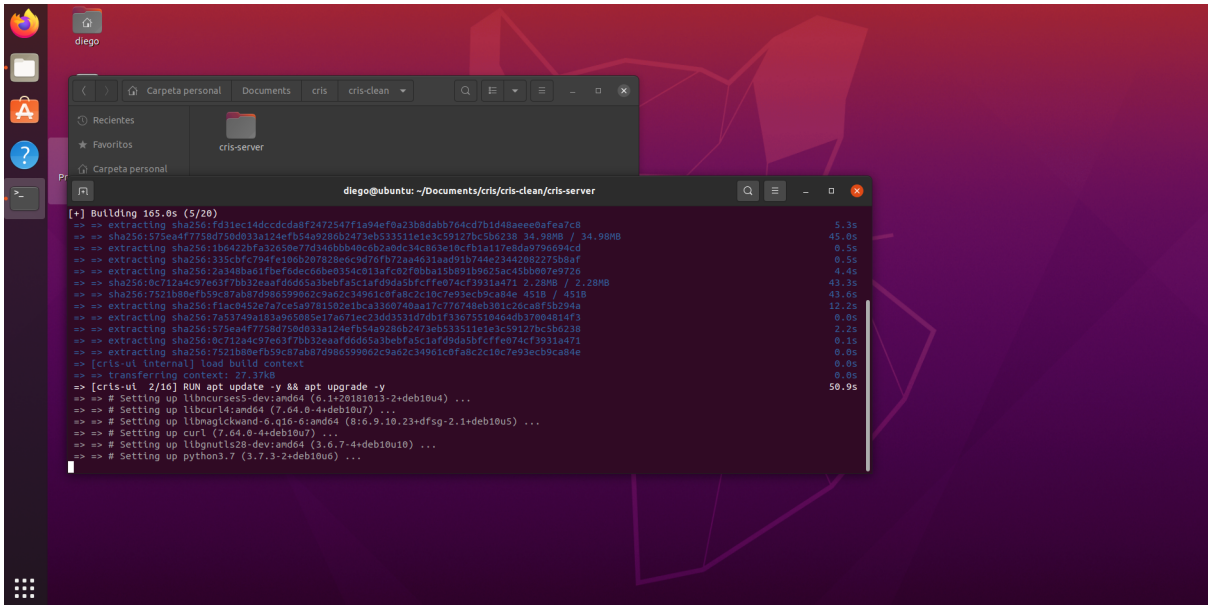


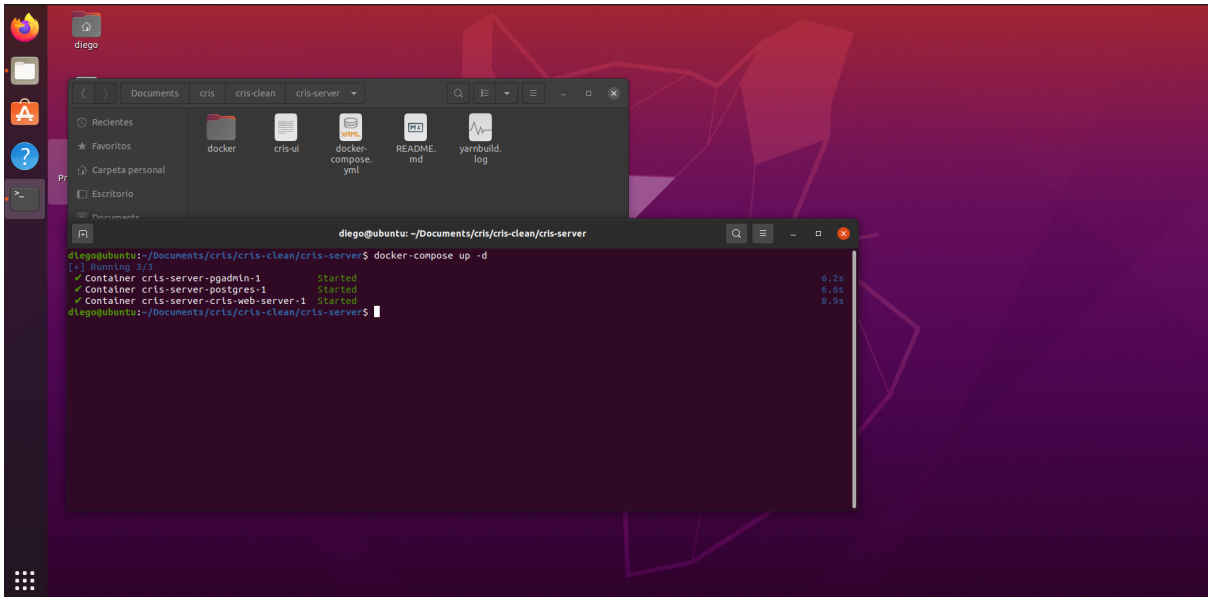
4. Nos movemos a la rama donde se encuentran unificados los servicios:
   git checkout unificado-cris

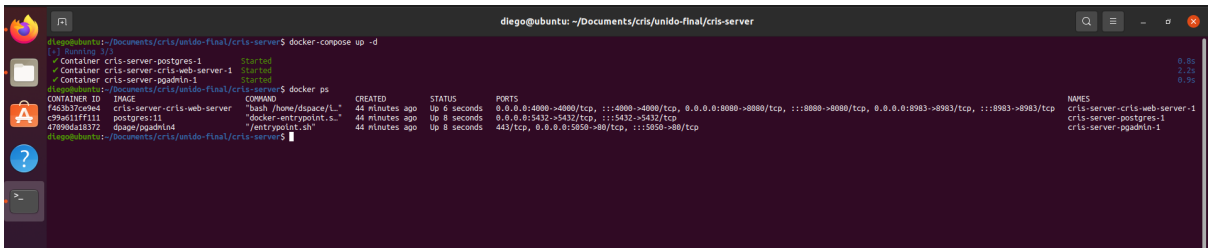


5. Ejecuta el Docker-Compose
   docker-compose up -d



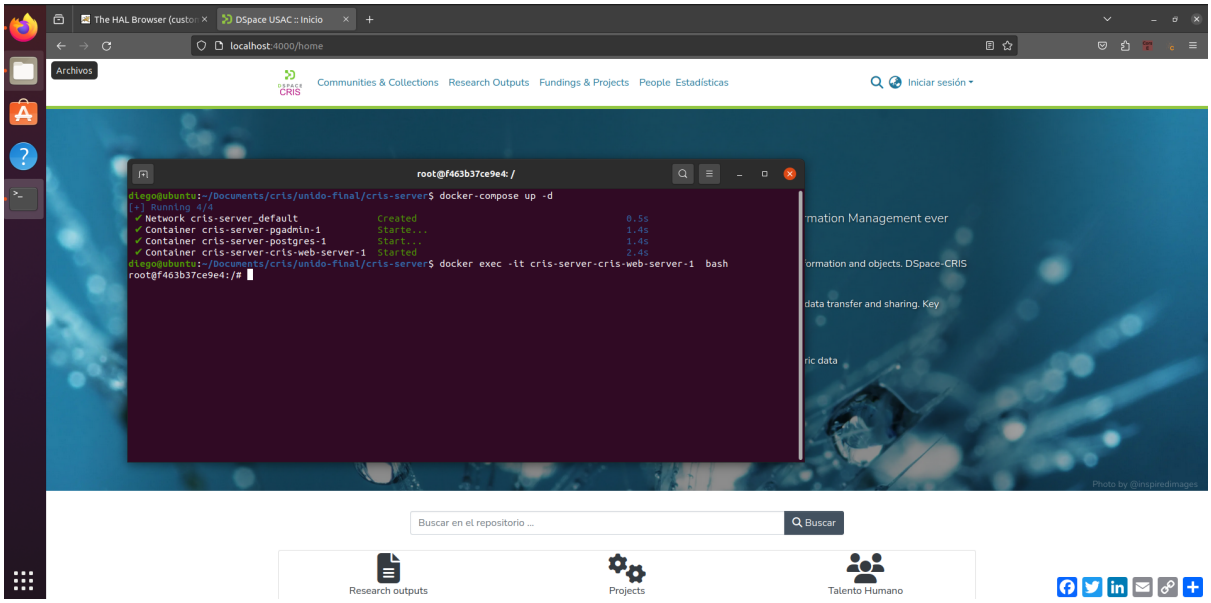




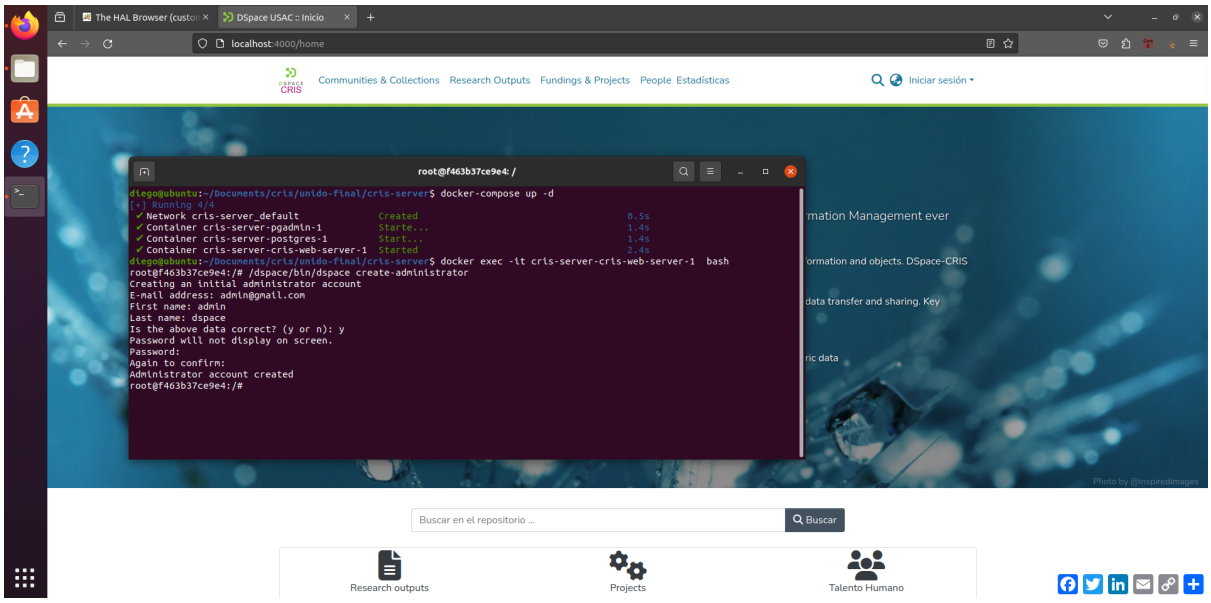


6. Verificar que todo se encuentre funcionando el el puerto configurado.

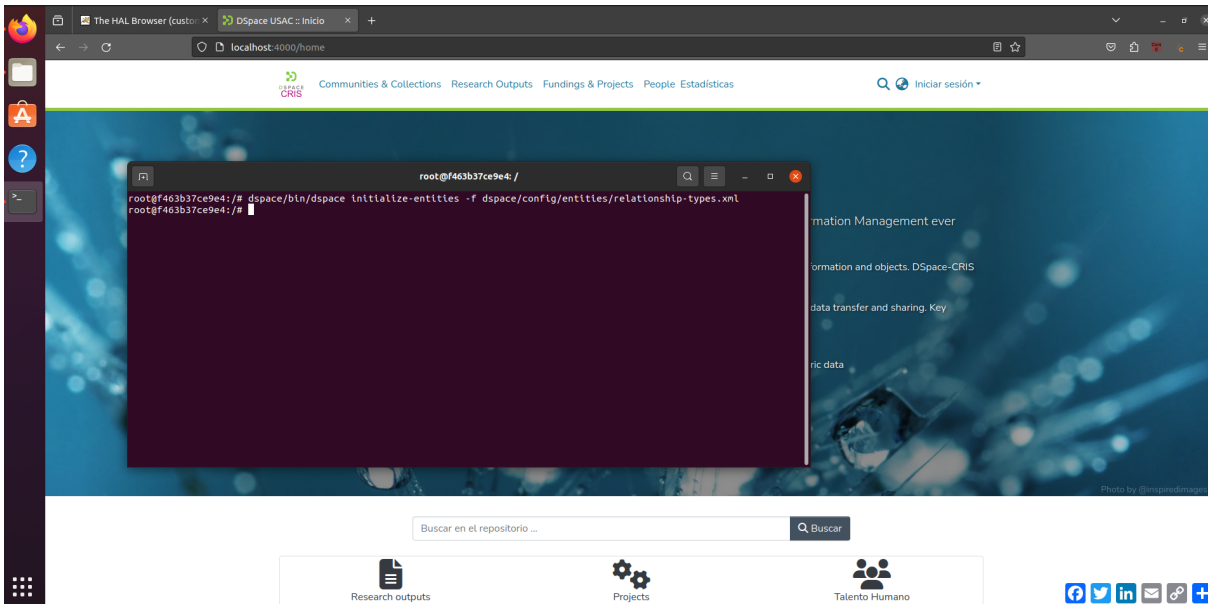
7. Procedemos a ingresar al contenedor web para crear el usuario administrador.
   docker exec -it cris-server-cris-web-server-1 bash



8. Creamos una cuenta de administrador para entrar el sistema DSpace:
   /dspace/bin/dspace create-administrator

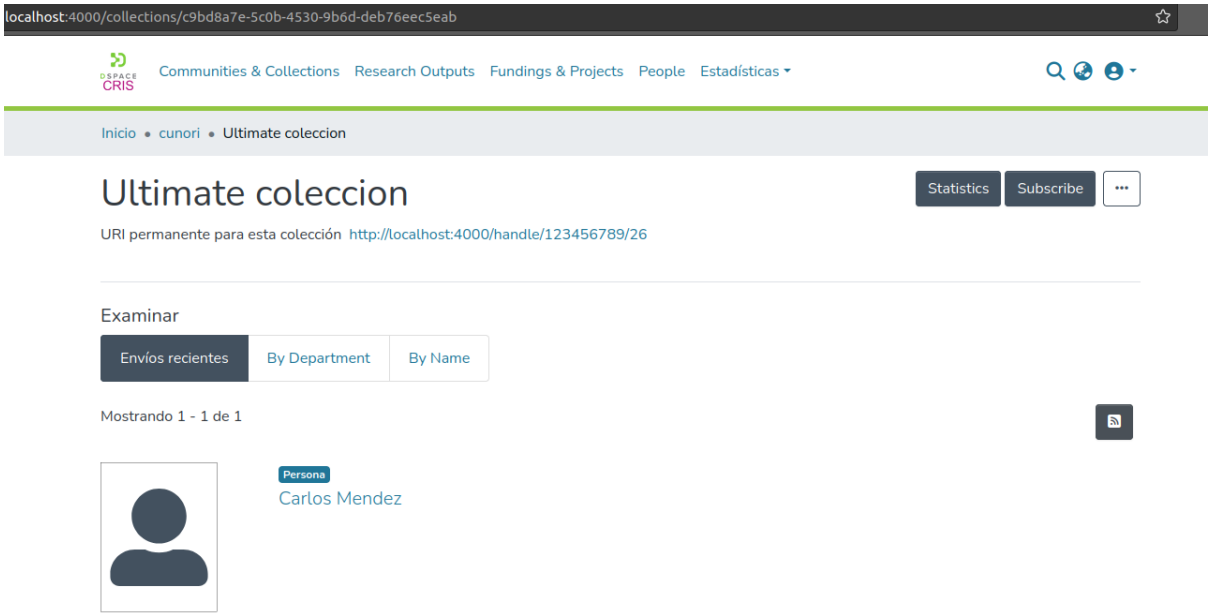


9. Inicializamos las entidades
   dspace/bin/dspace initialize-entities -f dspace/config/entities/relationship-types.xml



10. Corremos este comando para solucionar futuros errores sobre el tipo de items agregados (dentro de /dspace/bin):
    ./dspace dsrun org.dspace.app.util.InitializeEntities -f /dspace/config/entities/correction-relationship-types.xml

11. Finalmente el sistema debería permitir ingresar comunidades, colecciones e ítems de los diferentes tipos de entidades previamente inicializados.

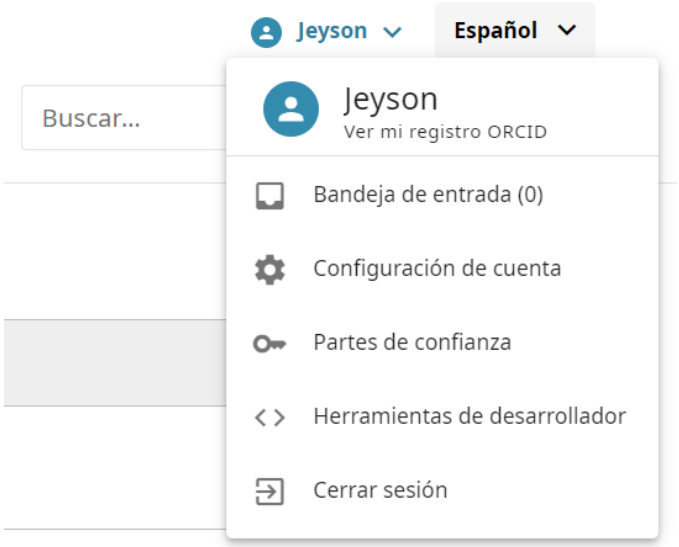


Guía de integración DSPACE-ORCID

Crear una cuenta en sandbox ORCID

Es importante mencionar que sandbox.orcid.org es un sitio web de pruebas. orcid.com es el sitio web oficial. Sandbox solo envía mensajes por correo electrónico a **mailinator.com.**

Luego hacer clic en su nombre en la esquina superior derecha para desplegar el siguiente menú, y seleccionar la opción "Herramientas de desarrollador".



Nos mostrará la siguiente ventana, en la cual debemos proporcionar la información requerida y al llenar todos los campos presionar el botón "Guardar aplicación"

URI de redireccionamiento

Una vez que el usuario haya autorizado su aplicación, volverá a la URI que usted especifique. Debe indicar estas URI de antemano o los usuarios de su integración experimentarán un error.

Tenga en cuenta lo siguiente:

- Solo **las URI HTTPS** se aceptan en la producción
- Los dominios registrados **DEBEN** coincidir exactamente con los dominios utilizados, incluidos los subdominios
- **Registre todas las URI de redireccionamiento en su totalidad cuando sea posible.** Es la opción más segura y la que recomendamos. Para obtener más información acerca de las URI de redireccionamiento, consulte nuestras preguntas frecuentes sobre URI de redireccionamiento

http://127.0.0.1:8080/server/api/authn/orcid

Añadir otra URI de redireccionamiento

Ejemplo de código

- **Autorizar solicitud**
- **Solicitud de token**
- **Solicitud implícita/OpenID**

Guardar aplicación

Luego de guardar los datos se generarán las credenciales para poder hacer uso de la API pública de sandbox ORCID.

Herramientas de desarrollador Volver a mi registro

ID de cliente
APP-7GP7TL9JX32ZWE2P

Código secreto del cliente
26ff9b2d-faa9-4e9f-acd6-1b9b0ff012fd Generar un nuevo código secreto de cliente

Configuración dentro del contenedor

Ingresar al contenedor utilizando el siguiente comando

docker exec -it cris-server_cris-web-server_1 bash

Navegar hasta la siguiente ruta /dspace/config buscar el archivo local.cfg y localizar esta sección

Ingresar las credenciales de la API pública de ORCID

```
##########################################
# ORCID Integration                      #
##########################################
orcid.domain-url= https://sandbox.orcid.org
orcid.api-url = https://api.sandbox.orcid.org/v3.0
orcid.public-url = https://pub.sandbox.orcid.org/v3.0
orcid.webhook-url = https://api.sandbox.orcid.org/
orcid.application-client-id =
orcid.application-client-secret =
orcid.webhook.registration-token =
```

Realizar las siguientes modificaciones

orcid.application-client-id = <TU-SANDBOX-ORCID-CLIENTE-ID>

orcid.application-client-secret = <TU-SANDBOX-ORCID-CLIENTE-SECRETO>

orcid.webhook.registration-token =

authentication-orcid.can-self-register = true

#Las siguientes propiedades son necesarias solo para la parte de sincronizacion / vinculacion (no para la autenticacion o importacion)

orcid.synchronization-enabled = true

#Es necesario habilitar el consumidor orcidqueue para realizar un seguimiento de lo que necesita ser sincronizado entre DSpace y ORCID

event.dispatcher.default.consumers = versioning, discovery, eperson, orcidqueue

Habilitar los permisos para el inicio de sesión

Navegar hasta la siguiente ruta /dspace/config/modules/ buscar el archivo orcid.cfg y localizar esta sección

```
# The scopes to be granted by the user during the login on ORCID
orcid.scope = /authenticate
#orcid.scope = /read-limited
#orcid.scope = /activities/update
#orcid.scope = /person/update
```

Realizar las siguientes modificaciones

Los ámbitos que otorgará el usuario durante el inicio de sesión en ORCID

orcid.scope = /authenticate

Los siguientes alcances SOLO son válidos si tiene una clave API de miembro. Deben comentarse si solo tiene una clave de API pública

#orcid.scope = /read-limited

#orcid.scope = /activities/update

#orcid.scope = /person/update

Habilitar el plugin para iniciar sesión con ORCID

Para habilitar **la autenticación ORCID,** debe descomentar la siguiente línea en el archivo /dspace/config/modules/authentication.cfg

ORCID certificate authentication.

plugin.sequence.org.dspace.authenticate.AuthenticationMethod = org.dspace.authenticate.OrcidAuthentication

```
  GNU nano 6.2                                                    authentication.cfg

# ORCID certificate authentication.
plugin.sequence.org.dspace.authenticate.AuthenticationMethod = org.dspace.authenticate.OrcidAuthentication
```

Cuando un usuario inicia sesión a través de ORCID, el sistema intentará reutilizar una cuenta existente buscando por correo electrónico. Si no se encuentra ninguno, se crea una nueva cuenta en DSpace. Es posible deshabilitar la creación de nuevas cuentas configurando la siguiente propiedad en false en el archivo /dspace/config/modules/authentication-orcid.cfg

```
GNU nano 6.2                                                    authentication-orcid.cfg
#----------------------------------------------------------------#
#------------ORCID AUTHENTICATION CONFIGURATIONS----------------#
#----------------------------------------------------------------#
# Configuration properties used by the ORCID                    #
# Authentication plugin, when it is enabled.                    #
#----------------------------------------------------------------#

# Should we allow new users to be registered automatically?
authentication-orcid.can-self-register = true
```

Las propiedades a continuación son necesarias solo para la parte de sincronización/vinculación (no para autenticación o importación) configurarla en el archivo /dspace/config/modules/orcid.cfg

```
GNU nano 6.2                                                          orcid.cfg

### Property to enable or disable the synchronization between DSpace and ORCID
orcid.synchronization-enabled = true
```

Debe permitir que el consumidor de orcidqueue realice un seguimiento de lo que debe sincronizarse entre DSpace y el evento de ORCID esto debe configurarse en el archivo /dspace/config/modules/dspace.cfg

```
GNU nano 6.2                                                          dspace.cfg
# Add doi here if you are using org.dspace.identifier.DOIIdentifierProvider to generate DOIs.
# Adding doi here makes DSpace send metadata updates to your doi registration agency.
# Add rdf here, if you are using dspace-rdf to export your repository content as RDF.
event.dispatcher.default.consumers = versioning, discovery, eperson, dedup, crisconsumer, orcidqueue,
```

Para proporcionar mensajes más significativos al usuario, DSpace implementa una validación local antes de intentar enviar el registro a ORCID. Esta validación verifica los datos utilizando las reglas anteriores para que se muestre un mensaje específico al usuario. Si por algún motivo ORCID devuelve otro error, se muestra un mensaje genérico al usuario y el mensaje técnico exacto recibido por ORCID se registra en el archivo dspace.log y se almacena en la tabla orcidhistory.

```
GNU nano 6.2                                                          orcid.cfg
#-----------------------------------------------------------------#
#-------------ORCID OBJECTS VALIDATION CONFIGURATIONS-------------#
#-----------------------------------------------------------------#

orcid.validation.affiliation.enabled = true
orcid.validation.work.enabled = true
orcid.validation.funding.enabled = true
```

Informe final de proyecto de investigación. Año 2023

Mapeo de los elementos de persona de DSpace a obras ORCID

En el archivo /dspace/config/spring/api/orcid-services.xml

```
GNU nano 6.2                                                                    orcid-services.xml
  <bean class="org.dspace.orcid.service.impl.OrcidProfileSectionFactoryServiceImpl">
      <constructor-arg name="sectionFactories">
              <list>

                      <bean class="org.dspace.orcid.model.factory.impl.OrcidAffiliationFactory">
                                      <constructor-arg name="sectionType" value="AFFILIATION" />
                                      <constructor-arg name="preference" value="AFFILIATION" />
                                      <property name="organizationField" value="${orcid.mapping.affiliation.name}" />
                                      <property name="roleField" value="${orcid.mapping.affiliation.role}" />
                                      <property name="startDateField" value="${orcid.mapping.affiliation.start-date}" />
                                      <property name="endDateField" value="${orcid.mapping.affiliation.end-date}" />
                              </bean>

                      <bean class="org.dspace.orcid.model.factory.impl.OrcidAffiliationFactory">
                                      <constructor-arg name="sectionType" value="QUALIFICATION" />
                                      <constructor-arg name="preference" value="EDUCATION" />
                                      <property name="organizationField" value="${orcid.mapping.qualification.name}" />
                                      <property name="roleField" value="${orcid.mapping.qualification.role}" />
                                      <property name="startDateField" value="${orcid.mapping.qualification.start-date}" />
                                      <property name="endDateField" value="${orcid.mapping.qualification.end-date}" />
                              </bean>

                      <bean class="org.dspace.orcid.model.factory.impl.OrcidAffiliationFactory">
                                      <constructor-arg name="sectionType" value="EDUCATION" />
                                      <constructor-arg name="preference" value="EDUCATION" />
                                      <property name="organizationField" value="${orcid.mapping.education.name}" />
                                      <property name="roleField" value="${orcid.mapping.education.role}" />
                                      <property name="startDateField" value="${orcid.mapping.education.start-date}" />
                                      <property name="endDateField" value="${orcid.mapping.education.end-date}" />
                              </bean>

                      <bean class="org.dspace.orcid.model.factory.impl.OrcidSimpleValueObjectFactory">
                                      <constructor-arg name="sectionType" value="OTHER_NAMES" />
                                      <constructor-arg name="preference" value="BIOGRAPHICAL" />
                                      <property name="metadataFields" value="${orcid.mapping.other-names}" />
                              </bean>

                      <bean class="org.dspace.orcid.model.factory.impl.OrcidSimpleValueObjectFactory">
                                      <constructor-arg name="sectionType" value="KEYWORDS" />
                                      <constructor-arg name="preference" value="BIOGRAPHICAL" />
                                      <property name="metadataFields" value="${orcid.mapping.keywords}" />
                              </bean>
```

```
                      <bean class="org.dspace.orcid.model.factory.impl.OrcidPersonExternalIdentifierFactory">
                                      <constructor-arg name="sectionType" value="EXTERNAL_IDS" />
                                      <constructor-arg name="preference" value="IDENTIFIERS" />
                                      <property name="externalIds" value="${orcid.mapping.person-external-ids}" />
                              </bean>

                      <bean class="org.dspace.orcid.model.factory.impl.OrcidSimpleValueObjectFactory">
                                      <constructor-arg name="sectionType" value="RESEARCHER_URLS" />
                                      <constructor-arg name="preference" value="IDENTIFIERS" />
                                      <property name="metadataFields" value="${orcid.mapping.researcher-urls}" />
                              </bean>

              </list>
      </constructor-arg>
  </bean>
```

La configuración anterior vincula cada pieza de información que se puede sincronizar desde DSpace a ORCID con una preferencia que el usuario puede administrar en el lado de DSpace (es decir, la sincronización de las palabras clave está vinculada a la preferencia BIOGRAFICA) y define qué metadatos de DSpace se utilizarán para Complete el campo ORCID. El bean lee el mapeo de metadatos del config/modules/orcid.cfg

```
  GNU nano 6.2                                    orcid.cfg
### Other names mapping ###
orcid.mapping.other-names = crisrp.name.variant
orcid.mapping.other-names = crisrp.name.translated

### Keywords mapping ###
orcid.mapping.keywords = dc.subject

### Country mapping ###
orcid.mapping.country = crisrp.country
orcid.mapping.country.converter =

### Person External ids mapping ###
##orcid.mapping.person-external-ids syntax is <metadatafield>::<type>
orcid.mapping.person-external-ids = person.identifier.scopus-author-id::SCOPUS
orcid.mapping.person-external-ids = person.identifier.rid::RID

### Researcher urls mapping ###
orcid.mapping.researcher-urls = oairecerif.identifier.url
```

Mapeo de elementos de publicación DSpace a ORCID Works

En el archivo /dspace/config/spring/api/orcid-services.xml

```
  GNU nano 6.2                                    orcid-services.xml
        <bean id="orcidWorkFactoryFieldMapping" class="org.dspace.orcid.model.OrcidWorkFieldMapping" >
                <property name="contributorFields" value="${orcid.mapping.work.contributors}" />
                <property name="externalIdentifierFields" value="${orcid.mapping.work.external-ids}" />
                <property name="publicationDateField" value="${orcid.mapping.work.publication-date}" />
                <property name="titleField" value="${orcid.mapping.work.title}" />
                <property name="journalTitleField" value="${orcid.mapping.work.journal-title}" />
                <property name="shortDescriptionField" value="${orcid.mapping.work.short-description}" />
                <property name="subTitleField" value="${orcid.mapping.work.sub-title}" />
                <property name="languageField" value="${orcid.mapping.work.language}" />
                <property name="languageConverter" ref="${orcid.mapping.work.language.converter}" />
                <property name="typeField" value="${orcid.mapping.work.type}" />
                <property name="typeConverter" ref="${orcid.mapping.work.type.converter}" />
                <property name="fundingField" value="${orcid.mapping.work.funding}" />
                <property name="fundingUrlField" value="${orcid.mapping.work.funding.url}" />
                <property name="fundingExternalIdType" value="${orcid.mapping.work.funding.external-id.type}" />
                <property name="fundingExternalId" value="${orcid.mapping.work.funding.external-id.value}" />
                <property name="fundingEntityExternalId" value="${orcid.mapping.work.funding.external-id.entity-value}" />
                <property name="citationType" value="${orcid.mapping.work.citation.type}" />
                <property name="citationCrosswalks">
                        <util:map>
                                <entry key="bibtex" value-ref="cslItemDataCrosswalkBibtex" />
                        </util:map>
                </property>
        </bean>
```

Lee el mapeo del archivo config/modules/orcid.cfg

```
GNU nano 6.2                                    orcid.cfg
### Work (Publication) mapping ###
orcid.mapping.work.title = dc.title
orcid.mapping.work.sub-title =

orcid.mapping.work.short-description = dc.description.abstract
orcid.mapping.work.publication-date = dc.date.issued
orcid.mapping.work.language = dc.language.iso
orcid.mapping.work.language.converter = mapConverterDSpaceToOrcidLanguageCode
orcid.mapping.work.journal-title = dc.relation.ispartof
orcid.mapping.work.type = dc.type
orcid.mapping.work.type.converter = mapConverterDSpaceToOrcidPublicationType
orcid.mapping.work.citation.type = bibtex

##orcid.mapping.work.contributors syntax is <metadatafield>::<role>
orcid.mapping.work.contributors = dc.contributor.author::author
orcid.mapping.work.contributors = dc.contributor.editor::editor

##orcid.mapping.work.external-ids syntax is <metadatafield>::<type> or $simple-handle::<type>
##The full list of available external identifiers is available here https://pub.orcid.org/v3.0/identifiers
orcid.mapping.work.external-ids = dc.identifier.doi::doi
orcid.mapping.work.external-ids = dc.identifier.scopus::eid
orcid.mapping.work.external-ids = dc.identifier.pmid::pmid
orcid.mapping.work.external-ids = $simple-handle::handle
orcid.mapping.work.external-ids = dc.identifier.isi::wosuid
orcid.mapping.work.external-ids = dc.identifier.issn::issn
```

El valor de orcid.mapping.work.type.converter coincide con el nombre de un bean definido en el archivo config/spring/api/orcid-services.xml

```
GNU nano 6.2                                    orcid-services.xml
    <bean name="mapConverterDSpaceToOrcidPublicationType" class="org.dspace.util.SimpleMapConverter" init-method="init">
            <property name="converterNameFile" value="orcid/mapConverter-dspace-to-orcid-publication-type.properties" />
            <property name="configurationService" ref="org.dspace.services.ConfigurationService" />
            <property name="defaultValue" value="other"/>
        </bean>
```

Asignación de elementos del proyecto DSpace a la financiación ORCID

En el archivo /dspace/config/spring/api/orcid-services.xml

```
GNU nano 6.2                                    orcid-services.xml
        <bean id="orcidFundingFactoryFieldMapping" class="org.dspace.orcid.model.OrcidFundingFieldMapping" >
                <property name="contributorFields" value="${orcid.mapping.funding.contributors}" />
                <property name="externalIdentifierFields" value="${orcid.mapping.funding.external-ids}" />
                <property name="titleField" value="${orcid.mapping.funding.title}" />
                <property name="typeField" value="${orcid.mapping.funding.type}" />
                <property name="typeConverter" ref="${orcid.mapping.funding.type.converter}" />
                <property name="amountField" value="${orcid.mapping.funding.amount}" />
                <property name="amountCurrencyField" value="${orcid.mapping.funding.amount.currency}" />
                <property name="amountCurrencyConverter" ref="${orcid.mapping.funding.amount.currency.converter}" />
                <property name="descriptionField" value="${orcid.mapping.funding.description}" />
                <property name="startDateField" value="${orcid.mapping.funding.start-date}" />
                <property name="endDateField" value="${orcid.mapping.funding.end-date}" />
                <property name="organizationField" value="${orcid.mapping.funding.organization}" />
        </bean>
```

Lee el mapeo del archivo config/modules/orcid.cfg

```
GNU nano 6.2                                        orcid.cfg
### Funding mapping ###
orcid.mapping.funding.title = dc.title
orcid.mapping.funding.type = dc.type
orcid.mapping.funding.type.converter = mapConverterDSpaceToOrcidFundingType
##orcid.mapping.funding.external-ids syntax is <metadatafield>::<type>
##The full list of available external identifiers is available here https://pub.orcid.org/v3.0/identifiers
orcid.mapping.funding.external-ids = oairecerif.internalid::other-id
orcid.mapping.funding.external-ids = crisfund.award.url::uri
orcid.mapping.funding.external-ids = oairecerif.funding.identifier::grant_number
orcid.mapping.funding.description = dc.description
orcid.mapping.funding.start-date = oairecerif.funding.startDate
orcid.mapping.funding.end-date = oairecerif.funding.endDate
##orcid.mapping.funding.contributors syntax is <metadatafield>::<type>
orcid.mapping.funding.contributors = crisfund.investigators::lead
orcid.mapping.funding.contributors = crisfund.coinvestigators::co-lead
orcid.mapping.funding.organization = oairecerif.funder
orcid.mapping.funding.amount = oairecerif.amount
orcid.mapping.funding.amount.currency = oairecerif.amount.currency
orcid.mapping.funding.amount.currency.converter = mapConverterDSpaceToOrcidAmountCurrency
```

El valor de orcid.mapping.funding.type.converter coincide con el nombre de un bean definido en config/spring/api/orcid-services.xml lo mismo se aplica para la moneda orcid.mapping.funding.amount.currency.converter = mapConverterDSpaceToOrcidAmountCurrency

```
GNU nano 6.2                                        orcid-services.xml
  <bean name="mapConverterDSpaceToOrcidFundingType" class="org.dspace.util.SimpleMapConverter" init-method="init">
            <property name="converterNameFile" value="orcid/mapConverter-dspace-to-orcid-funding-type.properties" />
            <property name="configurationService" ref="org.dspace.services.ConfigurationService" />
            <property name="defaultValue" value="grant"/>
      </bean>

  <bean name="mapConverterDSpaceToOrcidAmountCurrency" class="org.dspace.util.SimpleMapConverter" init-method="init">
            <property name="converterNameFile" value="orcid/mapConverter-dspace-to-orcid-amount-currency.properties" />
            <property name="configurationService" ref="org.dspace.services.ConfigurationService" />
            <property name="defaultValue" value=""/>
      </bean>
```

Configurar las funciones de importación

El siguiente bean se utiliza para configurar la **importación de registros de personas desde ORCID**. Se activa como fuente externa en config/spring/api/external-services.xml

```
GNU nano 6.2                                        external-services.xml
  <bean class="org.dspace.external.provider.impl.OrcidV3AuthorDataProvider" init-method="init">
      <property name="sourceIdentifier" value="orcid"/>
      <property name="orcidUrl" value="${orcid.domain-url}" />
      <property name="clientId" value="${orcid.application-client-id}" />
      <property name="clientSecret" value="${orcid.application-client-secret}" />
      <property name="OAUTHUrl" value="${orcid.token-url}" />
      <property name="orcidRestConnector" ref="orcidRestConnector"/>
      <property name="supportedEntityTypes">
          <list>
              <value>Person</value>
          </list>
      </property>
  </bean>
```

El mapeo entre ORCID Person y DSpace Person Item es el siguiente, actualmente codificado

ORCID	DSpace
Name/FamilyName	person.firstName
Name/GivenName	person.givenName
Name/Path	person.identifier.orcid
ORCID Profile URL	dc.identifier.uri

El siguiente bean se utiliza para configurar la importación de registros de publicación desde ORCID (Trabajo)

```
GNU nano 6.2                                                    external-services.xml
  <bean id="orcidPublicationDataProvider" class="org.dspace.external.provider.impl.OrcidPublicationDataProvider">
      <property name="sourceIdentifier" value="orcidWorks"/>
      <property name="fieldMapping" ref="orcidPublicationDataProviderFieldMapping"/>
      <property name="supportedEntityTypes">
          <list>
              <value>Publication</value>
          </list>
      </property>
  </bean>

  <bean id="vufindLiveImportDataProvider" class="org.dspace.external.provider.impl.LiveImportDataProvider">
      <property name="metadataSource" ref="vufindImportService"/>
      <property name="sourceIdentifier" value="vufind"/>
      <property name="recordIdMetadata" value="dc.identifier.other"/>
      <property name="supportedEntityTypes">
          <list>
              <value>Publication</value>
          </list>
      </property>
  </bean>
```

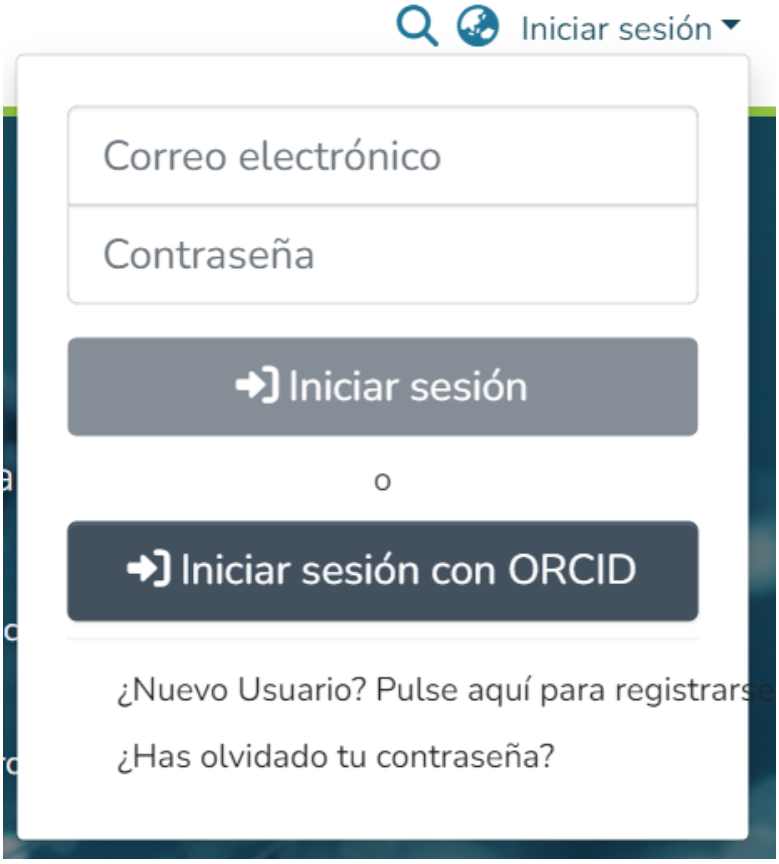
El mapeo de los metadatos de ORCID Work a los metadatos de DSpace se realiza mediante el siguiente bean en config/spring/api/orcid-services.xml

```
GNU nano 6.2                                                    orcid-services.xml
      <bean id="orcidPublicationDataProviderFieldMapping" class="org.dspace.orcid.model.OrcidWorkFieldMapping" >
              <property name="contributorFields" value="${orcid.external-data.mapping.publication.contributors}" />
              <property name="externalIdentifierFields" value="${orcid.external-data.mapping.publication.external-ids}" />
              <property name="publicationDateField" value="${orcid.external-data.mapping.publication.issued-date}" />
              <property name="titleField" value="${orcid.external-data.mapping.publication.title}" />
              <property name="journalTitleField" value="${orcid.external-data.mapping.publication.is-part-of}" />
              <property name="shortDescriptionField" value="${orcid.external-data.mapping.publication.description}" />
              <property name="languageField" value="${orcid.external-data.mapping.publication.language}" />
              <property name="languageConverter" ref="${orcid.external-data.mapping.publication.language.converter}" />
              <property name="typeField" value="${orcid.external-data.mapping.publication.type}" />
              <property name="typeConverter" ref="${orcid.external-data.mapping.publication.type.converter}" />
      </bean>
```

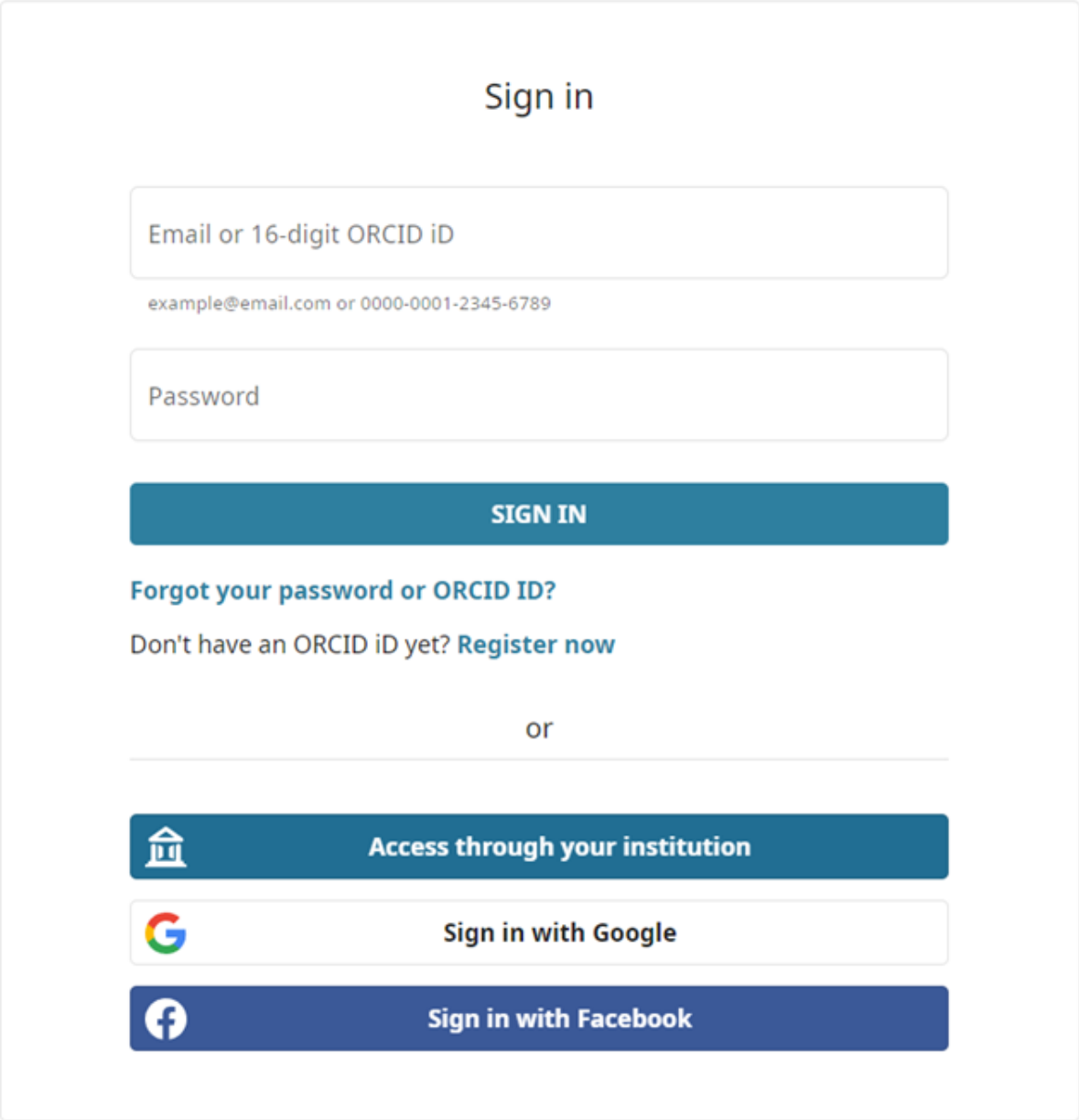
Que lee el mapeo del archivo config/modules/orcid.cfg

```
  GNU nano 6.2                                                    orcid.cfg *
### Work (Publication) external-data.mapping ###
orcid.external-data.mapping.publication.title = dc.title

orcid.external-data.mapping.publication.description = dc.description.abstract
orcid.external-data.mapping.publication.issued-date = dc.date.issued
orcid.external-data.mapping.publication.language = dc.language.iso
orcid.external-data.mapping.publication.language.converter = mapConverterOrcidToDSpaceLanguageCode
orcid.external-data.mapping.publication.is-part-of = dc.relation.ispartof
orcid.external-data.mapping.publication.type = dc.type
orcid.external-data.mapping.publication.type.converter = mapConverterOrcidToDSpacePublicationType

##orcid.external-data.mapping.publication.contributors syntax is <metadatafield>::<role>
orcid.external-data.mapping.publication.contributors = dc.contributor.author::author
orcid.external-data.mapping.publication.contributors = dc.contributor.editor::editor

##orcid.external-data.mapping.publication.external-ids syntax is <metadatafield>::<type> or $simple-handle::<type>
##The full list of available external identifiers is available here https://pub.orcid.org/v3.0/identifiers
orcid.external-data.mapping.publication.external-ids = dc.identifier.doi::doi
orcid.external-data.mapping.publication.external-ids = dc.identifier.scopus::eid
orcid.external-data.mapping.publication.external-ids = dc.identifier.pmid::pmid
orcid.external-data.mapping.publication.external-ids = dc.identifier.isi::wosuid
orcid.external-data.mapping.publication.external-ids = dc.identifier.issn::issn
```

Visualización en el frontend



Informe final de proyecto de investigación. Año 2023



Informe final de proyecto de investigación. Año 2023

Exploración DSpace CRIS (Versión 7)

Fuentes consultadas para la instalación

Documentación oficial
https://wiki.lyrasis.org/display/DSPACECRIS/DSpace-CRIS+Homehttps://wiki.lyrasis.org/display/DSPACECRIS/Installation

Github https://github.com/4Science/DSpace-angular/tree/dspace-cris-7#configuring

Información de terceros
https://mohammedh.io/2021/07/05/install-dspace-cris-7-along-with-free-ssl-and-nginx/https://www.youtube.com/watch?v=2FwaaM3NTBs

Stack de lenguajes/tecnologías

Backend: Java -Spring Solr Postgres

Frontend: Angular Bootstrap

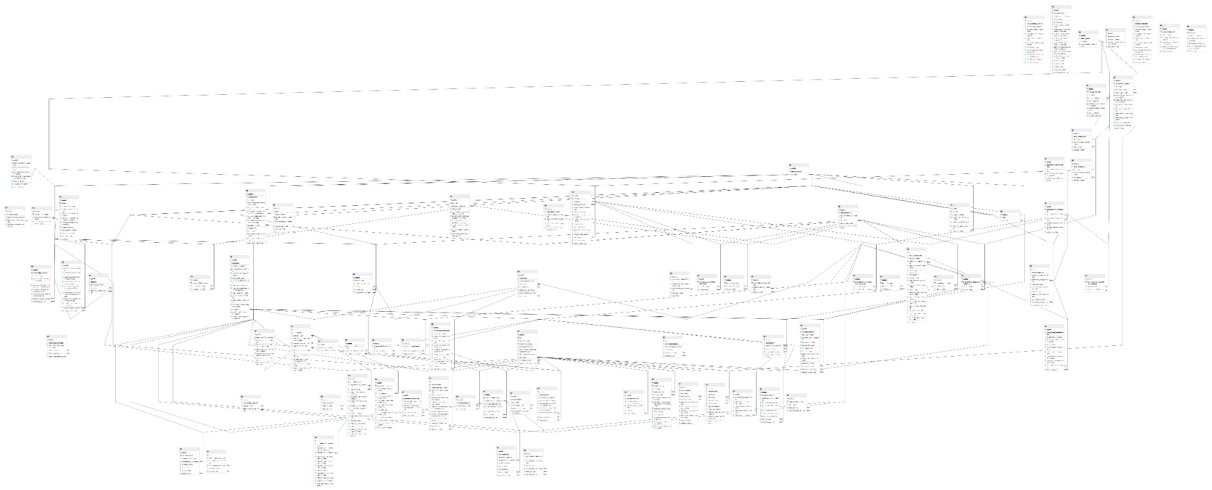
Status: Backend API Rest: OK Frontend:OK

Docker (Unificado fronted y backend)

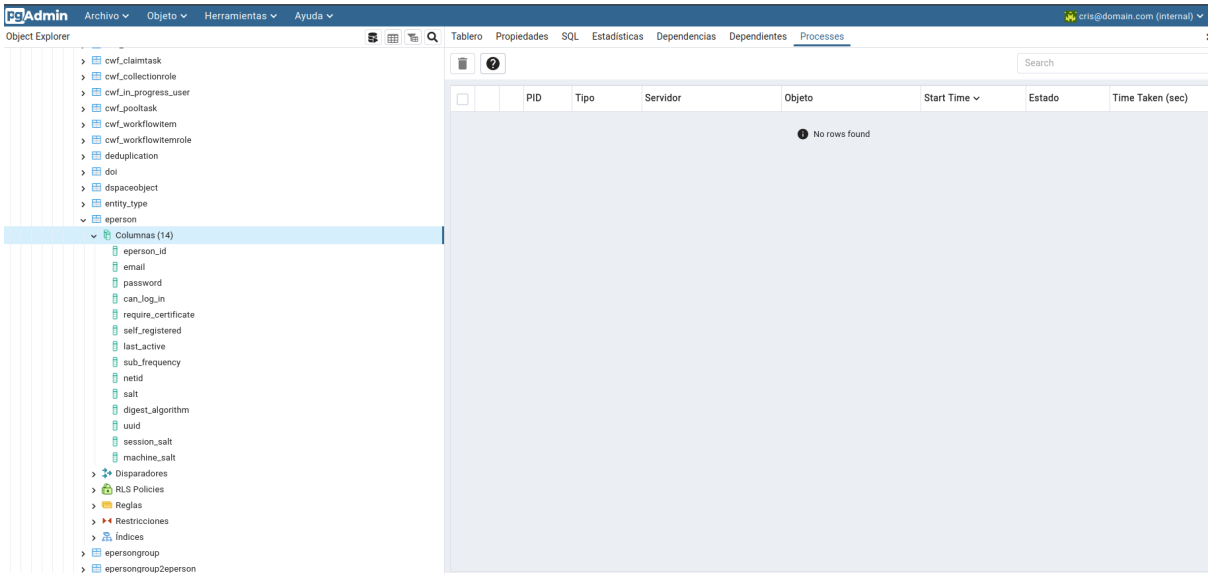
```
diego@ubuntu:~/Documents/cris/unido-final/cris-server$ docker-compose -f docker-
compose.yml start
[+] Running 3/3
 ✔ Container cris-server-pgadmin-1          Started                    2.1s
 ✔ Container cris-server-postgres-1         Started                    2.3s
 ✔ Container cris-server-cris-web-server-1  Started                    3.4s
```

Informe final de proyecto de investigación. Año 2023

Diagrama ER de la base de datos Dspace



Tablas de la base de datos en la interfaz de pgAdmin



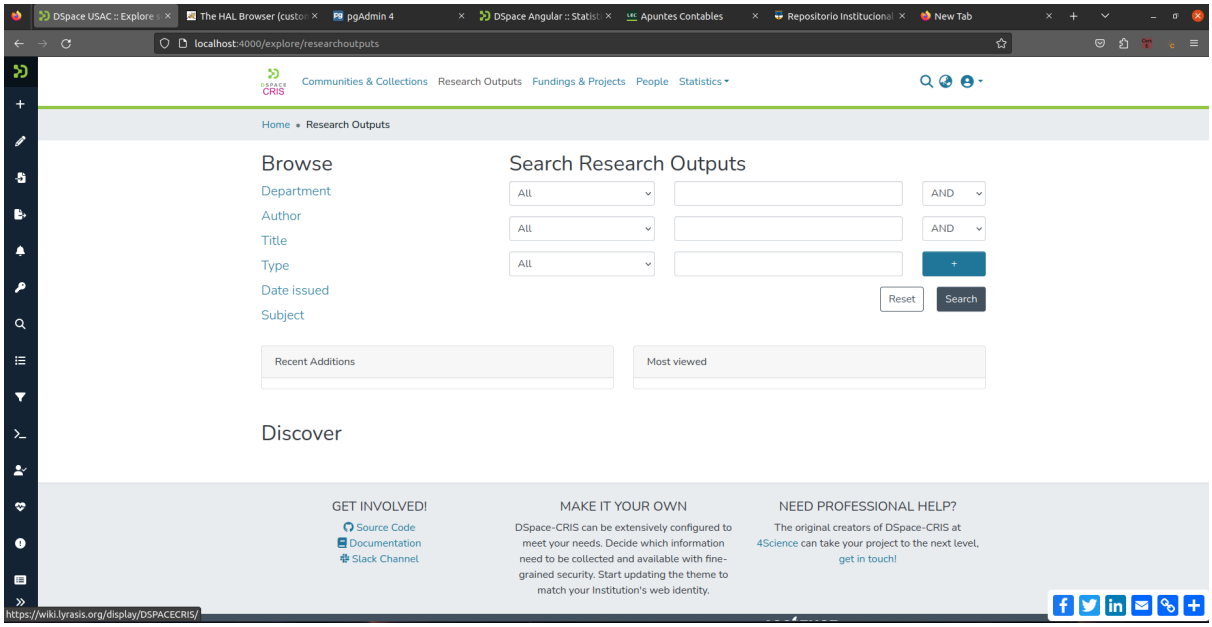
Como ver las tablas en el contenedor de postgres

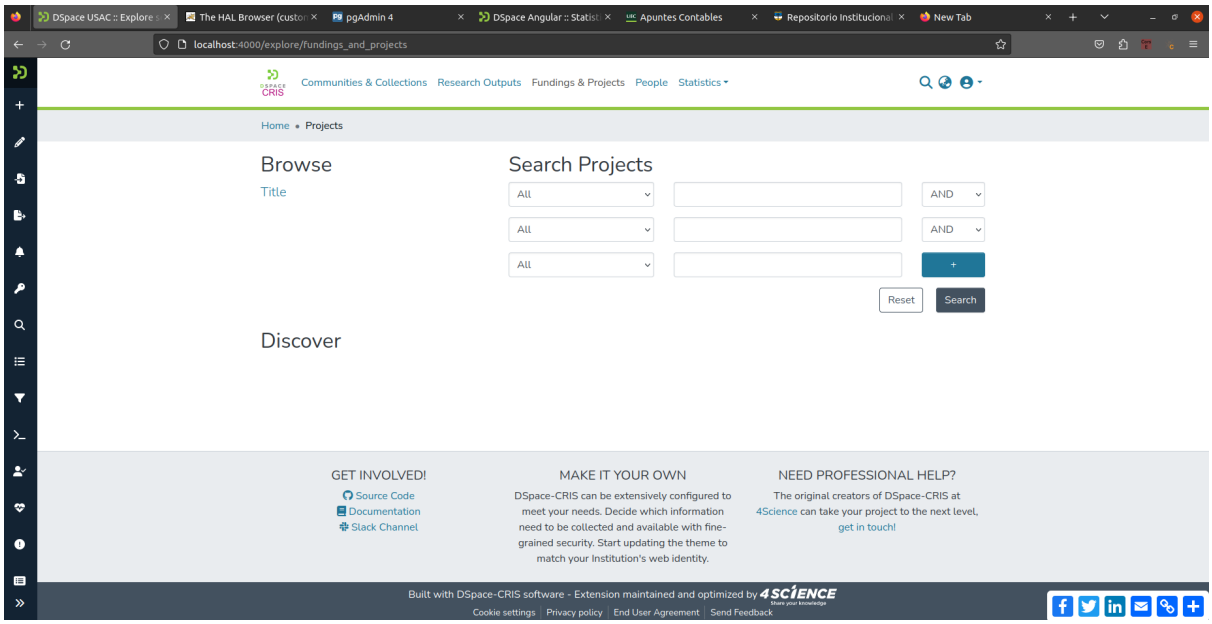
docker exect -it cris-server_postgres_1 bash

```
root@b02b22d95e7c:/# psql -U postgres -W
Password:
psql (11.16 (Debian 11.16-1.pgdg90+1))
Type "help" for help.

postgres=# \l
                                  List of databases
   Name    |  Owner   | Encoding |  Collate   |   Ctype    |   Access privileges
-----------+----------+----------+------------+------------+-----------------------
 dspace    | dspace   | UTF8     | en_US.utf8 | en_US.utf8 | =Tc/dspace           +
           |          |          |            |            | dspace=CTc/dspace
 postgres  | postgres | UTF8     | en_US.utf8 | en_US.utf8 |
 template0 | postgres | UTF8     | en_US.utf8 | en_US.utf8 | =c/postgres          +
           |          |          |            |            | postgres=CTc/postgres
 template1 | postgres | UTF8     | en_US.utf8 | en_US.utf8 | =c/postgres          +
           |          |          |            |            | postgres=CTc/postgres
(4 rows)

postgres=#
```

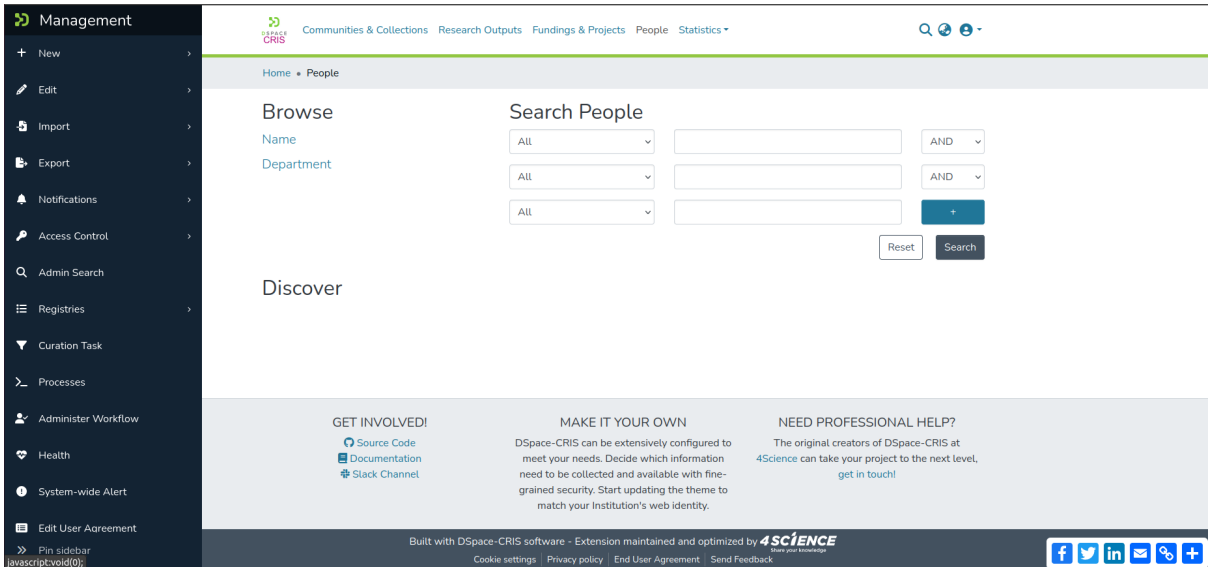
```
root@b02b22d95e7c:/# psql -U postgres -d dspace
psql (11.16 (Debian 11.16-1.pgdg90+1))
Type "help" for help.

dspace=# \dt
                    List of relations
 Schema |                Name                | Type  | Owner
--------+------------------------------------+-------+--------
 public | bitstream                          | table | dspace
 public | bitstreamformatregistry            | table | dspace
 public | bkp_schema_version                 | table | dspace
 public | bundle                             | table | dspace
 public | bundle2bitstream                   | table | dspace
 public | checksum_history                   | table | dspace
 public | checksum_results                   | table | dspace
 public | collection                         | table | dspace
 public | collection2item                    | table | dspace
 public | community                          | table | dspace
 public | community2collection               | table | dspace
 public | community2community                | table | dspace
 public | cris_layout_box                    | table | dspace
 public | cris_layout_box2securitygroup      | table | dspace
 public | cris_layout_box2securitymetadata   | table | dspace
 public | cris_layout_cell                   | table | dspace
 public | cris_layout_field                  | table | dspace
 public | cris_layout_field2nested           | table | dspace
 public | cris_layout_metric2box             | table | dspace
 public | cris_layout_row                    | table | dspace
 public | cris_layout_tab                    | table | dspace
 public | cris_layout_tab2securitygroup      | table | dspace
 public | cris_layout_tab2securitymetadata   | table | dspace
 public | cris_metrics                       | table | dspace
 public | cwf_claimtask                      | table | dspace
 public | cwf_collectionrole                 | table | dspace
 public | cwf_in_progress_user               | table | dspace
 public | cwf_pooltask                       | table | dspace
 public | cwf_workflowitem                   | table | dspace
--More--
```

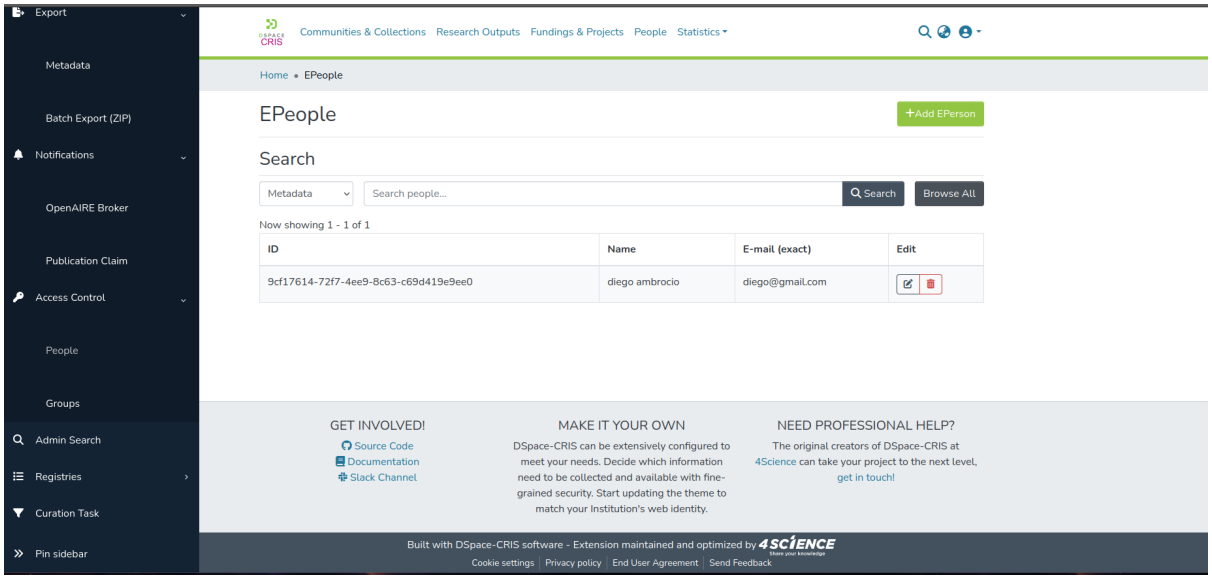
Informe final de proyecto de investigación. Año 2023



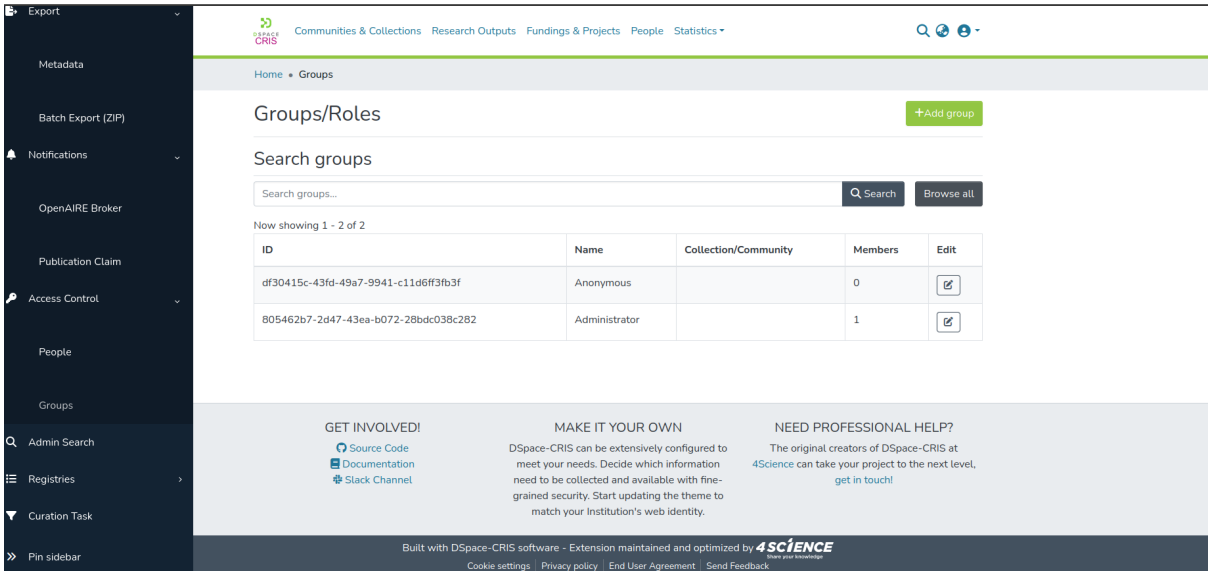


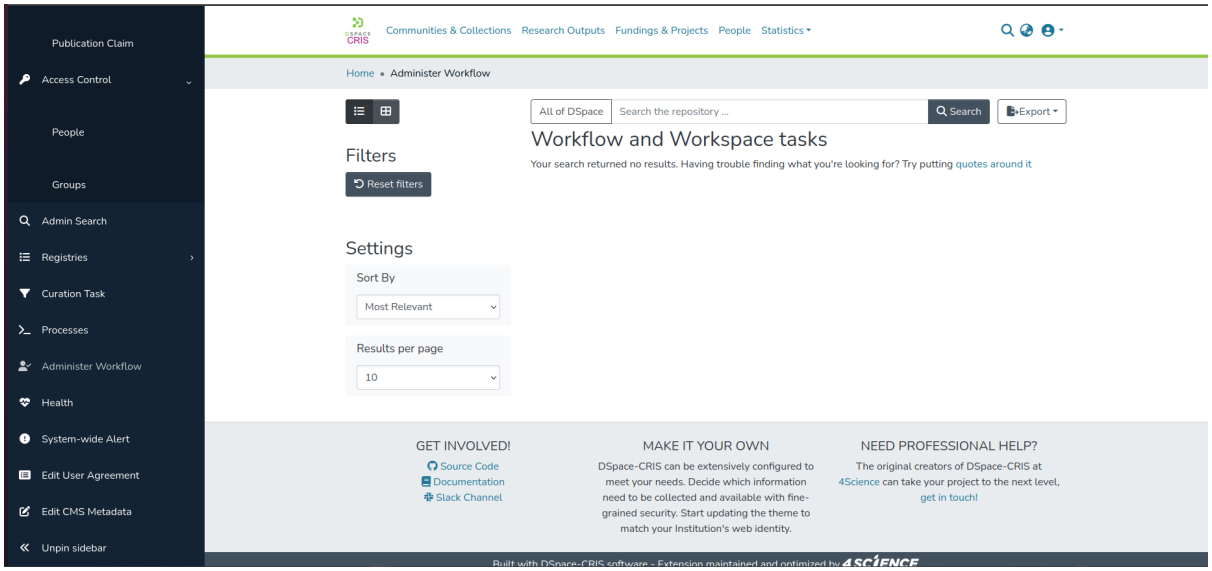
Informe final de proyecto de investigación. Año 2023



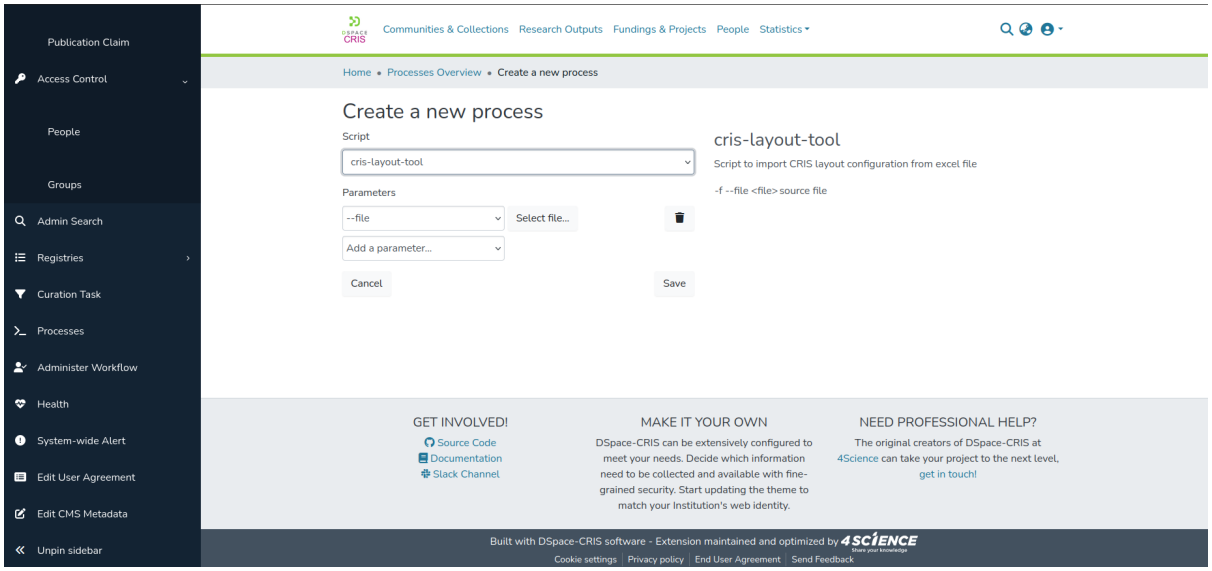


Informe final de proyecto de investigación. Año 2023





Informe final de proyecto de investigación. Año 2023




```
diego@ubuntu:~/Documents/cris/unido-final/cris-server$ docker-compose -f docker-
compose.yml start
[+] Running 3/3
 ✔ Container cris-server-pgadmin-1          Started                    2.1s
 ✔ Container cris-server-postgres-1         Started                    2.3s
 ✔ Container cris-server-cris-web-server-1  Started                    3.4s
```

Informe final de proyecto de investigación. Año 2023

Guía para el uso de la API pública de ORCID

Que datos es posible obtener de la API de ORCID

No.	Dato	Fuente primaria	API pública de ORCID	API de miembros ORCID
1	Número de personal	Siif	No	No
2	CUI	Formulario	No	No
3	RUI	Autogenerado	No	No
4	ORCID	Formulario	Sí	Sí
5	Colegio profesional	Siif	No	No
6	Colegiado	Siif	No	No
7	Email	Siif	Sí	Sí
8	Primer nombre	Siif	Sí	Sí
9	Segundo nombre	Siif	Sí	Sí
10	Tercer nombre	Siif	Sí	Sí
11	Primer apellido	Siif	Sí	Sí
12	Segundo apellido	Siif	Sí	Sí

13	Apellido de casada	Siif	No	No
14	Grupo étnico	Formulario	No	No
15	Sexo	Siif	No	No
16	Nacionalidad	Siif	Si	Si
17	Idioma materno	Formulario	No	No
18	Fecha de nacimiento	Siif	No	No
19	Teléfono	Siif	No	No
20	Historial académico	ORCID	Sí	Sí
21	Experiencia laboral	Formulario	Sí	Sí
22	Red de investigadores	Formulario	No	Sí
23	Experiencia en investigación	Módulo de investigaciones	No	No
24	Temáticas de trabajo	ORCID	No	Sí
25	Temáticas de interés	ORCID	No	Sí
26	Investigaciones en desarrollo	Módulo de investigaciones	No	No

27	Área de especialización	Formulario	Sí	Sí
28	Activo	Autogenerado	No	No
29	Publicaciones	ORCID	Sí	Si

Documentación consumo de datos de la API de ORCID desde Postman

Formulario para solicitud de credenciales de cliente en el sitio sandbox de ORCID

Datos de la aplicación

Nombre de la aplicación

DSPACE USAC

Nombre mostrado a los usuarios en la pantalla de autorización de OAuth

URL de aplicación

http://127.0.0.1:4000/

Descripción de la aplicación

Integración de ORCID al sistema DSPACE CRIS de la Universidad de San Carlos de Guatemala

Descripción mostrada a los usuarios en la pantalla de autorización de OAuth. Máximo 1000 caracteres.

URI de redireccionamiento

Una vez que el usuario haya autorizado su aplicación, volverá a la URI que usted especifique. Debe indicar estas URI de antemano o los usuarios de su integración experimentarán un error.

Tenga en cuenta lo siguiente:

- Solo **las URI HTTPS** se aceptan en la producción
- Los dominios registrados **DEBEN** coincidir exactamente con los dominios utilizados, incluidos los subdominios
- **Registre todas las URI de redireccionamiento en su totalidad cuando sea posible.** Es la opción más segura y la que recomendamos. Para obtener más información acerca de las URI de redireccionamiento, consulte nuestras preguntas frecuentes sobre URI de redireccionamiento

http://127.0.0.1:8080/server/api/authn/orcid

Si toda la información es correcta, ORCID otorgará las credenciales de cliente

ID de cliente
APP-7GP7TL9JX32ZWE2P

Código secreto del cliente
e8137b87-f5d8-4550-9912-7efe5ff331ab

Generar un nuevo código secreto de cliente

Solicitud de un token de acceso usando las credenciales de cliente en Postman

Realizar una solicitud POST como la siguiente

POST https://sandbox.orcid.org/oauth/token

En el body de la solicitud agregar la siguiente información

Params | Authorization | Headers (8) | Body | Pre-request Script | Tests | Settings

none | form-data | x-www-form-urlencoded | raw | binary | GraphQL

	Key	Value
☑	client_id	APP-7GP7TL9JX32ZWE2P
☑	client_secret	e8137b87-f5d8-4550-9912-7efe5ff331ab
☑	grant_type	client_credentials
☑	scope	/read-public
	Key	Value

Si las credenciales de cliente son correctas y todo lo demás está configurado como la imagen anterior, al realizar la solicitud deberíamos obtener el token de acceso.


```
1 {
2     "access_token": "ee906604-8a86-411a-9208-4fcaabeb053e",
3     "token_type": "bearer",
4     "refresh_token": "3df5e6a6-30b2-4314-8c02-153347b991cb",
5     "expires_in": 631138518,
6     "scope": "/read-public",
7     "orcid": null
8 }
```


Escoger los endpoint según los datos que se desean obtener de la API, puede usarse como guía la siguiente tabla obtenida de este sitio web: https://info.orcid.org/documentation/api-tutorials/api-tutorial-read-data-on-a-record/

Endpoint	Description
/record	Summary view of the full ORCID record
/person	Biographical section of the ORCID record, including through /researcher-urls below
/address	The researcher,s countries or regions
/email	The email address(es) associated with the record
/external-identifiers	Linked external identifiers in other systems
/keywords	Keywords related to the researcher and their work

/other-names	Other names by which the researcher is know
/personal-details	Personal details: the researcher,s name, credit (published) name, and biography
/researcher-urls	Links to the researcher,s personal or profile pages
/activities	Summary of the activities section of the ORCID record, including through /works below.
/educations	Education affiliations
/employments	Employment affiliations
/fundings	Summary of funding activities
/peer-reviews	Summary of peer review activities
/works	Summary of research works
/research-resources	Summary of research resources
/services	Summary of services
/qualifications	Summary of qualifications
/memberships	Summary of memberships
/distinctions	Summary of distinctions
/invited-positions	Summary of invited positions

El header completo de la solicitud debería verse como esto. Todas las solicitudes deben tener la siguiente configuración en el header para poder obtener la información.

Params | Authorization | Headers (7) | Body | Pre-request Script | Tests | Settings

Headers | Hide auto-generated headers

	Key	Value
☑	Postman-Token	<calculated when request is sent>
☑	Host	<calculated when request is sent>
☑	User-Agent	PostmanRuntime/7.33.0
☑	Accept	*/*
☑	Accept-Encoding	gzip, deflate, br
☑	Connection	keep-alive
☑	Authorization	Bearer 7176d9d9-e394-4d39-a2a2-737b9865e042

Asegurarse de escribir correctamente el token luego de la palabra Bearer.

☑	Authorization	Bearer 7176d9d9-e394-4d39-a2a2-737b9865e042

Las URL para realizar las peticiones de datos a la API de ORCID deben tener la siguiente estructura:

https://pub.sandbox.orcid.org/v3.0/0009-0005-9587-5049/endpoint

Donde **https://pub.sandbox.orcid.org/v3.0** no debe modificarse.

Sustituir **/0009-0005-9587-5049/** por su clave propia de ORCID.

Sustituir **/endpoint** por el endpoit del cual deseamos obtener información

Ejemplos de solicitudes a endpoints

Solicitud GET para obtener todos los datos personales del perfil

https://pub.sandbox.orcid.org/v3.0/0009-0005-9587-5049/person

Informe final de proyecto de investigación. Año 2023

Resultado de la solicitud

Body Cookies Headers (17) Test Results Status: 200 OK Time: 658 ms Size: 5.08 KB Save as Example

Pretty Raw Preview Visualize XML


```xml
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<person:person path="/0009-0005-9587-5049/person" xmlns:internal="http://www.orcid.org/ns/internal" xmlns:education="http://www.orcid.org/ns/edu
    <common:last-modified-date>2023-10-04T21:11:09.916Z</common:last-modified-date>
    <person:name visibility="public" path="0009-0005-9587-5049">
        <common:created-date>2023-08-16T19:59:44.289Z</common:created-date>
        <common:last-modified-date>2023-08-23T23:24:18.294Z</common:last-modified-date>
        <personal-details:given-names>Jeyson</personal-details:given-names>
        <personal-details:family-name>Martínez</personal-details:family-name>
    </person:name>
    <other-name:other-names path="/0009-0005-9587-5049/other-names"/>
    <person:biography visibility="public" path="/0009-0005-9587-5049/biography">
        <common:created-date>2023-09-06T22:25:17.566Z</common:created-date>
        <common:last-modified-date>2023-09-06T22:25:17.566Z</common:last-modified-date>
        <personal-details:content>Estudiante de Ingeniería en Ciencias y Sistemas</personal-details:content>
    </person:biography>
    <researcher-url:researcher-urls path="/0009-0005-9587-5049/researcher-urls"/>
    <email:emails path="/0009-0005-9587-5049/email">
        <common:last-modified-date>2023-08-17T18:54:24.445Z</common:last-modified-date>
        <email:email visibility="public" verified="true" primary="true">
            <common:created-date>2023-08-16T20:30:37.420Z</common:created-date>
            <common:last-modified-date>2023-08-17T18:54:24.445Z</common:last-modified-date>
            <common:source>
                <common:source-orcid>
                    <common:uri>https://sandbox.orcid.org/0009-0005-9587-5049</common:uri>
                    <common:path>0009-0005-9587-5049</common:path>
                    <common:host>sandbox.orcid.org</common:host>
                </common:source-orcid>
                <common:source-name>Jeyson Martínez</common:source-name>
            </common:source>
            <email:email>jeyson.pruebas@mailinator.com</email:email>
        </email:email>
    </email:emails>
    <address:addresses path="/0009-0005-9587-5049/address">
        <common:last-modified-date>2023-10-04T21:11:09.916Z</common:last-modified-date>
        <address:address put-code="7805" visibility="public" path="/0009-0005-9587-5049/address/7805" display-index="1">
            <common:created-date>2023-10-04T21:11:09.916Z</common:created-date>
            <common:last-modified-date>2023-10-04T21:11:09.916Z</common:last-modified-date>
            <common:source>
                <common:source-orcid>
                    <common:uri>https://sandbox.orcid.org/0009-0005-9587-5049</common:uri>
                    <common:path>0009-0005-9587-5049</common:path>
                    <common:host>sandbox.orcid.org</common:host>
                </common:source-orcid>
                <common:source-name>Jeyson Martínez</common:source-name>
            </common:source>
            <address:country>GT</address:country>
        </address:address>
    </address:addresses>
    <keyword:keywords path="/0009-0005-9587-5049/keywords"/>
    <external-identifier:external-identifiers path="/0009-0005-9587-5049/external-identifiers"/>
</person:person>
```

Solicitud GET para obtener el correo del perfil

https://pub.sandbox.orcid.org/v3.0/0009-0005-9587-5049/email

Resultado de la solicitud

Body Cookies Headers (17) Test Results Status: 200 OK Time: 529 ms Size: 2.85 KB Save as Example

Pretty Raw Preview Visualize XML


```xml
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<email:emails path="/0009-0005-9587-5049/email" xmlns:internal="http://www.orcid.org/ns/internal" xmlns:education="http://www.orcid.org/ns/educa
    <common:last-modified-date>2023-08-17T18:54:24.445Z</common:last-modified-date>
    <email:email visibility="public" verified="true" primary="true">
        <common:created-date>2023-08-16T20:30:37.420Z</common:created-date>
        <common:last-modified-date>2023-08-17T18:54:24.445Z</common:last-modified-date>
        <common:source>
            <common:source-orcid>
                <common:uri>https://sandbox.orcid.org/0009-0005-9587-5049</common:uri>
                <common:path>0009-0005-9587-5049</common:path>
                <common:host>sandbox.orcid.org</common:host>
            </common:source-orcid>
            <common:source-name>Jeyson Martínez</common:source-name>
        </common:source>
        <email:email>jeyson.pruebas@mailinator.com</email:email>
    </email:email>
</email:emails>
```


En el perfil de ORCID el correo (email) se visualiza de la siguiente manera

Solicitud GET para obtener dirección del perfil

https://pub.sandbox.orcid.org/v3.0/0009-0005-9587-5049/address

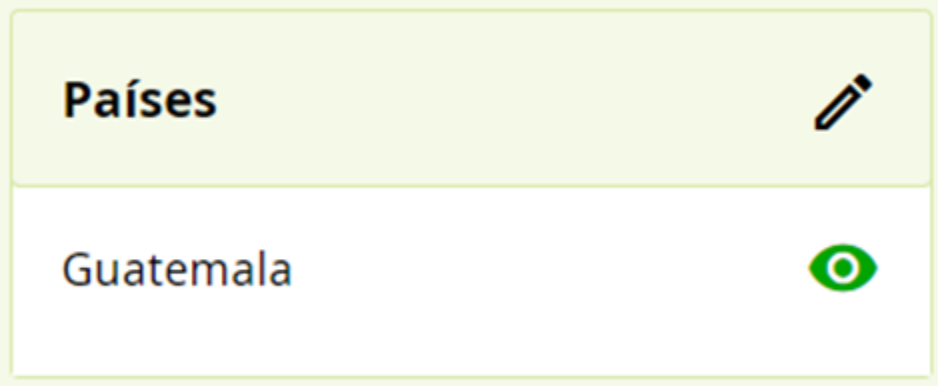
Resultado de la solicitud

Body Cookies Headers (17) Test Results Status: 200 OK Time: 350 ms Size: 2.89 KB Save as Example

Pretty Raw Preview Visualize XML


```xml
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<address:addresses path="/0009-0005-9587-5049/address" xmlns:internal="http://www.orcid.org/ns/internal" xmlns:education="http://www.orcid.org/n
    <common:last-modified-date>2023-10-04T21:11:09.916Z</common:last-modified-date>
    <address:address put-code="7805" visibility="public" path="/0009-0005-9587-5049/address/7805" display-index="1">
        <common:created-date>2023-10-04T21:11:09.916Z</common:created-date>
        <common:last-modified-date>2023-10-04T21:11:09.916Z</common:last-modified-date>
        <common:source>
            <common:source-orcid>
                <common:uri>https://sandbox.orcid.org/0009-0005-9587-5049</common:uri>
                <common:path>0009-0005-9587-5049</common:path>
                <common:host>sandbox.orcid.org</common:host>
            </common:source-orcid>
            <common:source-name>Jeyson Martínez</common:source-name>
        </common:source>
        <address:country>GT</address:country>
    </address:address>
</address:addresses>
```


En el perfil de ORCID la dirección (address) se visualiza de la siguiente manera



Solicitud GET para obtener la educación del perfil

https://pub.sandbox.orcid.org/v3.0/0009-0005-9587-5049/educations

Resultado de la solicitud

Body Cookies Headers (17) Test Results Status: 200 OK Time: 383 ms Size: 4.08 KB Save as Example

Pretty Raw Preview Visualize XML


```xml
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<activities:educations path="/0009-0005-9587-5049/educations" xmlns:internal="http://www.orcid.org/ns/internal" xmlns:education="http://www.orci
    <common:last-modified-date>2023-10-04T20:59:18.267Z</common:last-modified-date>
    <activities:affiliation-group>
        <common:last-modified-date>2023-10-04T20:59:18.267Z</common:last-modified-date>
        <common:external-ids/>
        <education:education-summary put-code="58629" display-index="1" path="/0009-0005-9587-5049/education/58629" visibility="public">
            <common:created-date>2023-10-04T20:59:18.267Z</common:created-date>
            <common:last-modified-date>2023-10-04T20:59:18.267Z</common:last-modified-date>
            <common:source>
                <common:source-orcid>
                    <common:uri>https://sandbox.orcid.org/0009-0005-9587-5049</common:uri>
                    <common:path>0009-0005-9587-5049</common:path>
                    <common:host>sandbox.orcid.org</common:host>
                </common:source-orcid>
                <common:source-name>Jeyson Martínez</common:source-name>
            </common:source>
            <common:department-name>Carreras de Ingeniería</common:department-name>
            <common:role-title>Licenciatura en Ingeniería en Ciencias y Sistemas</common:role-title>
            <common:start-date>
                <common:year>2018</common:year>
                <common:month>01</common:month>
                <common:day>15</common:day>
            </common:start-date>
            <common:end-date>
                <common:year>2023</common:year>
                <common:month>11</common:month>
                <common:day>30</common:day>
            </common:end-date>
            <common:organization>
                <common:name>Centro Universitario de Oriente</common:name>
                <common:address>
                    <common:city>Chiquimula</common:city>
                    <common:region>Chiquimula</common:region>
                    <common:country>GT</common:country>
                </common:address>
            </common:organization>
        </education:education-summary>
    </activities:affiliation-group>
</activities:educations>
```


En el perfil de ORCID la educación (educations) se visualiza de la siguiente manera

Informe final de proyecto de investigación. Año 2023

Educación y titulaciones (1) Agregar Ordenar

Centro Universitario de Oriente: Chiquimula, Chiquimula, GT Público

2018-01-15 hasta 2023-11-30 | Licenciatura en Ingeniería en Ciencias y Sistemas (Carreras de Ingeniería) Mostrar más detalles
Educación

Fuente: Jeyson Martínez

Solicitud GET para obtener las obras (publicaciones) del perfil

https://pub.sandbox.orcid.org/v3.0/0009-0005-9587-5049/works

Resultado de la solicitud


```xml
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<activities:works path="/0009-0005-9587-5049/works" xmlns:internal="http://www.orcid.org/ns/internal" xmlns:education="http://www.orcid.org/ns
    <common:last-modified-date>2023-10-05T20:37:49.558Z</common:last-modified-date>
    <activities:group>
        <common:last-modified-date>2023-10-05T20:37:49.558Z</common:last-modified-date>
        <common:external-ids>
            <common:external-id>
                <common:external-id-type>doi</common:external-id-type>
                <common:external-id-value>10.23640/07243.c.4232246</common:external-id-value>
                <common:external-id-normalized transient="true">10.23640/07243.c.4232246</common:external-id-normalized>
                <common:external-id-url>https://doi.org/10.23640/07243.c.4232246</common:external-id-url>
                <common:external-id-relationship>self</common:external-id-relationship>
            </common:external-id>
        </common:external-ids>
        <work:work-summary put-code="1798284" path="/0009-0005-9587-5049/work/1798284" visibility="public" display-index="1">
            <common:created-date>2023-10-04T21:10:54.472Z</common:created-date>
            <common:last-modified-date>2023-10-05T20:37:49.558Z</common:last-modified-date>
            <common:source>
                <common:source-orcid>
                    <common:uri>https://sandbox.orcid.org/0009-0005-9587-5049</common:uri>
                    <common:path>0009-0005-9587-5049</common:path>
                    <common:host>sandbox.orcid.org</common:host>
                </common:source-orcid>
                <common:source-name>Jeyson Martínez</common:source-name>
            </common:source>
            <work:title>
```

Informe final de proyecto de investigación. Año 2023


```xml
            <common:title>Integración de ORCID con DSPACE-CRIS</common:title>
        </work:title>
        <common:external-ids>
            <common:external-id>
                <common:external-id-type>doi</common:external-id-type>
                <common:external-id-value>10.23640/07243.c.4232246</common:external-id-value>
                <common:external-id-normalized transient="true">10.23640/07243.c.4232246</common:external-id-normalized>
                <common:external-id-url>https://doi.org/10.23640/07243.c.4232246</common:external-id-url>
                <common:external-id-relationship>self</common:external-id-relationship>
            </common:external-id>
        </common:external-ids>
        <work:type>journal-article</work:type>
        <common:publication-date>
            <common:year>2023</common:year>
            <common:month>07</common:month>
            <common:day>10</common:day>
        </common:publication-date>
        <work:journal-title>Perspectiva</work:journal-title>
      </work:work-summary>
    </activities:group>
</activities:works>
```


En el perfil de ORCID las obras (works) se visualizan de la siguiente manera

Informe final de proyecto de investigación. Año 2023

Proceso para instalar VIVO v1.8.1 con contenedores de Docker

Fuentes consultadas para la instalación

Documentación oficial

https://wiki.lyrasis.org/display/VIVO/Docker+for+VIVO

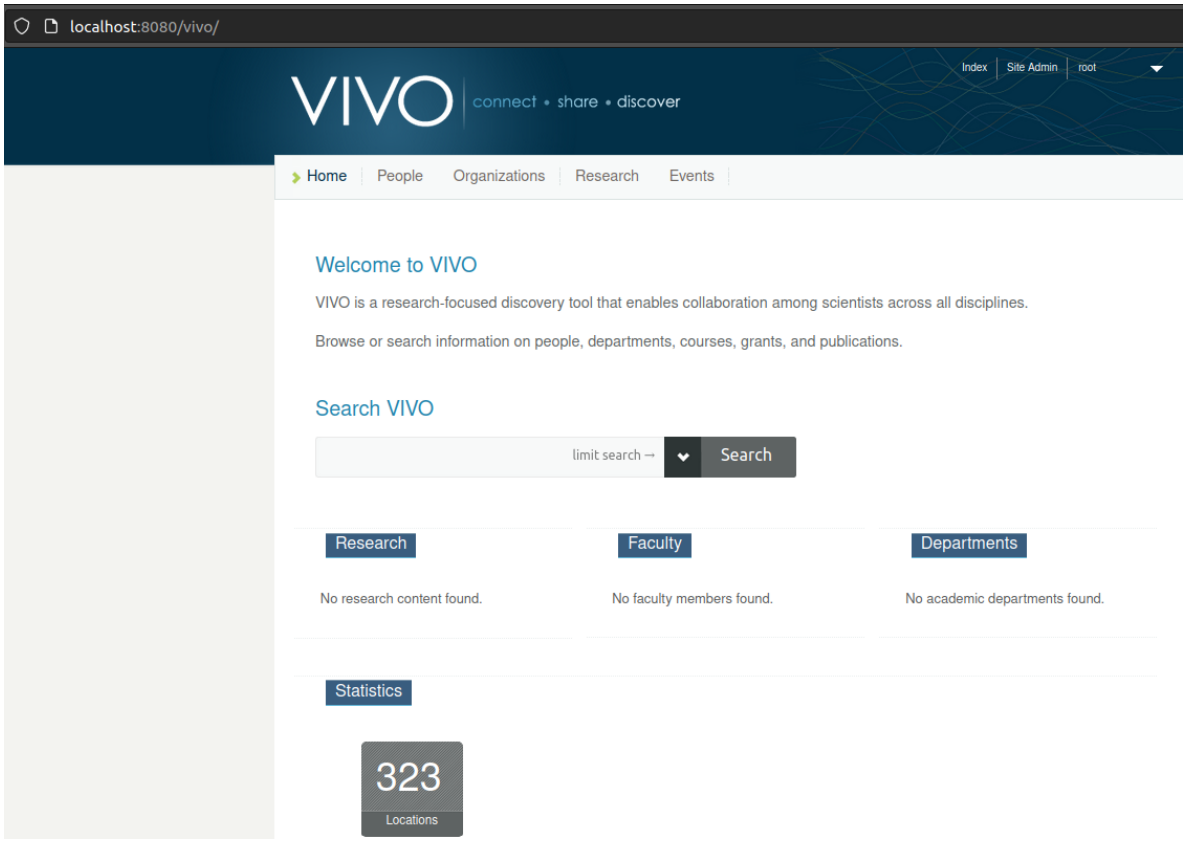
Repositorio de GitHub

https://github.com/gwu-libraries/vivo-docker

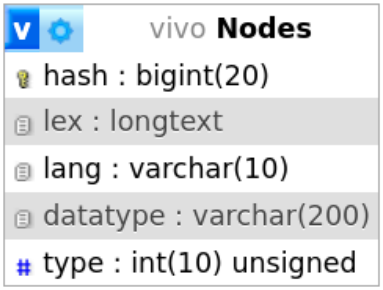
Estado de la instalación: OK

REPOSITORY	TAG	IMAGE ID	CREATED	SIZE
phpmyadmin/phpmyadmin	latest	e4be0d1537f5	3 months ago	517MB
gwul/vivo_app	1.8.1	b8b8e510f193	6 years ago	1.43GB
gwul/vivo_tomcat	latest	72b58fe34e7e	6 years ago	555MB
gwul/vivo_db	latest	43952fca7ebf	6 years ago	327MB

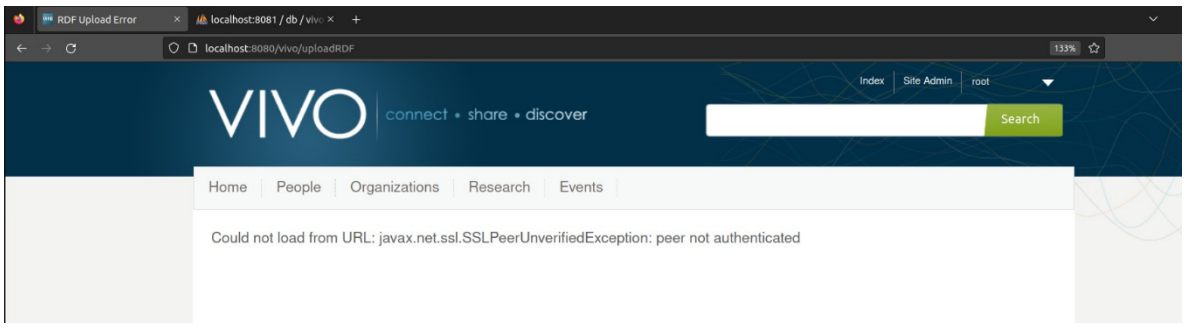
CONTAINER ID	IMAGE	COMMAND	CREATED	STATUS	PORTS	NAMES
277fea4edbd8	phpmyadmin/phpmyadmin	"/docker-entrypoint.…"	26 minutes ago	Up 26 minutes	0.0.0.0:8081->80/tcp, :::8081->80/tcp	vivo_phpmyadmin_1
c706f9c5f49a	gwul/vivo_tomcat	"/bin/sh -c 'python …"	26 minutes ago	Up 26 minutes	0.0.0.0:8080->8080/tcp, :::8080->8080/tcp	vivo_vivotomcat_1
4b1d24939fb3	gwul/vivo_db	"docker-entrypoint.s…"	26 minutes ago	Up 26 minutes	3306/tcp	vivo_vivodb_1



Base de Datos

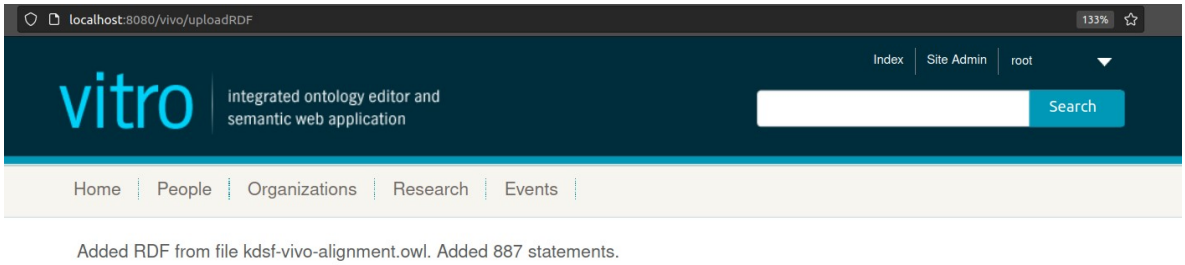


Error anterior al cargar los Datos de muestra



El error se presentó cuando se cargaron los datos por medio de un enlace.

Cargando datos de muestra por medio de un archivo .owl



Fue el primer intento exitoso de cargar datos, pero los datos no se reflejaban en la página, había que agregarlos manualmente uno por uno.

Datos de Muestra

Fuentes consultadas para la prueba de datos de muestra

Documentación oficial: https://wiki.lyrasis.org/display/VIVODOC114x/Sample+Data

Datos de muestra:
https://raw.githubusercontent.com/vivo-project/sample-data/master/sample-data.n3

Fuentes de terceros:

https://av.tib.eu/search?q=vivo

https://github.com/VIVO-DE/VIVO-KDSF-Integration

https://github.com/hudajkhan/VIVO17Output

Los datos de la muestra incluyen una universidad ficticia, departamentos ficticios, profesores y colaboradores ficticios, publicaciones y subvenciones ficticias, membresías y otros elementos académicos comunes en VIVO. Los datos de muestra no incluyen un ejemplo de todo tipo de cosas que se pueden almacenar en VIVO, ni contienen muchas facultades. Está destinado a demostrar los elementos más comunes de VIVO.

- No se agregaron los datos de muestra a nuestros datos.

- VIVO uso datos de una base de datos que se especificó.
- Se creo una base de datos para los datos de muestra y se le dijo a VIVO que use esa base de datos cuando cargue y use los datos de muestra.

1. Ingresar al contenedor de Docker en donde se instaló la base de datos.

```
docker exec -it vivo_vivodb_1 /bin/bash
```

2. *Crear una nueva base de datos de muestra en MySQL*

```
$ mysql -u root -p

Welcome to the MySQL monitor.  Commands end with ; or \\g.

Your MySQL connection id is 2

Server version: 5.7.9 MySQL Community Server (GPL)

Copyright (c) 2000, 2015, Oracle and/or its affiliates. All rights reserved.

Oracle is a registered trademark of Oracle Corporation and/or its

affiliates. Other names may be trademarks of their respective

owners.

Type 'help;' or '\\h' for help. Type '\\c' to clear the current input statement.

mysql> CREATE DATABASE sampledb CHARACTER SET utf8;

mysql> GRANT ALL ON sampledb.* TO 'vivo_user'@'localhost' IDENTIFIED
BY 'ghE35fkEm4dfe3';
```

```
root@373ca1db79ab:/# mysql -u root -p
Enter password:
Welcome to the MySQL monitor.  Commands end with ; or \g.
Your MySQL connection id is 53
Server version: 5.6.34 MySQL Community Server (GPL)

Copyright (c) 2000, 2016, Oracle and/or its affiliates. All rights reserved.

Oracle is a registered trademark of Oracle Corporation and/or its
affiliates. Other names may be trademarks of their respective
owners.

Type 'help;' or '\h' for help. Type '\c' to clear the current input statement.

mysql> show databases;
+--------------------+
| Database           |
+--------------------+
| information_schema |
| mysql              |
| performance_schema |
| vivo               |
+--------------------+
4 rows in set (0.00 sec)

mysql> CREATE DATABASE sampledb CHARACTER SET utf8;
Query OK, 1 row affected (0.00 sec)

mysql> GRANT ALL ON sampledb.* TO 'vitrodbUsername'@'localhost' IDENTIFIED BY 'vitrodbPassword';
Query OK, 0 rows affected (0.00 sec)

mysql> show databases;
+--------------------+
| Database           |
+--------------------+
| information_schema |
| mysql              |
| performance_schema |
| sampledb           |
| vivo               |
+--------------------+
5 rows in set (0.00 sec)
```

3. Ingresar al contenedor de Docker en donde se instalo Tomcat.

```
docker exec -it vivo_vivotomcat_1 /bin/bash
```

4. Descargar e instalar el editor de archivos "nano".

   Link: http://deb.debian.org/debian/pool/main/n/nano/

```
wget <http://deb.debian.org/debian/pool/main/n/nano/nano-tiny_3.2-3_amd64.deb>
dpkg -i nano-tiny_3.2-3_amd64.deb
```

```
root@04c56a0ea5e3:/usr/local/vivo# wget http://deb.debian.org/debian/pool/main/n/nano/nano-tiny_3.2-3_amd64.deb
--2023-07-23 20:43:49--  http://deb.debian.org/debian/pool/main/n/nano/nano-tiny_3.2-3_amd64.deb
Resolving deb.debian.org (deb.debian.org)... 146.75.34.132, 2a04:4e42:79::644
Connecting to deb.debian.org (deb.debian.org)|146.75.34.132|:80... connected.
HTTP request sent, awaiting response... 200 OK
Length: 101180 (99K) [application/vnd.debian.binary-package]
Saving to: 'nano-tiny_3.2-3_amd64.deb'

nano-tiny_3.2-3_amd64.deb                      100%[===================================================================================================>]  98.81K  --.-KB/s   in 0.1s

2023-07-23 20:43:49 (919 KB/s) - 'nano-tiny_3.2-3_amd64.deb' saved [101180/101180]

root@04c56a0ea5e3:/usr/local/vivo# dpkg -i nano-tiny_3.2-3_amd64.deb
Selecting previously unselected package nano-tiny.
(Reading database ... 19035 files and directories currently installed.)
Preparing to unpack nano-tiny_3.2-3_amd64.deb ...
Unpacking nano-tiny (3.2-3) ...
Setting up nano-tiny (3.2-3) ...
update-alternatives: using /bin/nano-tiny to provide /usr/bin/editor (editor) in auto mode
```

5. Dirigirse a la carpeta en donde se encuentran los archivos de configuración de VIVO. Editar el archivo runtime.properties para decirle a VIVO el nombre de la base de datos que debe usar. En este caso sampledb.

```
cd /usr/local/vivo/home
nano-tiny runtime.properties
```

```
root@04c56a0ea5e3:/usr/local/vivo/home# ls -l
total 44
drwxrwsrwx 2 root staff  4096 Jul 23 20:30 config
-rw-rw-rw- 1 root staff 14347 Jul 23 20:30 example.runtime.properties
drwxrwsrwx 9 root staff  4096 Jul 23 20:30 rdf
-rw-r--r-- 1 root staff  1240 Jul 23 20:30 runtime.properties
drwxrwsrwx 4 root staff  4096 Jul 23 20:37 solr
drwxr-sr-x 2 root staff  4096 Jul 23 20:30 tdbModels
drwxr-sr-x 3 root staff  4096 Jul 23 20:35 upgrade
drwxr-sr-x 3 root staff  4096 Jul 23 20:30 uploads
```

```
  GNU nano 3.2                                                    runtime.properties

#
#Sun Jul 23 20:30:29 UTC 2023
Vitro.defaultNamespace=http\://vivo.gwu.edu/individual/
rootUser.emailAddress=vivo_root@gwu.edu
VitroConnection.DataSource.url=jdbc\:mysql\://db/sampledb
VitroConnection.DataSource.username=vivo_user
VitroConnection.DataSource.password=ghE35fkEm4dfe3
email.smtpHost=
email.replyTo=
vitro.local.solr.url=http\://localhost\:8080/vivosolr
selfEditing.idMatchingProperty=http\://vivo.gwu.edu/ns#networkId
VitroConnection.DataSource.pool.maxActive=40
VitroConnection.DataSource.pool.maxIdle=10
VitroConnection.DataSource.dbtype=MySQL
VitroConnection.DataSource.driver=com.mysql.jdbc.Driver
VitroConnection.DataSource.validationQuery=SELECT 1
visualization.temporal=enabled
proxy.eligibleTypeList=http\://xmlns.com/foaf/0.1/Person, http\://xmlns.com/foaf/0.1/Organization
Vitro.reconcile.defaultTypeList=http\://vivoweb.org/ontology/core#Role, core\:Role; http\://vivoweb.org/ontology/core#AcademicDegree, core\:Academic Degree; http\://purl.org/NET/c4dm/event.owl#Event, evS
```

6. Reiniciar los contenedores en donde se instalo VIVO.

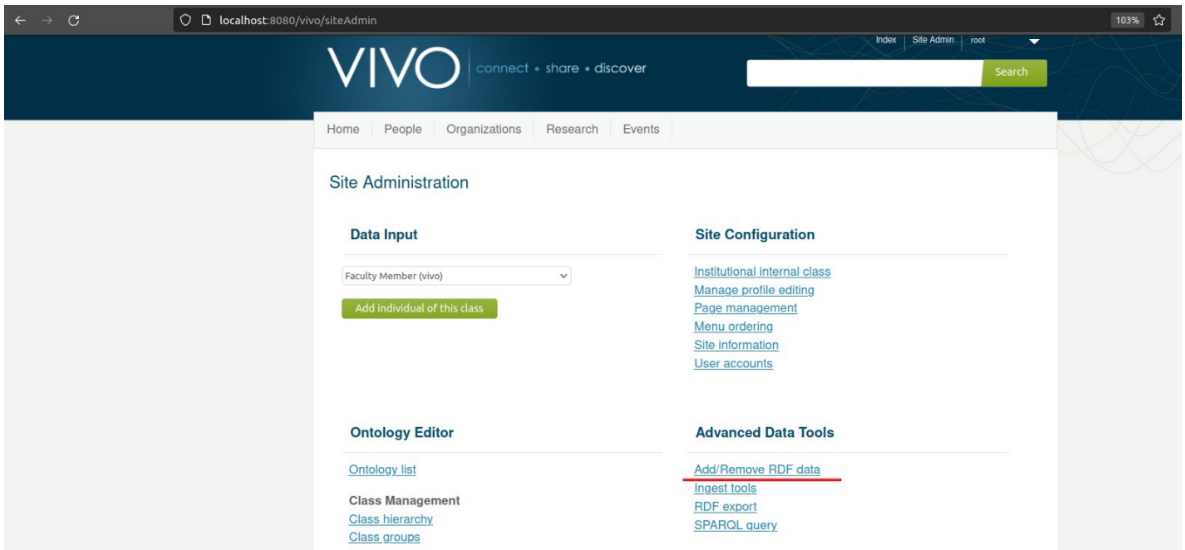
```
ubuntu22-04@ubuntu2204-virtual-machine:~/Escritorio/ViVo$ docker ps
CONTAINER ID   IMAGE                 COMMAND                  CREATED          STATUS                       PORTS                                          NAMES
04c56a0ea5e3   gwul/vivo_tomcat      "/bin/sh -c 'python …"   22 minutes ago   Up 22 minutes                0.0.0.0:8080->8080/tcp, :::8080->8080/tcp      vivo_vivotomcat_1
78247024ab90   gwul/vivo_app:1.8.1   "/bin/sh -c '[ ! -f …"   22 minutes ago   Restarting (1) 17 seconds ago                                               vivo_vivoapp_1
7d00c47d378f   phpmyadmin/phpmyadmin "/docker-entrypoint.…"   22 minutes ago   Up 22 minutes                0.0.0.0:8081->80/tcp, :::8081->80/tcp          vivo_phpmyadmin_1
373ca1db79ab   gwul/vivo_db          "docker-entrypoint.s…"   22 minutes ago   Up 22 minutes                3306/tcp                                       vivo_vivodb_1
ubuntu22-04@ubuntu2204-virtual-machine:~/Escritorio/ViVo$ docker stop 373ca1db79ab 04c56a0ea5e3 78247024ab90 373ca1db79ab
373ca1db79ab
04c56a0ea5e3
78247024ab90
373ca1db79ab
ubuntu22-04@ubuntu2204-virtual-machine:~/Escritorio/ViVo$ docker ps
CONTAINER ID   IMAGE                 COMMAND                  CREATED          STATUS          PORTS                                          NAMES
7d00c47d378f   phpmyadmin/phpmyadmin "/docker-entrypoint.…"   23 minutes ago   Up 23 minutes   0.0.0.0:8081->80/tcp, :::8081->80/tcp          vivo_phpmyadmin_1
ubuntu22-04@ubuntu2204-virtual-machine:~/Escritorio/ViVo$ docker stop 7d00c47d378f
7d00c47d378f
ubuntu22-04@ubuntu2204-virtual-machine:~/Escritorio/ViVo$ docker ps
CONTAINER ID   IMAGE     COMMAND   CREATED   STATUS    PORTS     NAMES
ubuntu22-04@ubuntu2204-virtual-machine:~/Escritorio/ViVo$ docker restart 373ca1db79ab 04c56a0ea5e3 78247024ab90 7d00c47d378f
373ca1db79ab
04c56a0ea5e3
78247024ab90
7d00c47d378f
ubuntu22-04@ubuntu2204-virtual-machine:~/Escritorio/ViVo$ docker ps
CONTAINER ID   IMAGE                 COMMAND                  CREATED          STATUS                                  PORTS                                          NAMES
04c56a0ea5e3   gwul/vivo_tomcat      "/bin/sh -c 'python …"   24 minutes ago   Up 3 seconds                            0.0.0.0:8080->8080/tcp, :::8080->8080/tcp      vivo_vivotomcat_1
78247024ab90   gwul/vivo_app:1.8.1   "/bin/sh -c '[ ! -f …"   24 minutes ago   Restarting (1) Less than a second ago                                          vivo_vivoapp_1
7d00c47d378f   phpmyadmin/phpmyadmin "/docker-entrypoint.…"   24 minutes ago   Up 2 seconds                            0.0.0.0:8081->80/tcp, :::8081->80/tcp          vivo_phpmyadmin_1
373ca1db79ab   gwul/vivo_db          "docker-entrypoint.s…"   24 minutes ago   Up 4 seconds                            3306/tcp                                       vivo_vivodb_1
```

Al completar estos pasos, VIVO estará utilizando una base de datos "vacía".

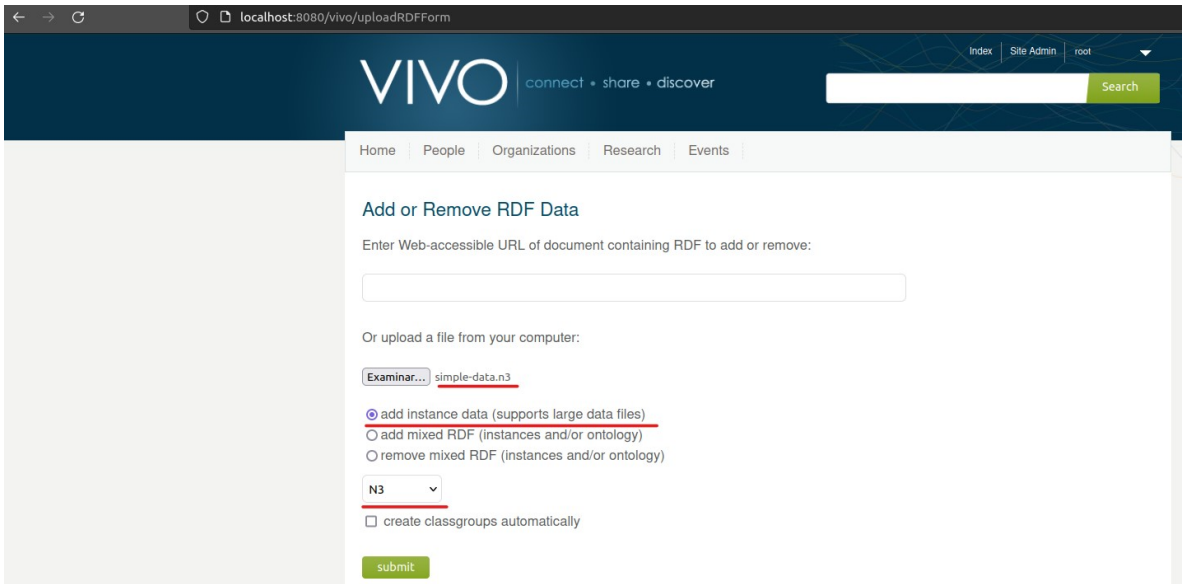
Esperar a que VIVO se inicie y actualizar el navegador, ya que VIVO carga automáticamente los datos de los archivos distribuidos con VIVO. VIVO puede tardar varios minutos en reiniciarse, cargar los datos de la primera vez e indexarlos. Cuando VIVO haya completado su trabajo y se haya actualizado el navegador, la página de inicio de VIVO se verá igual cuando se instaló por primera vez.

Cargando los datos de muestra

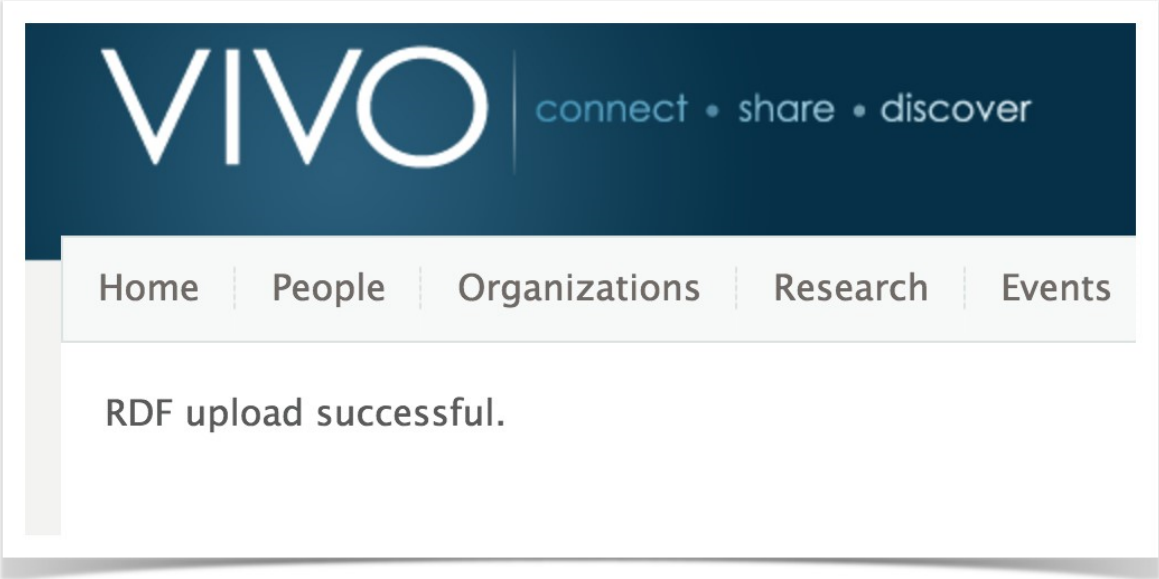
1. Iniciar sesión en VIVO como administrador del sitio.
2. Ir a Site Admin / Add / Remove RDF Data



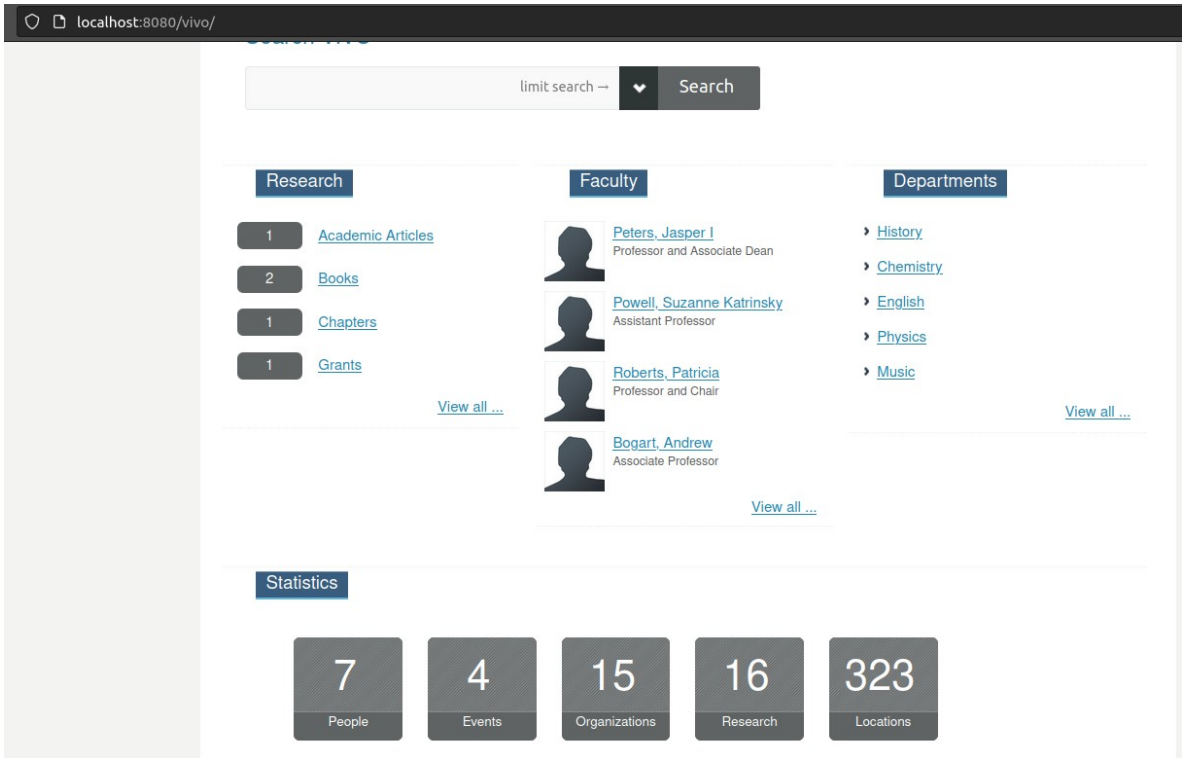
3. Guardar en un archivo los datos de muestra de la dirección siguiente: https://raw.githubusercontent.com/vivo-project/sample-data/master/sample-data.n3
4. Dar clic en "Examinar" y seleccionar el archivo que se guardo anteriormente.
5. Seleccionar "agregar datos de instancia" y establezca el tipo de archivo en "N3". Presionar Enviar.



6. Después de una breve carga, se verá:

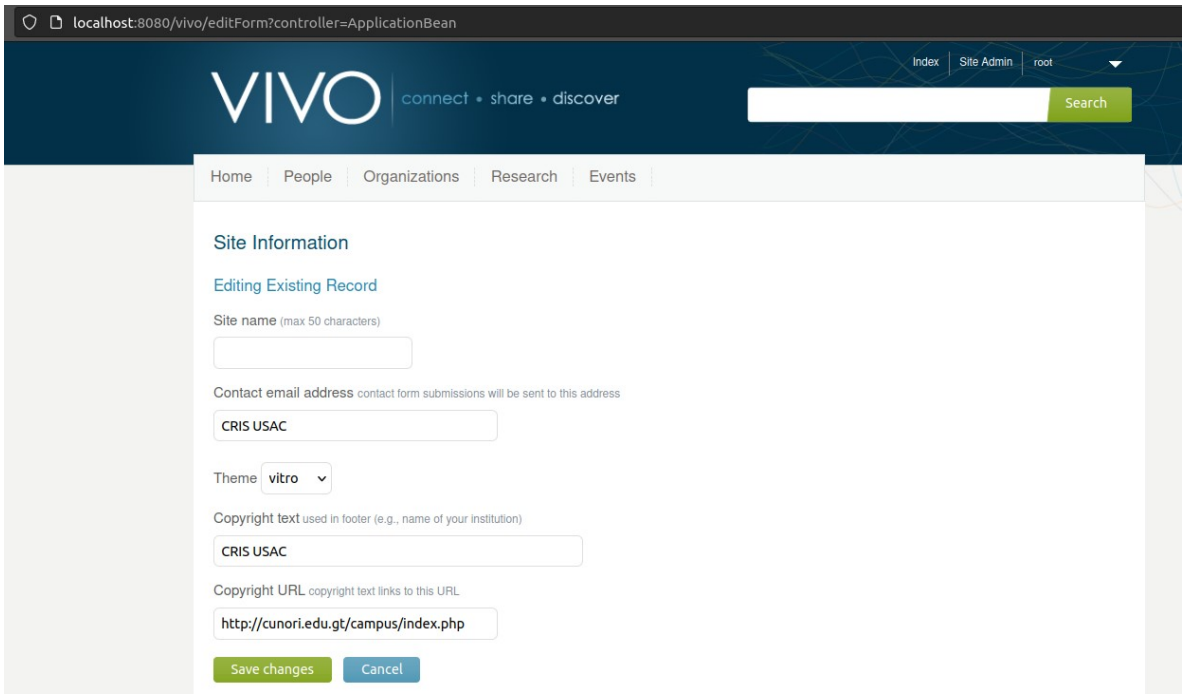


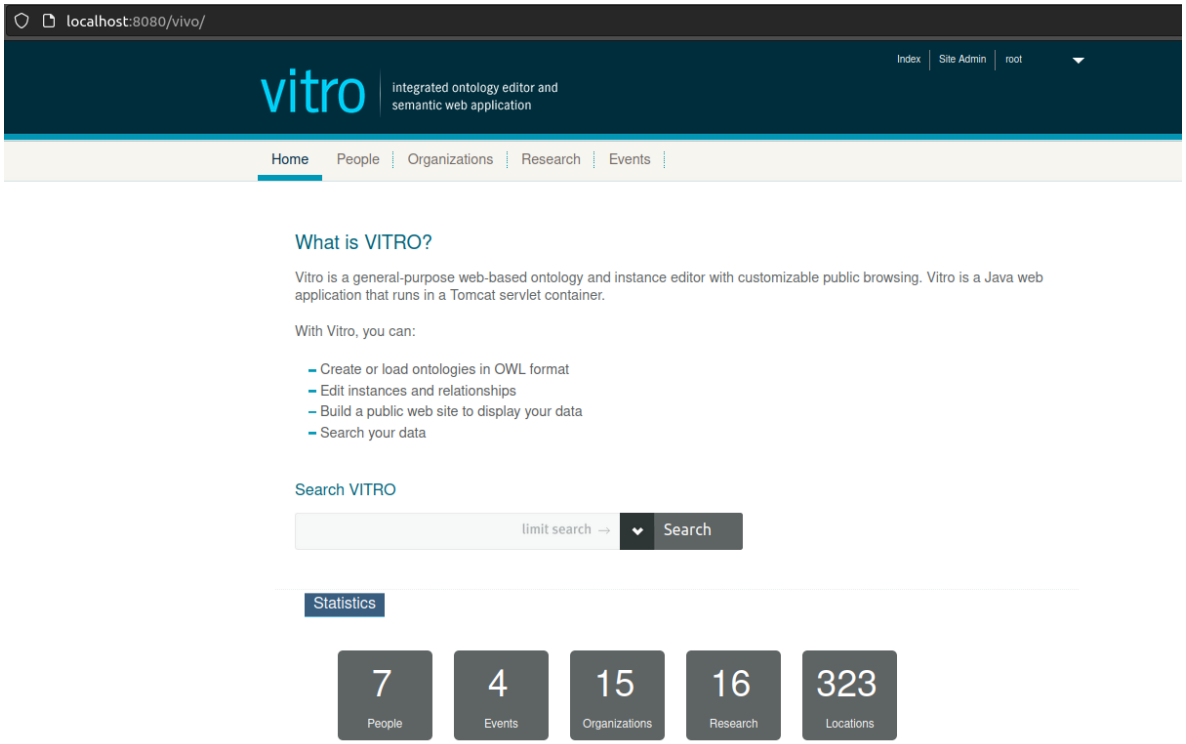
7. Navegar a la página de inicio. ¡Eso es todo! Comenzar a explorar.



Personalización

1. Ir a Site Admin / Site Configuration / Site Information





DSpace-VIVO

Como referencia, VIVO anteriormente se instaló exitosamente utilizando contenedores de Docker.

Además por aparte se instaló exitosamente DSpace-CRIS también utilizando Docker.

Cómo seguimiento a este proceso:

Ahora se procede a experimentar usando DSpace y ViVO.

Tomando la siguiente ruta de guia:

Documentación Oficial:

https://wiki.lyrasis.org/display/VIVO/DSpace-VIVO

Guía de Instalación:

https://wiki.lyrasis.org/display/VIVO/Migrate+DSpace-Metadata+from+DSpace+to+VIVO

Repositorio en GitHub:

https://github.com/vivo-community/DSpace-VIVO

Esta página presenta el procedimiento para migrar datos de DSpace a VIVO.

Requisitos de Software (Instalación local)

- Linux Ubuntu
- Linux bash
- jdk 11
- maven 3.6.3
- solr y tomcat

Configuración de los recursos necesarios para ejecutar VIVO

Configuración del proyecto

```
mkdir -p ~/dspace-vivo-prj/00-GIT
cd ~/dspace-vivo-prj/00-GIT
```

Recuperar el código fuente

```
git clone --depth 1 --branch Beta-1.1 <https://github.com/vivo-community/DSpace-VIVO>
```

Instalar Solr + Tomcat

```
./DSpace-VIVO/releng/org.vivoweb.dspacevivo.installer/00-INIT/install-tomcat-solr-app.sh
```

Nota: Antes de ejecutar está instrucción se debe de colocar una versión reciente de Tomcat, la versión que tiene el archivo por defecto no existe en el repositorio de donde se descarga.

```
export SCRIPT_DIR="$( cd "$( dirname "${BASH_SOURCE[0]}" )" &> /dev/null && pwd -P)"
source $SCRIPT_DIR/../00-env.sh
cd $DEPLOY
TOMCAT_VER=9.0.78
SOLR_VER=8.11.2

######################
# Downloading packages
#
wget    https://dlcdn.apache.org/tomcat/tomcat-9/v$TOMCAT_VER/bin/apache-tomcat-$TOMCAT_VER.tar.gz \
        https://dlcdn.apache.org/lucene/solr/$SOLR_VER/solr-$SOLR_VER.tgz
```

Abrir y editar el archivo: install-tomcat-solr-app.sh

Colocar la versión que se indica.

Instalando/compilando VIVO

```
./DSpace-VIVO/releng/org.vivoweb.dspacevivo.installer/01-VIVO/vivo-git-clone.sh

./DSpace-VIVO/bundles/org.vivoweb.dspacevivo/script/vivo-compile-and-deploy-for-tomcat.sh
```

Run - Start/Stop VIVO

Starting VIVO

```
cd ./DSpace-VIVO/bundles/org.vivoweb.dspacevivo/script
```

source ./00-env.sh

solr-start.sh

tomcat-start.sh

**Para mostrar VIVO en un navegador web (<http://localhost:8080/vivo-dspace/>)

y SOLR (<http://localhost:8983/solr/#/>)**

browse-vivo.sh

Por detener VIVO

tomcat-stop.sh

solr-stop.sh

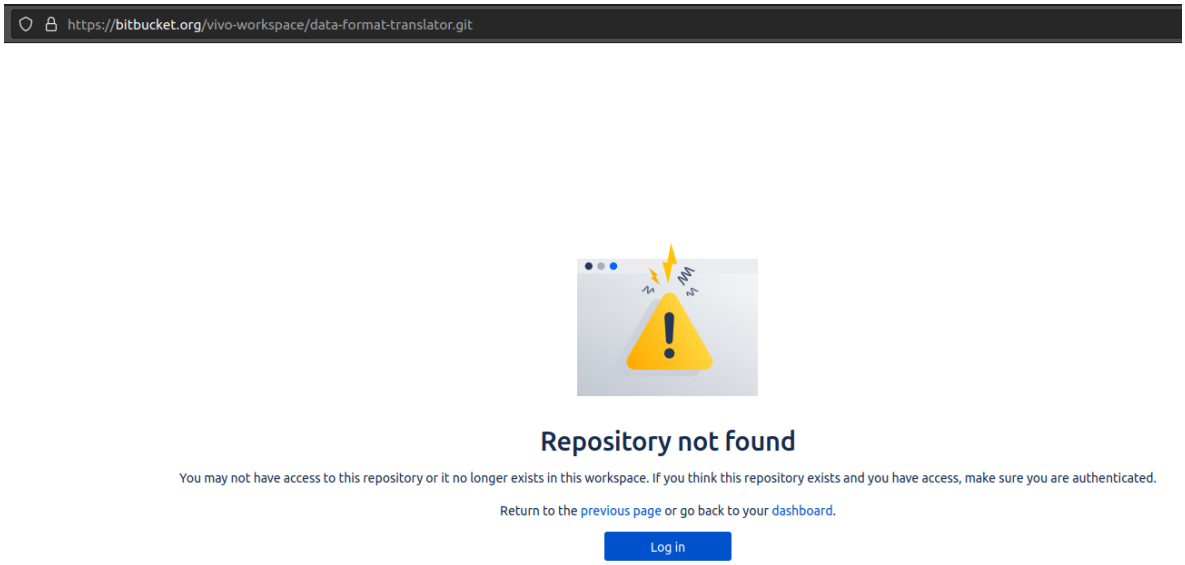
Configuración de los recursos necesarios para ejecutar Dspace

Instalación de Apache Jena y sus otras herramientas asociadas

./DSpace-VIVO/releng/org.vivoweb.dspacevivo.installer/99-OTHER_TOOLS/jena-git-clone-and-deploy.sh

Al ejecutar la instrucción nos muestra un mensaje que el repositorio de donde se quieren descargar las utilidades no existe.

```
Clonando en 'data-format-translator'...
remote: You may not have access to this repository or it no longer exists in this workspace. If you think this repository exists and you have access, make sure you are authenticated.
fatal: repositorio 'https://bitbucket.org/vivo-workspace/data-format-translator.git/' no encontrado
./DSpace-VIVO/releng/org.vivoweb.dspacevivo.installer/99-OTHER_TOOLS/jena-git-clone-and-deploy.sh: linea 16: cd: data-format-translator: No existe el archivo o el directorio
./DSpace-VIVO/releng/org.vivoweb.dspacevivo.installer/99-OTHER_TOOLS/jena-git-clone-and-deploy.sh: linea 17: install-translator.sh: orden no encontrada
./DSpace-VIVO/releng/org.vivoweb.dspacevivo.installer/99-OTHER_TOOLS/jena-git-clone-and-deploy.sh: linea 18: sparql: orden no encontrada
```



Por lo tanto no se puede continuar con la instalación porque al ejecutar los comandos siguientes que nos ofrece la guía, se compilan con errores. La instrucción siguiente de la guía es:

Compilación/Instalación de DSpace-VIVO-EXEMPLE y sus bibliotecas de código

```
./DSpace-VIVO/test/org.vivoweb.dspacevivo.etlexample/script/mvn_install_example.sh
```

Al ejecutar nos aparece el siguiente error, muy probablemente sea porque no se ejecutó la instrucción anterior debido a que no existe el repositorio y no se descargan e instalan los requerimientos necesarios.

```
[INFO] ------------------------------------------------------------------------
[INFO] Reactor Summary for org.vivoweb.dspacevivo.etlexample.installer 0.0.1-SNAPSHOT:
[INFO]
[INFO] org.vivoweb.dspacevivo.model.ontologie ............. SUCCESS [  1.219 s]
[INFO] org.vivoweb.dspacevivo.model.openapi ............... FAILURE [  2.642 s]
[INFO] org.vivoweb.dspacevivo.transformation .............. SKIPPED
[INFO] org.vivoweb.dspacevivo.etlexample.installer ........ SKIPPED
[INFO] ------------------------------------------------------------------------
[INFO] BUILD FAILURE
[INFO] ------------------------------------------------------------------------
[INFO] Total time:  4.029 s
[INFO] Finished at: 2023-08-01T12:18:47-06:00
[INFO] ------------------------------------------------------------------------
[ERROR] Failed to execute goal org.apache.maven.plugins:maven-compiler-plugin:3.3:compile (default-compile) on project org.vivoweb.dspacevivo.model.openapi: Compilation failure -> [Help 1]
[ERROR]
[ERROR] To see the full stack trace of the errors, re-run Maven with the -e switch.
[ERROR] Re-run Maven using the -X switch to enable full debug logging.
[ERROR]
[ERROR] For more information about the errors and possible solutions, please read the following articles:
[ERROR] [Help 1] http://cwiki.apache.org/confluence/display/MAVEN/MojoFailureException
[ERROR]
[ERROR] After correcting the problems, you can resume the build with the command
[ERROR]   mvn <args> -rf :org.vivoweb.dspacevivo.model.openapi
[INFO] Scanning for projects...
```

```
[INFO]
[INFO] ------< org.vivoweb.dspacevivo:org.vivoweb.dspacevivo.etlexample >------
[INFO] Building org.vivoweb.dspacevivo.etlexample 0.0.1-SNAPSHOT
[INFO] --------------------------------[ jar ]---------------------------------
[INFO] ------------------------------------------------------------------------
[INFO] BUILD FAILURE
[INFO] ------------------------------------------------------------------------
[INFO] Total time:  0.824 s
[INFO] Finished at: 2023-08-01T12:18:49-06:00
[INFO] ------------------------------------------------------------------------
[ERROR] Failed to execute goal on project org.vivoweb.dspacevivo.etlexample: Could not resolve dependencies for project org.vivoweb.dspacevivo:org.vivoweb.dspacevivo.etlexample:jar:0.0.1-SNAPSHOT: The fo
llowing artifacts could not be resolved: org.vivoweb.dspacevivo:org.vivoweb.dspacevivo.transformation:jar:0.0.1-SNAPSHOT, org.vivoweb.dspacevivo:org.vivoweb.dspacevivo.model.openapi:jar:0.0.1-SNAPSHOT: C
ould not find artifact org.vivoweb.dspacevivo:org.vivoweb.dspacevivo.transformation:jar:0.0.1-SNAPSHOT at specified path /home/ubuntu22-04/dspace-vivo-prj/00-GIT/DSpace-VIVO/test/org.vivoweb.dspacevivo.e
tlexample/../../bundles/org.vivoweb.dspacevivo.transformation/target/org.vivoweb.dspacevivo.transformation-0.0.1-SNAPSHOT.jar -> [Help 1]
[ERROR]
[ERROR] To see the full stack trace of the errors, re-run Maven with the -e switch.
[ERROR] Re-run Maven using the -X switch to enable full debug logging.
[ERROR]
[ERROR] For more information about the errors and possible solutions, please read the following articles:
[ERROR] [Help 1] http://cwiki.apache.org/confluence/display/MAVEN/DependencyResolutionException
```

Importación de Datos de CSV a RDF

Documentación Oficial:

https://wiki.lyrasis.org/display/VIVODOC111x/Using+the+Convert+CSV+to+RDF+ingest+tool

https://wiki.lyrasis.org/display/VIVO/Using+the+Karma+data+integration+tool

Repositorio de GitHub (Karma):

https://github.com/usc-isi-i2/Web-Karma/wiki/Installation%3A-Source-Code

Convertir CSV a RDF utilizando Ingest Tool de VIVO

1. **Utilizar los archivos de plantilla CSV que proporciona VIVO**

organización.csv

position.csv

people.csv

2. **Crear modelos de espacio de trabajo**

Crearemos dos modelos de datos temporales titulados 'csv-ingest' y 'csv-construct' para mantener nuestro trabajo separado de los modelos principales de VIVO.

1. Seleccione "Ingest Tools" en el menú Advanced Tools
2. Seleccione "Manage Jena Models"
3. Haga clic en el botón "Create Model" y luego escriba un nombre para su modelo, 'csv-ingest'.
4. Repita el paso 3 y nombre el modelo 'csv-construct'

Ingest Menu

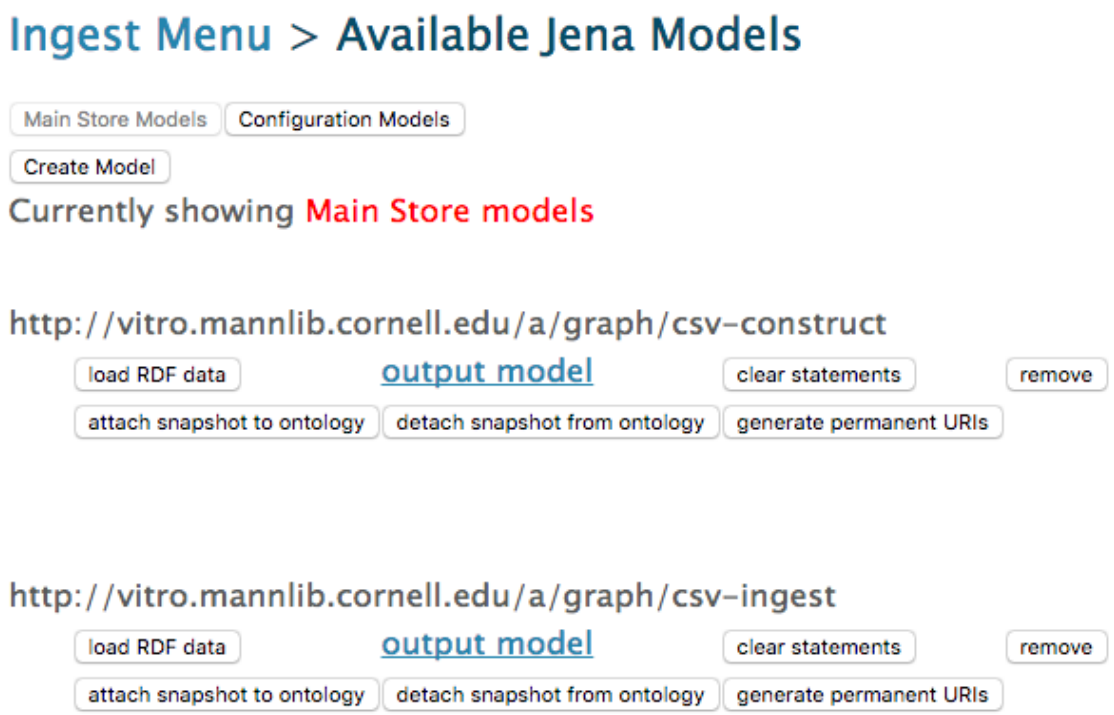
Manage Jena Models
Subtract One Model from Another

Convert CSV to RDF
Convert XML to RDF

Execute SPARQL CONSTRUCT
Generate TBox
Name Blank Nodes
Smush Resources
Merge Resources
Change Namespace of Resources
Process Property Value Strings
Split Property Value Strings into Multiple Property Values

Execute Workflow

Dump or Restore the knowledge base



3. **Importar archivo CSV a RDF**

Ahora haga clic en 'Convert CSV to RDF' en Ingest Tool. Comience proporcionando una URL para su archivo CSV o cargando el archivo directamente desde su computadora. Complete los campos en el formulario.

En el espacio en blanco que se encuentra justo debajo de:

Namespace for Classes and Properties ('in what namespace should these properties be created?'

colocar: http://localhost/vivo/ y en el segundo cuadro en blanco colocar cualquiera de las siguientes opciones: **"ws_ppl"**, "ws_org", y "ws_post"

Y en modelo destino buscar y seleccionar uno de los modelos creados anteriormente, en este caso: "csv-ingest".

Después de hacer clic en convertir, podemos verificar nuestra conversión RDF a través de dos métodos, siendo el primer método el más rápido.

1. Desde el Menú de ingesta, seleccione 'Administrar modelos de Jena' y luego seleccione el 'modelo de salida' del modelo de ingesta. Esto es algo que puede abrir con WordPad y ver los triples creados para su archivo CSV.

2. También desde Manage Jena Model podemos adjuntar el modelo de ingesta a la aplicación web actual. Luego, podemos volver al Administrador del sitio, seleccionar el enlace Jerarquía de clases, seleccionar 'Todas las clases' y navegar hasta el Nombre de la clase que creó (individual, por ejemplo).

Haga clic en 'Siguiente paso'. Aquí creará la URI para sus nuevos individuos.

Seleccione el prefijo URI

Es más conveniente si coincide con la URL de su instancia de VIVO más '/individual/', por ejemplo, http://vivo.university.edu/individual/ .

Ahora, haga clic en 'Convertir CSV'. Los datos CSV ahora deben convertirse en RDF temporales e insertarse en el modelo 'csv-ingest'.

4. **Confirmar los URI de propiedades de datos y la estructura RDF**

La herramienta CSV to RDF convierte el CSV en RDF temporal que podemos consultar mediante SPARQL. Este RDF temporal no se puede mostrar en VIVO. El RDF temporal lo podemos descargar en un archivo volviendo a la página 'Manage Jena Models' y haciendo clic en 'output model' debajo de csv-ingest.

5. **Convertir RDF temporal en ontología VIVO/ISF usando SPARQL CONSTRUCT**

Nota: Este paso no se pudo realizar debido a que los modelos creados no aparecían para ser seleccionados al hacer la consulta SPARQL.

6. **Cargar en la aplicación web**

Ir a la página de administración del sitio

Seleccione Add/Remove RDF

Buscar el RDF temporal creado en el paso 4 y colocarle de extensión a ese archivo .n3 o .rdf

Seleccionar N3 como formato de importación.

La confirmación debe indicar 'RDF agregado del archivo people_rdf'.

Los datos ingeridos ahora deberían mostrarse tanto en el índice como en los resultados de búsqueda.

Uso de la herramienta de integración de datos Karma

Seguir las instrucciones del repositorio de GitHub mencionado al principio de esta página para instalar Karma.

Hay que tener instalado previamente Java 1.7 y Maven 3.0.

Descargar el archivo de instalación desde: https://github.com/InformationIntegrationGroup/Web-Karma/archive/master.zip.

Descomprimir su archivo .zip en el directorio que desee.

Abrir la terminal e ir al directorio donde colocó Karma.

Escribir el comando: mvn clean install

Si se instaló correctamente aparecerá:

```
ubuntu22-04@ubuntu2204-virtual-machine: ~/Descargas/Source.code/Web-Karma
[INFO] ------------------------------------------------------------------------
[INFO] Reactor Summary for ISI Karma Maven Webapp 0.0.1-SNAPSHOT:
[INFO]
[INFO] ISI Karma Maven Webapp ............................ SUCCESS [  4.772 s]
[INFO] karma-util ........................................ SUCCESS [ 27.516 s]
[INFO] karma-semanticlabeling ............................ SUCCESS [05:22 min]
[INFO] karma-typer ....................................... SUCCESS [ 23.252 s]
[INFO] karma-common ...................................... SUCCESS [01:55 min]
[INFO] karma-commands .................................... SUCCESS [  0.011 s]
[INFO] commands-common ................................... SUCCESS [  1.189 s]
[INFO] commands-update ................................... SUCCESS [  1.020 s]
[INFO] commands-worksheet ................................ SUCCESS [  1.695 s]
[INFO] commands-alignment ................................ SUCCESS [  1.498 s]
[INFO] commands-python ................................... SUCCESS [  2.080 s]
[INFO] commands-alignment-openrdf ........................ SUCCESS [ 14.802 s]
[INFO] commands-publish .................................. SUCCESS [  2.542 s]
[INFO] karma-jdbc ........................................ SUCCESS [  7.352 s]
[INFO] commands-publish-jdbc ............................. SUCCESS [  0.171 s]
[INFO] commands-publish-alignment-openrdf ................ SUCCESS [  2.225 s]
[INFO] commands-update-jdbc .............................. SUCCESS [  1.261 s]
[INFO] commands-import ................................... SUCCESS [  0.014 s]
[INFO] import-common ..................................... SUCCESS [  0.780 s]
[INFO] import-csv ........................................ SUCCESS [  0.761 s]
[INFO] import-database ................................... SUCCESS [  0.937 s]
[INFO] import-excel ...................................... SUCCESS [  0.681 s]
[INFO] import-file ....................................... SUCCESS [  0.136 s]
[INFO] import-ontology ................................... SUCCESS [  0.632 s]
[INFO] import-json ....................................... SUCCESS [  0.583 s]
[INFO] import-service .................................... SUCCESS [  0.709 s]
[INFO] import-spatial .................................... SUCCESS [01:04 min]
[INFO] import-xml ........................................ SUCCESS [  0.649 s]
[INFO] import-avro ....................................... SUCCESS [  0.624 s]
[INFO] commands-include .................................. SUCCESS [  0.015 s]
[INFO] karma-web ......................................... SUCCESS [ 36.096 s]
[INFO] karma-offline ..................................... SUCCESS [01:08 min]
[INFO] commands-bloom .................................... SUCCESS [01:33 min]
[INFO] karma-web-services ................................ SUCCESS [  0.014 s]
[INFO] web-services-rdf .................................. SUCCESS [ 38.842 s]
[INFO] web-services-publish-es ........................... SUCCESS [01:08 min]
[INFO] karma-jsonld ...................................... SUCCESS [01:25 min]
[INFO] ------------------------------------------------------------------------
[INFO] BUILD SUCCESS
[INFO] ------------------------------------------------------------------------
[INFO] Total time:  16:32 min
[INFO] Finished at: 2023-08-13T23:34:54-06:00
```

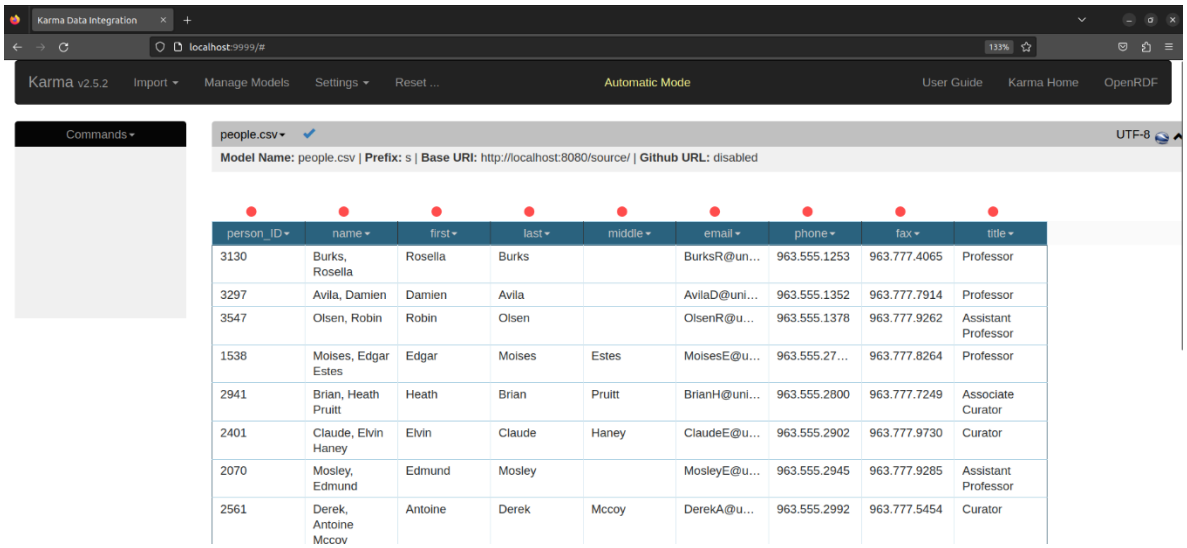
Luego de haberse instalado

Ir a cd karma-web

Inicie la parte del servidor de Karma escribiendo: mvn -Djetty.port=9999 jetty:run Espere hasta que vea [INFO] Servidor Jetty iniciado. Abra un navegador web y vaya a la página http://localhost:9999.

Nota: se puede utilizar cualquier otro puerto, si solo colocamos mvn jetty:run se iniciará en el puerto 8000 y nos dará error, porque VIVO ya está iniciado en ese puerto.

Debería ver una página similar a esta:



person_ID	name	first	last	middle	email	phone	fax	title
3130	Burks, Rosella	Rosella	Burks		BurksR@un...	963.555.1253	963.777.4065	Professor
3297	Avila, Damien	Damien	Avila		AvilaD@uni...	963.555.1352	963.777.7914	Professor
3547	Olsen, Robin	Robin	Olsen		OlsenR@u...	963.555.1378	963.777.9262	Assistant Professor
1538	Moises, Edgar Estes	Edgar	Moises	Estes	MoisesE@u...	963.555.27...	963.777.8264	Professor
2941	Brian, Heath Pruitt	Heath	Brian	Pruitt	BrianH@uni...	963.555.2800	963.777.7249	Associate Curator
2401	Claude, Elvin Haney	Elvin	Claude	Haney	ClaudeE@u...	963.555.2902	963.777.9730	Curator
2070	Mosley, Edmund	Edmund	Mosley		MosleyE@u...	963.555.2945	963.777.9285	Assistant Professor
2561	Derek, Antoine Mccoy	Antoine	Derek	Mccoy	DerekA@u...	963.555.2992	963.777.5454	Curator

15 Vinculación

Se estableció vinculación con la Dirección General de Investigación DIGI-USAC y diferentes instancias de la misma Universidad, por ser los principales interesados y beneficiarios. Asimismo, a través de la mediación de la DIGI, se entablaron vínculos con la Universidad Nacional de Costa Rica UNA para conocer su experiencia en la implementación de su sistema y obtener lineamientos de utilidad en la elaboración de este sistema. También se realizó el vínculo con EUROCRIS, una asociación de expertos en CRIS, dedicados a mejorar la disponibilidad de información de investigación, su representante brindó con su experiencia un valor agregado a este proceso.

Además, se realizaron consultas a expertos internacionales entre ellos:

Guillermo Esteban Murillo Goussen
guillermo.murillogoussen@ucr.ac.cr
Universidad de Costa Rica (UCR)

Mtra. Ana Lorena Jiménez París
ajimene@una.ac.cr
Dirección Ejecutiva
Vicerrectoría de Investigación
Universidad Nacional de Costa Rica

Vivian dos Santos Silva
vivian@ic.ufrj.br
Universidad Federal de Rio de Janeiro
Brasil

Anna Guillaumet
Reserch manager
SIGMA Gestión Universitaria, A.I.E, (M.P.)
https://www.sigmaaie.org/es
España

Participa en una agrupación de universidades públicas españolas, sin ánimo de lucro, medio propio de las universidades socias de la agrupación y les desarrollan y dan mantenimiento a los sistemas de información para el ámbito académico y de la investigación, entre ellos un sistema CRIS.

16 Estrategia de difusión y divulgación

Se hizo una presentación en la etapa intermedia del trabajo realizado junto al resto de proyectos de la convocatoria diferenciada 2023 según el programa asignado.

Se elaboró este informe de investigación con el detalle necesario que describe completamente el diseño realizado.

Se publicó un artículo en una de las revistas de CUNORI-USAC y se tiene en trámite un artículo en la revista de DIGI-USAC.

Los resultados se compartieron a través de una presentación a personal del Instituto de Investigación de CUNORI-USAC.

También se generó una nota de prensa que fue compartida en las redes sociales del Centro Universitario de Oriente y de la Dirección General de Investigación, se hizo uso de videos y material gráfico para este fin.

Se realizó una entrevista en el programa de DIGI que se compartió por varios medios para informar a la población en general.

Se creó un roll up publicitario para usar en cualquier presentación del proyecto.

17 Aporte de la propuesta de investigación a los Prioridades Nacionales de Desarrollo (PND) identificando su meta correspondiente:

Los resultados de la investigación contribuyen a generar el mapa completo de trabajo que aumentará la eficiencia y calidad en la gestión de la investigación. Además fomentará la colaboración entre investigadores e instituciones de investigación reduciendo la brecha en investigación y desarrollo en el país.

El diseño de este sistema es el primer paso para contribuir al monitoreo y la evaluación de las Prioridades Nacionales de Desarrollo (PND) al proporcionar información sobre la investigación y el desarrollo en áreas prioritarias lo cual a su vez puede ayudar al gobierno y otras organizaciones a identificar brechas y oportunidades para el desarrollo sostenible del país.

La Secretaría de Planificación y Programación de la Presidencia (SEGEPLAN), plantea las metas priorizadas por Guatemala. El objetivo número 4, "Educación de calidad" se propone "promover oportunidades de aprendizaje durante toda la vida" (Secretaría de Planificación y Programación de la Presidencia, 2016). En la meta priorizada 4.3, del mismo objetivo, Guatemala se compromete a brindar acceso igualitario a "formación técnica, profesional y superior de calidad" (SEGEPLAN, 2016). Debido a que, un sistema CRIS favorece la visibilidad de los procesos de investigación y sus

resultados, puede contribuir efectivamente en el alcance de la meta 4.3, su repercusión principal es académica al proveer información científica para facilitar el aprendizaje continuo y la formación de calidad de la ciudadanía.

La investigación fue aplicada, el grupo que se beneficiará con estos resultados es el personal de la Dirección General de Investigación DIGI-USAC , a mediano y largo plazo investigadores, funcionarios públicos, distintas organizaciones que proveen financiamiento a la investigación y sociedad en general.

18 Orden de pago final (incluir únicamente al personal con contrato vigente al 31 de diciembre de 2023)

Nombres y apellidos	Categoría (investigador /auxiliar)	Registro de personal	Procede pago de mes (Sí / No)	Firma

19 Declaración del Coordinador(a) del proyecto de investigación

El Coordinador de proyecto de investigación con base en el *Reglamento para el desarrollo de los proyectos de investigación financiados por medio del Fondo de Investigación*, artículos 13 y 20, deja constancia que el personal contratado para el proyecto de investigación que coordina ha cumplido a satisfacción con la entrega de informes individuales por lo que es procedente hacer efectivo el pago correspondiente.

Samy Eunice Pinto Castañeda **Nombre del coordinador del proyecto de investigación**	**Firma**
Fecha: 21/11/2023	

20 Aval del Director(a) del instituto, centro o departamento de investigación o Coordinador de investigación del centro regional universitario

De conformidad con el artículo 13 y 19 del *Reglamento para el desarrollo de los proyectos de investigación financiados por medio del Fondo de Investigación* otorgo el aval al presente informe mensual de las actividades realizadas en el proyecto (escriba el nombre del proyecto de investigación) en mi calidad de (indique: Director del instituto, centro o departamento de

investigación o Coordinador de investigación del centro regional universitario), mismo que ha sido revisado y cumple su ejecución de acuerdo a lo planificado.

M.Sc. Nery Waldemar Galdámez Cabrera **Vo.Bo. Nombre y cargo de quien da el aval al informe**	**Firma**
Fecha: 21/11/2023	

21 Visado de la Dirección General de Investigación

MSc. Inga. Andrea Rodas Morán **Vo.Bo. Nombre Coordinador(a) del Programa Universitario de Investigación**	**Firma**
Fecha: 21/11/2023	

Ing. Agr. MARN Julio Rufino Salazar **Vo.Bo. Nombre Coordinador General de Programas Universitarios de Investigación**	**Firma**
Fecha: 28/11/2023	

/Digi2023