

## Capítulo 2

# RITEC & CRIS: Interoperabilidad para visibilidad y medición del impacto de la producción científica energética

*Laura Icela González Pérez  
María Soledad Ramírez-Montoya  
Francisco José García-Peñalvo  
Héctor Gibrán Ceballos  
Erika A. Juárez Ibarra*

### RESUMEN

El Tecnológico de Monterrey cuenta con dos plataformas para la gestión y la difusión del conocimiento científico, las cuales son el Sistema de Información de Investigación Actual (Current Research Information System – CRIS) para la gestión institucional interna y el Repositorio Institucional (RI) desarrollado como parte del movimiento educativo abierto para dar visibilidad externa en formato abierto a la investigación desarrollada por la institución. Los investigadores capturan su producción científica, como son artículos de revista, libros, capítulos de libro, artículos de conferencias, tesis entre otros, en los dos sistemas y esto hace que la captura de la misma información se duplique. El objetivo de este capítulo es dar a conocer la estrategia que se llevó a cabo, para encontrar puntos en común de ambos sistemas y así permitir la interoperabilidad entre ellos, a través de la iniciativa del proyecto “Laboratorio Binacional para la Gestión Inteligente de la Sustentabilidad Energética y la Formación Tecnológica”. Para abordar el tema se planteó la siguiente pregunta de investigación, ¿cómo se relacionan los CRIS y los del RI para que sean interoperables? Para contestar a ello, se utilizó el método comparativo. Los resultados apuntan a que su relación se entrelaza al homogeneizar los metadatos utilizados en el CRIS y en el RI a través de protocolos de interoperabilidad (e.g. SWORD, OAI) y enviar de forma unidireccional la captura de la producción científica del CRIS hacia el RI, en primera instancia, para la captura de la producción científica del proyecto mencionado, a través de una interfaz única.

“La mayoría de las ideas fundamentales de la ciencia son esencialmente sencillas y, por regla general pueden ser expresadas en un lenguaje comprensible para todos”.  
Albert Einstein

### INTRODUCCIÓN

Los procesos inmersos para publicar un artículo científico, capítulo de libro, un libro, una ponencia, etc., atraviesan por diversas fases que van desde la definición del problema, búsqueda de información, metodología, recopilar, tratar y analizar datos y comunicar la investigación. El conocimiento científico se gestiona a través de sistemas tecnológicos denominados Current Research Information System (CRIS) y la comunicación se hace a través de diversas vías, una de ellas los

Repositorios Institucionales. Los repositorios apoyan la gestión del conocimiento con herramientas tecnológicas de acceso abierto, que ayudan a preservar y difundir la producción científica y académica generada en las instituciones (Ramírez-Montoya y Ceballos, 2017).

El movimiento educativo abierto se fundamenta en prácticas educativas que incentivan y formulan actividades para hacer accesible dicha producción científica generada dentro de las Universidades y Centros de Investigación de manera abierta. Con este estudio se presenta el proceso que se llevó a cabo para establecer la interoperabilidad entre el CRIS-Tec y el Repositorio Institucional (RITEC) del Tecnológico de Monterrey, que involucró una configuración de protocolos, metadatos y plataformas tecnológicas para aumentar la consulta y depósito de recursos de información científica, a partir del proyecto “Laboratorio Binacional para la Gestión Inteligente de la Sustentabilidad Energética y la Formación Tecnológica” en el Repositorio Institucional del Tecnológico de Monterrey. El proyecto está apoyado por la Secretaría de Energía de México y el Consejo Nacional de Ciencia y Tecnología.

El Laboratorio binacional está conformado por el Tecnológico de Monterrey, el Tecnológico Nacional de México, a través del Centro Nacional de Investigación y Desarrollo Tecnológico (CENIDET), el Instituto Investigaciones Eléctricas (IIE) -actualmente Instituto Nacional de Electricidad y Energías Limpias (INEEL)-, la Universidad Estatal de Arizona (ASU), y el Berkeley Energy and Climate Institute (BECI) de la Universidad de California en Berkeley. Ramírez-Montoya y Mendoza—Domínguez (2017) enuncian que estas instituciones participan colaborativamente, en conjunto con la Comisión Federal de Electricidad (CFE), en el desarrollo de líneas de investigación derivadas de los programas nacionales e identificadas como relevantes, particularmente en los siguientes temas:

- a) Modelos de decisión inteligentes para la gestión de la sustentabilidad energética.
- b) Modelos matemáticos y de conocimiento, optimización, simulación y visualización para la sustentabilidad energética.
- c) Tecnología inteligente para la cadena de valor de energía.
- d) Diseño y creación de laboratorios virtuales y físicos para entrenamiento técnico y tecnológico en temas que impacta la sustentabilidad energética.
- e) Tecnología educativa para cursos masivos abiertos en línea (MOOC) en temas de sustentabilidad energética, y conversión de ingenieros para el sector.

En este capítulo se dan a conocer las definiciones conceptuales de un RI y un CRIS y su relación, los esquemas actuales de estándares de metadatos y recomendaciones a nivel mundial. Después se presenta el caso particular del Tecnológico de Monterrey y la solución propuesta para hacer posible la interoperabilidad entre el CRIS-Tec y el RITEC, con las ventajas que han dado para maximizar la

gestión y visibilidad de la producción científica. En especial, se presenta el resultado de la interoperabilidad de los sistemas, para dar visibilidad abierta a la producción científica energética, generada a través del proyecto del Laboratorio Binacional.

## **1. MARCO TEÓRICO**

### **1.1 Repositorios Institucionales**

Los Repositorios Institucionales albergan los recursos educativos abiertos que se generan dentro de una Universidad o Centro de Investigación. González-Pérez, Glasserman, Ramírez-Montoya, y García-Peñalvo, (2017) mencionan que el término Repositorio Institucional (RI) es conocido como un conjunto de servicios web centralizados, cuyo objetivo es organizar, gestionar, preservar y divulgar principalmente producción científica y académica que pertenece a una comunidad académica en acceso abierto. La producción científica que se preserva en los RI está conformada desde artículos, conferencias, capítulos de libros, libros, tesis y producción académica como por ejemplo videos, monografías, presentaciones, infografías, presentaciones, entre otros, y cuya característica en común es que cuentan con un licenciamiento de acceso abierto que permite que estos materiales sean reutilizados, citados y reconocidos en diversos contextos.

Los RI son las plataformas que dan acceso libre a los recursos que son depositados en ellos, este depósito se lleva a cabo por sus autores a través de un flujo de trabajo que se relaciona con la estructura de un conjunto de metadatos detallados que facilitan y permiten que el descubrimiento de estos recursos. Yañez-Figueroa, Ramírez-Montoya, y García-Peñalvo, (2017) consideran que la finalidad de la diseminación del conocimiento científico es dar espacios digitales o físicos a los ciudadanos, para que ahí confluyan las ideas para fundamentar soluciones ante las demandas de la sociedad. Uno de estos espacios digitales es el Repositorio Institucional donde se accede a información científica que sea la base teórica y empírica para dichas propuestas.

### **1.2 Current Research Information System (CRIS)**

Los Sistemas de información e Investigación Actual (CRIS por sus siglas en inglés) son plataformas al servicio de los investigadores de una institución para gestionar sus proyectos

financiados y la producción científica que generan. La necesidad de una Institución por contar con un CRIS que realice una gestión sistematizada de los proyectos financiados para investigación y la generación de conocimiento aumenta porque se cuenta con los indicadores de la investigación, la innovación y la transferencia de conocimiento y tecnología que produce la Institución hacia otros sectores. Velásquez-Duran y Ramírez-Montoya (2018) señalan que los sistemas CRIS permiten la gestión global eficiente de todos los procesos relacionado con la investigación y su información derivada de una Institución y que permiten la interoperabilidad entre la información científica. Además, con las actuales reformas a las políticas de acceso abierto en 2014 en México para las instituciones educativas se vuelve estratégico contar con un sistema de gestión y visibilidad vinculado para dar seguimiento a la producción científica generada.

Para gestionar los sistemas CRIS se requiere que se involucren varios departamentos o servicios, y adecuarse a las necesidades de la alta dirección de la institución para tomar en cuenta aspectos relacionados con decisiones estratégicas a través de los resultados de la calidad de datos (Ribeiro, de Castro y Mennielli, 2016). Los CRIS adquieren un rol dentro del ecosistema institucional educativo debido a que gestionan la información científica y al hacerlos interoperables con sistemas abiertos como el RI, pueden brindar accesibilidad, captura y visualización centralizada de la información científica de manera inteligente. Gómez-Aguilar, García-Peñalvo, y Therón (2014) mencionan que ya existe una analítica visual disponible para obtener representaciones de información a modo de un modelo mental de datos y de nuevo conocimiento, por ello el integrar la este tipo de visualización a la información científica, permitirá identificar las relaciones que hay entre la información de diversos sistemas y así identificar puntos de encuentro de los datos que se capturan por parte de los investigadores y de instituciones en los sistemas de gestión.

### **1.3 Interoperabilidad entre sistemas de investigación**

Según la Asociación Española de Documentación e Información (SEDIC, 2004), el término “interoperabilidad” se describe de la siguiente manera:

*“[...] la habilidad que tiene un sistema o producto para trabajar con otros sistemas o productos sin un esfuerzo especial por parte del cliente. [...] es un concepto clave al hablar*

*de esquemas de metadatos y de la necesidad de compatibilizar todos ellos, para una recuperación de información integral en distintas colecciones de datos y metadatos distribuidos. La interoperabilidad entre distintos esquemas de metadatos puede realizarse de diversas formas, por ejemplo, a través del funcionamiento de un protocolo (tipo OAI – Open Archives Initiative) o bien a través del mapeo o establecimiento de correspondencias entre informaciones en diferentes formatos (por ej. MARC-DC, FGDC-DC, etc.), para la conversión de elementos de metainformación que permita hacerlos compatibles”. SEDIC, 2004*

Con la definición anterior, se puede argumentar que los metadatos son pieza clave para que funcione la interoperabilidad y los que tienen la capacidad de lograr que la recuperación de información sea integral entre las distintas colecciones de datos y metadatos distribuidos. Por otro lado, para autores como Santos. Hermosa, Ferran y Abadal, (2012), el uso de metadatos y estándares internacionales tiene la finalidad de dar la característica de interoperable a los Repositorios Institucionales, aunque los flujos de trabajo para su almacenamiento deben vincularse con los metadatos para asegurar su descubrimiento en las plataformas. Los autores Qin, Ball, y Greenberg (2012) señalan que los metadatos tienen una misión crítica para descubrir la información científica y llevar un control de la citación de los autores y señalan que deben contar con las siguientes *e-science properties*:

- a) Verificable: Sirve para asegurar la validez de la investigación y permite a los investigadores que no sean el propietario de los datos repita el estudio usando los mismos datos.
- b) Interfuncional: Los conjuntos de datos deben contener suficientes metadatos para facilitar su descubrimiento, selección, agregación o filtrado y reutilización.
- c) Analizable: Los conjuntos de datos deben encontrarse en un estado que permita una manipulación de datos mínima para identificar el estado de la investigación científica.
- d) Interoperable: Los conjuntos de datos deben cumplir con los estándares para que puedan ser comunicados y procesados por diferentes sistemas y herramientas de software.

En la actualidad, los metadatos utilizados para la representación y descripción de los datos científicos son el componente de la comunicación científica contemporánea. En las últimas décadas, las comunidades de las ciencias físicas, sociales y de la vida han desarrollado esquemas de metadatos

para facilitar la documentación, el intercambio, el archivo y la reutilización de los datos de investigación. A continuación, se describen (a) los metadatos de los CRIS, (b) los metadatos de los RI y (c) los esquemas para soportar la interoperabilidad entre las dos plataformas.

**a) Los metadatos de los CRIS (*CERIF*)**

Los conjuntos de datos y metadatos científicos utilizados por las plataformas de gestión de contenido, como los CRIS y los RI permiten el descubrimiento de los recursos a través de las búsquedas que identifican de forma más sencilla los conjuntos de colecciones, autores y proyectos de Investigación (Jeffery y Asserson, 2009). Los CRIS se basan en el estándar CERIF, el cual es un estándar para la gestión y el intercambio de datos de investigación, es decir, la información y datos sobre investigadores, organizaciones, proyectos, productos y financiación, derivados del proceso de investigación para proporcionar un modelo de datos que se puede utilizar para describir el dominio de la investigación, incluyendo las relaciones entre las partes constituyentes, y cómo estos cambian con el tiempo. La semántica formal de CERIF proporciona un núcleo definido y actual de entidades y da respuesta a las solicitudes de interoperabilidad con repositorios institucionales.

**b) Los metadatos de los RI (*Dublin Core*)**

Los Repositorios típicamente usan metadatos basados en el estándar de Dublin Core (DC) y con el uso del protocolo Open Archive Initiative – Protocol for Metadata Harvesting (OAI-PMH) se hace posible la interoperabilidad e indexación con Google Scholar. Los desarrolladores de repositorios que utilizan *Dublin Core*, *MARC*, etc. comienzan a darse cuenta de las limitaciones de estos formatos y cada vez más están adoptando CERIF como modelo.

**c) Esquemas de metadatos para soportar la interoperabilidad entre los CRIS y los RI**

Los estándares que se están utilizando para homogenizar los Repositorios Institucionales y CRIS se basan en una guía de OpenAIRE que brinda directrices para soportar la interoperabilidad. Esta guía fue creada por la Fundación Española para la Ciencia y la Tecnología (FECYT) y la Red de Bibliotecas Universitarias (REBIUN) de la CRUE, las cuales crearon la plataforma RECOLECTA (<https://recolecta.fecyt.es/coar>) que tiene el objetivo de crear una infraestructura nacional de repositorios científicos de acceso abierto y garantizar que sean interoperables según los estándares de la comunidad mundial (Recolector de ciencia abierta [RECOLECTA], 2018). De esta plataforma en

2012 se creó el proyecto OpenAIRE (Open Access Infrastructure for Research in Europe) que apoya a la política de acceso abierto en Europa, a ese proyecto le siguió el proyecto OpenairePlus (2nd Generation of Open Access Infrastructure for Research in Europe) el cual es un proyecto continuación de OpenAIRE cuyo objetivo es incorporar los sets de datos de investigación junto a los artículos científicos en países europeos.

De esa plataforma nacen la Confederación de Repositorios de Acceso Abierto (COAR) que une y representa a 90 instituciones de todo el mundo (Europa, América Latina, Asia, y América del Norte), cuya misión es mejorar la visibilidad y lograr una mayor aplicación de los resultados de investigación a través de redes globales de repositorios de acceso digital abierto. Príncipe, et al. (2014) analizaron la guía de OpenAIRE que soporta la interoperabilidad entre los repositorios institucionales, los repositorios de datos y los CRIS (Tablas 1 y 2) y sugieren el siguiente uso para cada uno de los campos:

Mandatory (M) = el campo siempre debe estar presente en el registro de metadatos. Un elemento vacío no está permitido.

Mandatory when applicable (MA) = cuando se puede obtener el valor del campo, debe estar presente en el registro de metadatos.

Recommended (R) = el uso del campo es recomendado.

Optional (O) = la propiedad puede ser utilizada para proporcionar información complementaria sobre el recurso

Tabla 1

*Metadatos para Repositorios Institucionales y Repositorios de datos científicos*

<b>Repositorios Institucionales</b>	<b>Repositorios de datos científicos</b>
1. Title (M)	1. Identifier (M)
2. Creator (M)	2. Creator (M)
3. Project Identifier (MA)	3. Title (M)
4. Access Level (M)	4. Publisher (M)
5. License Condition (R)	5. PublicationYear (M)
6. Embargo End Date (MA)	6. Subject (R)

7. Alternative Identifier (R)	7. Contributor (MA/O)
8. Publication Reference (R)	8. Date (M)
9. Dataset Reference (R)	9. Language (R)
10. Subject (MA)	10. ResourceType (R)
11. Description (MA)	11. AlternateIdentifier (O)
12. Publisher (MA)	12. RelatedIdentifier (MA)
13. Contributor (R)	13. Size (O)
14. Publication Date (M)	14. Format (O)
15. Publication Type (M)	15. Version (O)
16. Publication Version (R)	16. Rights (MA)
17. Format (R)	17. Description (MA)
18. Resource Identifier (M)	18. GeoLocation (O)
19. Source (R)	
20. Language (R)	
21. Relation (O)	
22. Coverage (R)	
23. Audience (R)	

Tabla 2  
*Metadatos de sistemas CRIS*

Publication	Product/Dataset	Person	Organization	Project	Funding	Equipment	Service
Internal Identifier	Internal Identifier	Internal Identifier	Internal Identifier	Internal Identifier	Internal Identifier	Internal Identifier	Internal Identifier
Product / Dataset	Name	Gender	Short name	Start Date	Start Date	Acronym	Name
Person	Description	First name(s), Family name(s) and Infix name(s)	Legal name	End Date	End Date	Name	Description

Organisation	Subject	Nationality	Organisation type	Acronym	Acronym	Description	Service Type
Project	Language	Federated Identifiers	Country	Title	Amount	Federated Identifiers	Website Address
	Product Type	Contact	Organisation Seat Location	Abstract	Name	<b>Relations hip(s) with</b>	OAI-PMH Base URL
	Product Type (CERIF)	<b>Relations hip(s) with</b>	Federated Identifiers	Subject	Description	Dataset	Subject Classification Resource Address
	Open Access Type	Publication	Organisation Website Address	Open Access Requirements	Subject		<b>Relations hip(s) with</b>
	License Type	Product / Dataset	Publication	Federated Identifiers	Funding Type		Organisation
	Federated Identifiers	Organisation	Product / Dataset	<b>Relations hip(s) with</b>	Open Access Requirements		
	Publication	Project	Person	Publication	Federated Identifiers		
	Product / Dataset	Funding	Project	Product / Dataset	<b>Relations hip(s) with</b>		
	Person		Funding	Person	Person		
	Organisation		Service	Organisation	Organisation		
	Project			Funding	Project		
	Equipment				Funding		

La unificación de los sistemas CRIS con los Repositorios está en proceso todavía, por ejemplo, en Noruega (FRIDA a NORA), Flandes (FRIS) y Dinamarca (PURE - CRIS). Esto se debe a que el CRIS, en una institución de investigación, está íntimamente vinculado al perfil y el trabajo del investigador y al flujo de trabajo organizacional, y gran parte de los metadatos requeridos para una publicación (autor, institución, derechos) ya están almacenados en el CRIS y no necesita el autor volverlos a ingresar. Además, los metadatos de publicación están rodeados de información contextual relevante de uso para el usuario final.

Los metadatos son la base de las tecnologías para la gestión científica y por ello se requiere de una visión a largo plazo que sustente la infraestructura de información científica en las instituciones. En

este sentido, los autores Wilkinson et al. (2016) proponen una guía para la gestión y proponen que los metadatos de todo recurso de información deben contar con cuatro características: encontrables, accesibles, interoperables y reutilizables.

Para lograr que los recursos sean encontrables se requiere que tengan un identificador global único y persistente, que sus metadatos estén enriquecidos y describan claramente al recurso, y que estén indexados de la mayor cantidad de formas posibles. Para que sean recuperables deben ser accesibles por su identificador usando un protocolo de comunicación estándar, abierto, gratuito y universalmente implementable; en los casos que sea necesario, el protocolo debe implementar un procedimiento de autenticación y autorización; y en caso de que el recurso no esté accesible, sus metadatos sí lo deben estar.

Sobre la base de un ciclo de vida de investigación típico, Wilkinson et al. (2016) definieron 10 tareas de usuario de datos (Tabla 3). Las primeras cuatro tareas de usuario son las mismas que las definidas por los estándares de catalogación de bibliotecas: descubrimiento, identificación, selección y obtención. Las seis tareas restantes son exclusivas de la investigación científica: verificar, analizar, gestionar, archivar, publicar y citar.

Tabla 3  
*Las 10 tareas de usuario de datos*

Tareas genéricas (catalogación de bibliotecas)	Tareas científicas	Tareas de datos	Tareas de difusión
descubrimiento	verificar	gestionar	publicar
identificación	analizar	archivar	Citar
selección			
obtención			

## 2. METODOLOGÍA

En este estudio surge la pregunta de investigación ¿cómo se relacionan los CRIS y los del RI para que sean interoperables? para ello se utilizó el método comparativo de metadatos, debido a que

se requiere encontrar los metadatos con los que se realiza la interconexión de los dos sistemas y así permitir la interoperabilidad entre ellos, a través de la iniciativa del proyecto “Laboratorio Binacional para la Gestión Inteligente de la Sustentabilidad Energética y la Formación Tecnológica”. El método comparativo es lógico y sistemático que para Liphajrt (1971) podría ser considerado como una estrategia de investigación, pero no como una técnica de investigación. Para Sartori (1984) el método comparativo tiene como objetivo la búsqueda de similitudes y disimilitudes, dado que la comparación se basa en el criterio de homogeneidad; siendo la identidad de clase el elemento que legitima la comparación, se compara entonces lo que pertenece al mismo género o especie. Las disimilaridades se presentan como lo que diferencia a la especie de su género, y esto no es lo mismo que señalar las variaciones internas de una misma clase; por lo cual se requiere de un trabajo sistemático y riguroso que implique la definición previa de las propiedades y los atributos posibles de ser comparados. Según Fideli (1998) el método comparativo es un método para confrontar dos o varias propiedades enunciadas en dos o más objetos, en un momento preciso o en un arco de tiempo más o menos amplio. De esta manera se comparan los metadatos que se utilizan en el RITEC y el Cris – Tec con el objetivo de diseñar los flujos de información de ambos sistemas y así automatizar la transferencia y almacenamiento de los datos funcionales que son capturados en el CRIS-TEC para que se almacenen de manera automática en el repositorio RITEC. Como se mencionó anteriormente, los metadatos son pieza clave para que funcione la interoperabilidad ya que tienen la capacidad de lograr que la recuperación de información sea integral entre las distintas colecciones de datos y metadatos distribuidos. Senso y De la Rosa Piñero (2003) mencionan que la definición para metadato dada por Tim Berners Lee va más allá de descripción de recursos y engloba las particularidades de cada recurso por lo que plantea tres tipos de metadatos: (1) el metadato que se encuentra dentro del documento mismo (2) El metadato que se produce durante una transferencia HTTP (HyperText Transfer Protocol) –cliente y servidor se envían información sobre el objeto que están transmitiendo por medio de metadatos – y (3) el metadato que se utiliza cuando se consulta en otro documento para comprobar si se puede acceder a él. Es por eso que la comparación se realiza teniendo en cuenta las características de origen y la descripción, normalización internacional e institucional de los datos que se emplean para cada uno.

### 3. Desarrollo de la implementación

#### 3.1 Interoperabilidad del RITEC y CRIS-TEC

El CRIS-Tec es un sistema que se ha desarrollado por más de 10 años en el Tecnológico de Monterrey (Cantú, Ceballos, Mora y Escoffí, 2005). Durante ese tiempo, se han incorporado múltiples fuentes de datos institucionales, alrededor del cual se ha construido un modelo de datos que soporta los elementos correspondientes a los sistemas CRIS descritos por Príncipe, et al. (2014). En sentido inverso, el CRIS-Tec se convirtió en proveedor de información institucional respecto a la investigación del Tec, tanto para departamentos internos, como para entidades externas. La Figura 1 muestra los sistemas a los cuales se proveía ya de información a partir del CRIS-Tec.



Figura 1. Clientes de datos del CRIS-Tec.

Al hacer interoperable el CRIS-Tec con el RITEC se analizaron dos opciones: 1) Utilizar el RITEC como interfaz única de captura de producción científica, y 2) utilizar el CRIS-Tec como interfaz única de captura de producción científica. En ambos casos, se tendrían que implementar un proceso de exportación-importación de datos de un sistema al otro.

Se inició por analizar el proceso de captura/dépósito en ambos sistemas tal como estaban antes de la integración. Las Figuras 2 y 3 ilustran dichos procesos. Como puede verse, en CRIS-Tec se importa información de bases de datos referenciales tales como Scopus, la cual es actualizada por cualquiera de los autores de la publicación. El autor también tiene la opción de registrar una publicación que esté indizada en Scopus. En ambos casos, CRIS-Tec sólo registraba metadatos de la publicación.

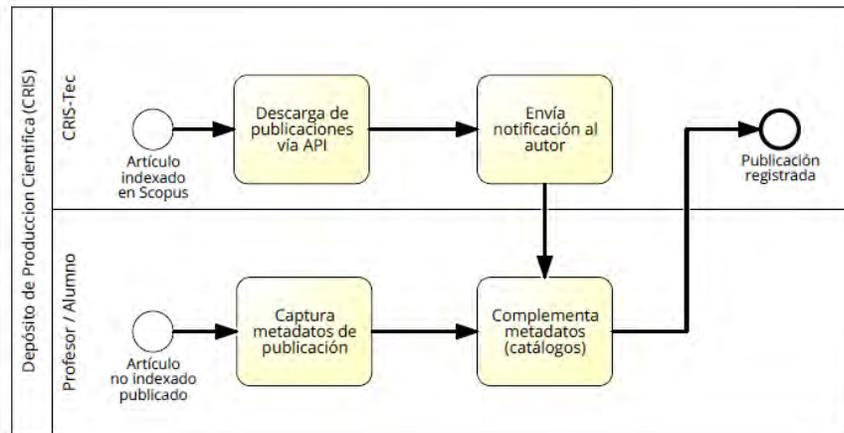


Figura 2. Flujo de registro de producción científica en CRIS-Tec.

El depósito en el Repositorio del Tecnológico de Monterrey (RITEC) parte del registro de los metadatos por parte del autor, y cuenta con un flujo de revisión y autorización por parte de un catalogador especializado en la vía verde de Acceso Abierto (Figura 3).

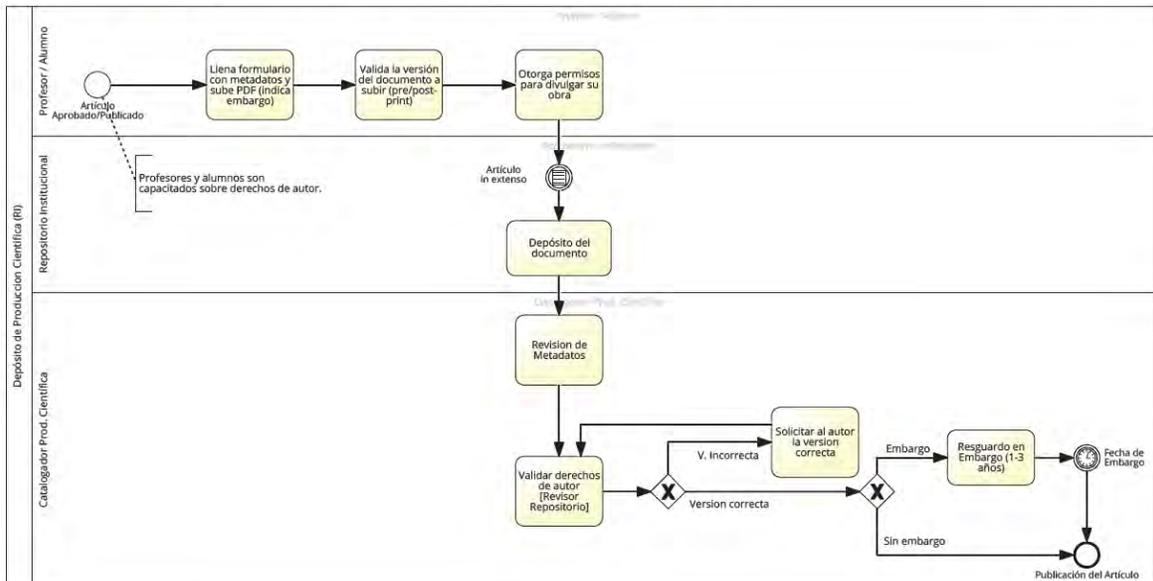


Figura 3. Flujo de registro de producción científica en el RI.

El autor o el catalogador pueden solicitar un periodo de embargo para respetar tanto procesos de protección de propiedad intelectual (patentes), como permisos de la editorial sobre el autoarchivo de producción científica de acceso restringido. La interfaz del RITEC se presenta en acceso abierto (Figura 4), donde una colección especial es la del Proyecto SENER que preservar y difunde en abierto la producción científica y académica generada en el proyecto “Laboratorio Binacional para la Gestión Inteligente de la Sustentabilidad Energética y la Formación Tecnológica”.

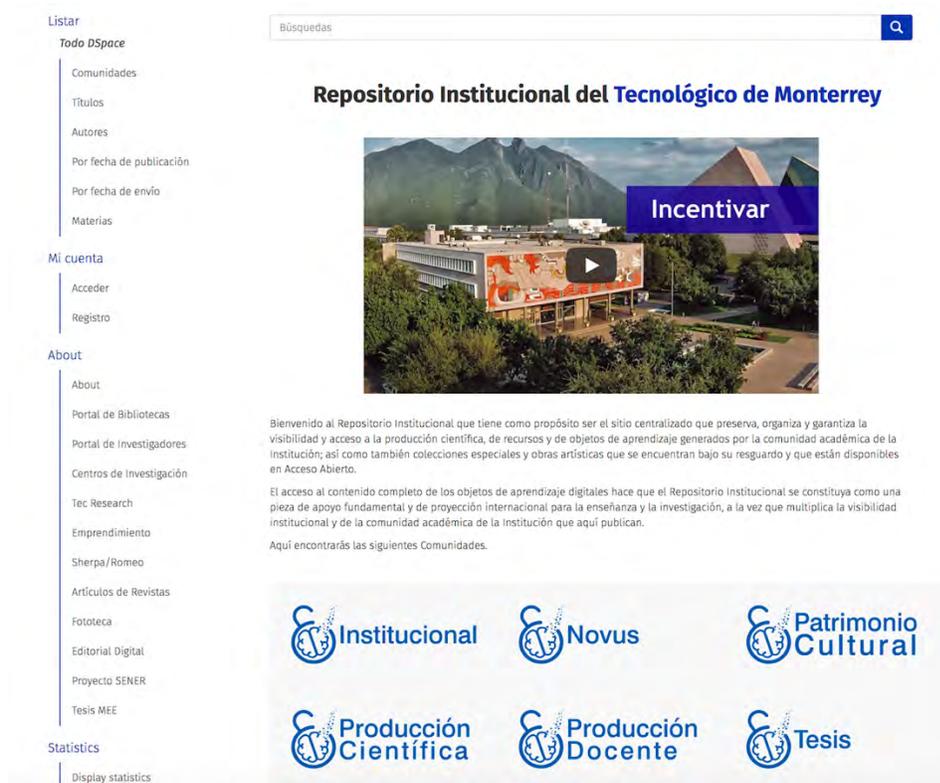


Figura 4. Portal del RITEC (Web: <https://repositorio.itesm.mx>)

Posteriormente se identificaron las ventajas y desventajas de ambos enfoques de integración. El resultado de este análisis se muestra en la Tabla 4.

Tabla 4  
Ventajas y desventajas de los enfoques de integración CRIS-Tec y RITEC

RITEC como interfaz de captura	CRIS-Tec como interfaz de captura
<b>Ventajas (Pros)</b> <ul style="list-style-type: none"> <li>● Se pueden configurar interfaces de captura por colección (tipo de publicación)</li> <li>● Se importan metadatos a partir del DOI.</li> <li>● Se puede vincular al autor con su ORCID.</li> <li>● Cuenta con flujos de validación y autorización configurables.</li> <li>● Se puede utilizar OAI-PMH para exportar metadatos</li> </ul>	<b>Ventajas (Pros)</b> <ul style="list-style-type: none"> <li>● Se tienen catálogos especializados (personas, afiliaciones, revistas, etc.).</li> <li>● Se tienen entidades con descripciones y relaciones a otras entidades.</li> <li>● Existe una relación entre la obra y cada uno de los autores.</li> <li>● Se puede exportar la información en el formato que se requiera.</li> <li>● Se puede utilizar el protocolo SWORD para depositar documentos en el RI.</li> </ul>
<b>Desventajas (Cons)</b> <ul style="list-style-type: none"> <li>● No se tienen catálogos especializados (personas, afiliaciones, revistas, etc.)</li> <li>● No se tiene una relación entre la obra y los autores, a excepción del depositante.</li> <li>● La información proveniente del RI tendría que homologarse con los catálogos del CRIS-Tec.</li> </ul>	<b>Desventajas (Cons)</b> <ul style="list-style-type: none"> <li>● Se requiere desarrollar la funcionalidad de depósito en el RI.</li> <li>● Sólo empleados y alumnos activos pueden registrar sus obras.</li> <li>● El RI pierde visibilidad al momento de depositar.</li> </ul>

A partir de este análisis se concluyó que la principal limitante para la captura única desde el RI es la falta de catálogos especializados y la necesidad de homologar datos con dichos catálogos, lo cual es una tarea que no se puede automatizar. Por ello se decidió extender las capacidades del CRIS-Tec para utilizarlo como interfaz única de captura para las colecciones de Producción Científica.

Para automatizar el depósito de documentos en el RI se optó por desarrollar en el CRIS-Tec una implementación del protocolo SWORD (Simple Web-service Offering Repository Deposit), que es un protocolo usado en repositorios para poder realizar envíos de contenidos desde otras aplicaciones o servicios.

### **3.2. Implementación de SWORD en CRIS-Tec**

La importancia que tiene la conexión entre dos sistemas de esta naturaleza tiene que ver por la magnitud del valor de la información científica que ambos gestionan hacia dentro y fuera del la Institución. de acuerdo con González et. Al. (2018) los repositorios institucionales deben cumplir con ciertos objetivos definidos en el contexto del Movimiento Educativo Abierto, uno de ellos tiene que ver con su evaluación de entidades internacionales, de gobierno locales y de lineamientos institucionales y una de las más importantes que es la de los usuarios quienes la usan.

Para lograr la activación de la conexión entre los dos sistemas se debe implementar el protocolo SWORD en el repositorio que debe tener instalada la herramienta DSpace la cual permite acceder, mediante un servicio o aplicación web, a recibir y enviar datos al repositorio. La importancia de hacerlo es que al configurar el servicio de envío se simplifica el proceso de envío de información y cualquier usuario registrado o no registrado en el repositorio pueda insertar sus contenidos de forma simple. Para Allinson, Francois, y Lewis (2004) el perfil de Atom Publishing Protocol se utiliza para implementar SWORD en interfaces de repositorios gestionados con herramientas como DSPACE, ePrints, Fedora, Intralibrary, que ofrece funcionalidades de realizar el depósito de recursos desde múltiples sistemas con la mediación del servicio por parte de bibliotecarios y catalogadores. Por ello existe la necesidad de implementar la interoperabilidad en el Repositorio y que sólo ciertos perfiles de usuarios capturen la información de cierto contenido.

El desarrollo de la interoperabilidad CRIS-Tec y la implementación de SWORD, permite la carga de información pueda subir información directamente al repositorio de manera automática. De esta forma el número de usuarios o los permisos no son relevantes, ya que ese proceso lo gestiona CRIS-Tec. Posterior al depósito, un usuario cualificado del repositorio, p.ej. bibliotecario, se encargará de validar y complementar los datos recibidos.

Por una parte, el depósito a través del protocolo SWORD requiere un conjunto de metadatos describiendo el recurso, el archivo a depositar (el texto completo de la publicación) y la colección en la que se depositará en el Repositorio. Por otra parte, la interfaz de captura de publicaciones en CRIS-Tec se reconfigura dependiendo del tipo de publicación a capturar, cambiando la lista de datos que se pueden capturar para cada uno de ellos. Por eso fue necesario tener dos tipos de mapeos:

1. Mapeo de tipo de publicación: Un catálogo de tipos de publicación se utiliza como referencia para mapear tanto los artículos importados de Scopus y de Web of Science como las publicaciones no indexadas capturadas en el CRIS-Tec , con las colecciones en el RITEC (identificadas por claves).
2. Mapeo de metadatos: Un catálogo de campos de captura en CRIS-Tec se utiliza para mapear cada dato con el metadato correspondiente en el RITEC, dependiendo del tipo de publicación

a registrar. Dicho catálogo tiene indicado un tipo de dato que es utilizado para formatear los metadatos en formato XML.

La Tabla 5 muestra el mapeo de metadatos para artículos de revista. Puede verse por ejemplo que la fecha de publicación se construye a partir de dos campos en el CRIS-Tec, mes y año. De forma similar, el metadato dc.format.extent se compone de una concatenación de la página inicial y final del artículo.

Tabla 5  
*Mapeo de metadatos para artículos de revista en CRIS-Tec*

<b>Campo CRIS-Tec</b>	<b>Tipo</b>	<b>Metadato RITEC</b>
Nombre del artículo	Cadena	dc.title
Autor(es)	Cadena	dc.contributor.author
Creador del registro	Cadena	dc.creator
Nombre de la revista	Cadena	dc.publisher
ISSN de la revista	Cadena	dc.identifier.issn
Volumen	Entero	bibo.volume
Número	Entero	bibo.issue
Página inicial	Entero	bibo.pageStart
Página final	Entero	bibo.pageEnd
Página inicial - Página final	Entero	dc.format.extent
01 - Mes - Año	Fecha	dc.date.issued
Categoría conspectus	Cadena	dc.subject
Abstract	Texto	dc.description.abstract

#### **4. Resultados**

A continuación se enuncian tres aspectos fundamentales para la relación entre CRIS-Tec y RITEC:

1) El depósito de los recursos educativos abiertos (REA) a través de la interfaz de captura del CRIS -Tec se realiza con los mismos metadatos de la colección Producción Científica del RITEC, por lo que la captura de estos metadatos en el RITEC se vuelve innecesaria. La redefinición del proceso de captura de la producción científica en el CRIS-Tec y en RITEC incluye la implementación del protocolo SWORD. Si bien Allinson, Francois, y Lewis (2004) señalan que con SWORD como parte de las interfaces de repositorios el depósito de recursos se puede hacer desde múltiples sistemas con la mediación del servicio por parte de bibliotecarios y catalogadores, ver en la Figura 5 el flujo de trabajo para el depósito de producción científica en el CRIS-Tec que involucra captura de información adicional a la que pide en el RITEC, la cual requiere una compatibilidad para que los nombres de autores registrados en el Consejo Nacional de Ciencia y Tecnología (CONACYT) en México, sean los mismos que se capturan en el CRIS-Tec y en el RITEC

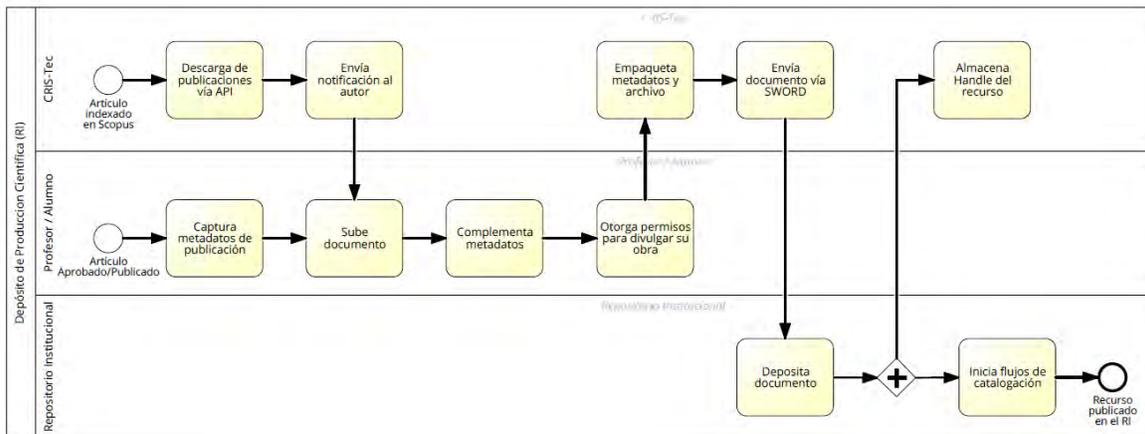


Figura 5. Flujo de registro de producción científica en CRIS-Tec y RITEC.

2) Para que los metadatos sean interoperables deben utilizar un lenguaje formal, accesible, compartido y ampliamente aplicable; en la medida de lo posible, los metadatos deben poder referenciar otros metadatos. Los metadatos son reutilizables en la medida que describen ricamente al recurso, con una pluralidad de atributos precisos y relevantes, se publican con una licencia de uso clara y accesible, están asociados con su origen o procedencia, y cumplen

con las normas relevantes para el dominio. Si bien para Príncipe, et al. (2014) se debe de utilizar una guía, como la de OpenAIRE, que soporta la interoperabilidad entre los repositorios institucionales, los repositorios de datos y los CRIS. Los metadatos de la investigación científica giran en torno a la toda la tarea del usuario y los convierte en un requisito para otras tareas para poder realizarlas. Por ello, la infraestructura de metadatos para datos y recursos científicos puede beneficiar enormemente las tareas de los usuarios. La Figura 6 muestra el proceso de actualización de un recurso desde el CRIS-Tec. Dado que un recurso publicado en RITEC no puede modificarse es necesario borrarlo y volverlo a crear. Para el autor este procedimiento es transparente, mientras que el catalogador debe volver a checar los metadatos y el documento.

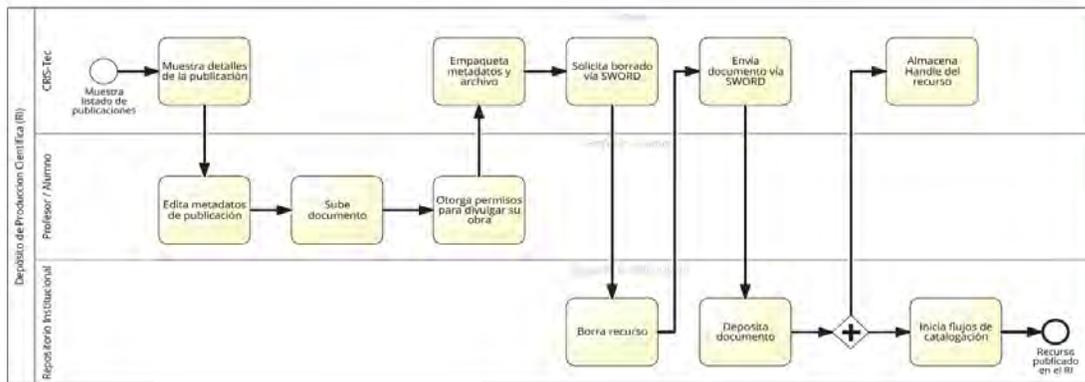


Figura 6. Flujo de actualización de producción científica en CRIS-Tec y RI.

- 3) El CRIS-Tec mantiene una relación (mapeo) entre publicaciones almacenadas en su base de datos y recursos publicados en RITEC lo que adicionalmente permitirá medir el efecto del autoarchivo en la citación de las publicaciones compartidas en acceso abierto y hoy se refleja en la visibilidad y medición del impacto de la producción científica generada en el proyecto “Laboratorio Binacional para la Gestión Inteligente de la Sustentabilidad Energética y la Formación Tecnológica”, a través de su colección ver Figura 7. Con ello se cumple el objetivo principal de un Repositorio Institucional que señalan los autores González-Pérez, Glasserman, Ramírez-Montoya, y García-Peñalvo, (2017) el cual consiste en organizar, gestionar, preservar y divulgar la producción científica y académica en formato abierto. La producción científica que se presenta en acceso abierto incluye artículos en revistas indizadas, artículos

de conferencias, capítulos de libros, reseñas de libros, informes técnicos y la producción académica incluye presentaciones del equipo de investigación en las reuniones de trabajo del proyecto, infográficos, diagramas, dispositivas, videos, ejercicios y experimentos.

The screenshot displays the search interface of the RITEC repository. On the left, there are navigation menus for 'Browse' (All of DSpace, Communities, Titles, Authors, Issue Date, Submit Date, Discipline) and 'My Account' (Login, Register). Below these is a 'Filter by Category' section with an 'Author' list including names like Ram3rez Montoya, Mar3a S., and others. The main search area features a search bar with the ID '266632' and a 'Go' button. Below the search bar, there are tabs for 'Publications' and 'Profiles', and a status 'Now showing items 1-10 of 97'. An 'Export' button and a settings icon are also visible. The search results list three items, each with a thumbnail, title, and author information.

Figura 7. Producci3n en RITEC del proyecto “Laboratorio Binacional para la Gesti3n Inteligente de la Sustentabilidad Energ3tica y la Formaci3n Tecnol3gica” (Web: <http://tiny.cc/RITecEnergia>)

Con los tres puntos anteriores que se destacan en los resultados, se responde a la pregunta de investigaci3n ¿c3mo se relacionan los CRIS y los del RI para que sean interoperables?, el primero con el protocolo SWORD configurado en el Dspace donde esta integrado el Repositorio Institucional, el segundo con metadatos que contengan un lenguaje formal, accesible, compartido y ampliamente aplicable entre los dos sistemas y el tercero con lograr que a trav3s de una misma interfaz de captura se organice, gestione, preserve y divulgue la producci3n cient3fica y acad3mica en formato abierto.

## 5. Conclusiones

Los sistemas de gesti3n de informaci3n cient3fica, como son el Repositorio Institucional y el

CRIS, permiten que el público en general pueda tener acceso al conocimiento en formato abierto de forma sistematizada. Por ello, el Repositorio Institucional es una puerta abierta a los recursos científicos generados por el proyecto “Laboratorio Binacional para la Gestión Inteligente de la Sustentabilidad Energética y la Formación Tecnológica”, ya lo señalan Rincón-Flores, Ramírez-Montoya, Serra-Barragán (2017) el dar a conocer las implicaciones de la reforma energética a través de plataformas abiertas permite que los ciudadanos mexicanos y latinoamericanos comprendan la creación de la reforma energética y accedan a formación y recursos orientados a formarlos en dichas temáticas, a través de cursos de formación masiva (MOOC) y Repositorios Institucionales y propiciar que se involucren de una manera activa y aporten nuevas soluciones.

Con la integración del sistema CRIS y el Repositorio Institucional, en el Tecnológico de Monterrey se eliminó la captura doble de la producción científica generada por autores de la comunidad académica y en particular a la producida por el proyecto “Laboratorio binacional de sustentabilidad energética”. Se puede decir que un aspecto importante para la interoperabilidad además de los metadatos, son los protocolos que se implantan en ambas plataformas, porque al contar con un proceso unificado se potencian las ventajas de ambas plataformas y se aumenta la experiencia de uso por parte de los investigadores del proyecto y se realiza la difusión de los productos científicos generados.

Este estudio es de utilidad para otras universidades y centros de investigación que requieren incorporar procesos de depósito de recursos científicos desde una única interface y hacer interoperables varias plataformas. Es un punto a favor que el investigador no vea sólo más sistemas que demanden capturar la misma información para hacerlos disponibles en Internet y puedan ser más visibles.

Por último, la integración de los dos sistemas contribuye a medir el impacto de los Repositorios Institucionales que están en acceso abierto (autoarchivo de producción científica) y hacer posible la comparación de la citación de artículos compartidos en acceso abierto con aquellos que están en acceso restringido para realizar estadísticas que permiten medir la citación.

## **REFERENCIAS**

- Allinson, J., François, S., & Lewis, S. (2008). SWORD: Simple Web-service offering repository deposit. *Ariadne*, (54).
- Asociación Española de Documentación e Información (SEDIC) (2004) *Conceptos asociados al uso de meta información* [versión web]. Recuperado de: <http://www.sedic.es/autoformacion/metadatos/tema2.htm>
- Cantú, F., Ceballos, H., Mora, S., & Escoffié, M. (2005). A Knowledge-based information system for managing research programs and value creation in a university environment. In Proceedings of the Americas Conference on Information Systems – AMCIS. USA. pp: 781-791. August 2005.
- Fideli, R. (1998). *La comparazione*. Milán. Agneli.
- Godin, B. (2006). The knowledge-based economy: conceptual framework or buzzword? *The Journal of technology transfer*, 31(1), 17-30.
- Gómez-Aguilar, D. A., García-Peñalvo, F. J., & Therón, R. (2014). Analítica visual en e-learning. *El profesional de la información*, 23(3).
- González-Pérez, L.I., Glasserman, L. D., Ramírez-Montoya, M.S. y García-Peñalvo, F.J. (2017). Repositorios como soportes para diseminar experiencias de innovación educativa. En Ramírez-Montoya, M.S y Valenzuela, J.R. (coords.), *Innovación Educativa. Investigación, formación, vinculación y visibilidad* (pp. 259-271). México: Editorial Síntesis.
- González-Pérez, L. I., Ramírez Montoya, M. S., Mercado Varela, M. A., Juárez Ibarra, E. A., y Ceballos, H. G. *Aportes de una herramienta de descubrimiento en un repositorio institucional: un estudio de caso*. CIIE 2017. Monterrey, México.
- Jeffery, K., y Asserson, A. (2009). Institutional repositories and current research information systems. *New Review of Information Networking*, 14(2), 71-83.
- Lijphart, A. (1971). Comparative politics and the comparative method. *American political science review*, 65(3), 682-693.
- Príncipe, P., Rettberg, N., Rodrigues, E., Elbæk, M. K., Schirrwagen, J., Houssos, N., y Jörg, B. (2014). OpenAIRE guidelines: supporting interoperability for literature repositories, data archives and CRIS. *Procedia Computer Science*, 33, 92-94.
- Qin, J., Ball, A., y Greenberg, J. (2012). Functional and architectural requirements for metadata: Supporting discovery and management of scientific data. In *International Conference on Dublin Core and Metadata Applications*. 62-71
- Ramírez-Montoya, M. S. y Ceballos, H. G. (2017). Institutional Repositories. En Cantú, F. (Coord.), *Research Analytics. Boosting University Productivity and Competitiveness through Scientometrics*. CRC Press. Disponible en: <https://repositori.itesm.mx/handle/11285/628020>
- Ramírez-Montoya, M. S. y Mendoza-Domínguez, A. (eds) (2017). *Innovación y sustentabilidad energética. Formación con MOOCs e investigación educativa*. Madrid: Narcea
- Recolector de Ciencia Abierta. (2018). *Recolecta*. [Versión web]. Recuperado de: <https://recolecta.fecyt.es>
- Ribeiro, L., de Castro, P., & Mennielli, M. (2016). EUNIS-EUROCRIS Joint Survey on CRIS and IR.

- Rincón-Flores, E., Ramírez-Montoya, M.S., Serra-Barragán, L.A. (2017). Gamificación basada en retos en el curso MOOC la reforma energética en México y sus oportunidades. En Ramírez-Montoya, M.S. y Mendoza-Domínguez, A. (eds), *Innovación y sustentabilidad energética. Formación con MOOCs e investigación educativa* (pp. 57-83). Madrid: Narcea Disponible en: <http://hdl.handle.net/11285/628012>
- Santos-Hermosa, G., Ferran Ferrer, N., y Abadal, E. (2012). Recursos educativos abiertos: repositorios y uso. *El profesional de la información*, 21(2). 136-145. ISSN: 1386-6710
- Sartori, G. (1984). *La política, lógica y método en las ciencias sociales*. México. Fondo de Cultura Económico.
- Velásquez-Duran, A., & Ramírez-Montoya, M. S. (2018). Research Management Systems: Systematic Mapping of Literature (2007-2017). *International Journal on Advanced Science, Engineering and Information Technology*. Version of the manuscript as preprint: [https://osf.io/7dv8c/?view\\_only=c43355c23f4b41f1b0792c9ad0911747](https://osf.io/7dv8c/?view_only=c43355c23f4b41f1b0792c9ad0911747) (in edition).
- Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., y Bouwman, J. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific data*, 3.
- Yañez-Figueroa, J. A, Ramírez-Montoya, M. S. y García-Peñalvo, F. J. (2017). Vinculación universidad-sociedad para la innovación educativa: Los casos de laboratorios ciudadanos. En Ramírez-Montoya, M.S., Valenzuela-González, J. R (eds.), *Innovación Educativa. Investigación, formación, vinculación y visibilidad* (pp. 201-222). Madrid: Síntesis. Disponible en: <http://hdl.handle.net/11285/628016>

## **SOBRE LOS AUTORES**

**Laura Icela González Pérez.** Cursa el doctorado de Formación en la Sociedad del Conocimiento por la Universidad de Salamanca, España. Obtuvo el grado de Maestra en Tecnología Educativa por el Tecnológico de Monterrey y de Licenciatura en Informática Administrativa por la Universidad Tecnológica de México. Su línea de investigación es desarrollo y uso de la tecnología educativa, colabora en el proyecto de investigación “Laboratorio Binacional para la Gestión Inteligente de la Sustentabilidad Energética y la Formación Tecnológica”. Su experiencia profesional radica en la implementación de sistemas para la gestión del conocimiento científico, planificación y gestión de proyectos I + D, desarrollo y uso de prácticas de Ciencia Abierta, diseño de prototipos basados en el Diseño Centrado en el Usuario (DCU) y diseño curricular de programas de formación a profesores e investigadores. Correo: [Laugonzalez@itesm.mx](mailto:Laugonzalez@itesm.mx)

**María Soledad Ramírez-Montoya.** Es Doctora en Filosofía y Ciencias de la Educación por la Universidad de Salamanca (USAL). Sus líneas de investigación están centradas en las estrategias de enseñanza, los recursos tecnológicos para la educación, la formación de investigadores educativos y el movimiento educativo abierto. Es directora de posgrado y educación continua e investigadora titular de la Escuela de Humanidades y Educación del Tecnológico de Monterrey (México) en los programas de maestría y doctorado de educación. Es directora de la Cátedra UNESCO: «Movimiento educativo abierto para América Latina», directora de la oficina del International Council for Open of Distance Education (ICDE): OER Latin America, investigadora principal en la Red Strengthening Information Society Research Capacity Alliance (SIRCA), organizadora principal de la Red Latinoamericana Abierta Regional de Investigación Social y Educativa (CLARISE) y coordinadora del grupo de investigación de «Grupo de Investigación e Innovación en Educación» del Tecnológico de Monterrey. Es miembro del Sistema Nacional de Investigadores en México. Participa en el «Laboratorio Binacional para la Gestión Inteligente de la Sustentabilidad Energética y la Formación Tecnológica» como responsable del subproyecto «Innovación abierta, interdisciplinaria y colaborativa para formar en sustentabilidad

energética a través de MOOCs», con actividades de coordinación del equipo de investigadores, dirección de tesis doctorales y producción científica y académica en el proyecto. Correo: [solramirez@tec.mx](mailto:solramirez@tec.mx)

**Francisco José García Peñalvo** es doctor en Ingeniería en Informática y profesor Titular de Universidad (acreditado a Catedrático de Universidad) del Departamento de Informática y Automática en la Universidad de Salamanca (España) y Profesor Distinguido de la Escuela de Humanidades y Educación del Tecnológico de Monterrey (México), con 3 sexenios de investigación y 3 quinquenios docentes reconocidos. Desde 2006 es el director del Grupo de Investigación Reconocido por la USAL (y Grupo de Excelencia de Castilla y León) GRIAL (GRupo de investigación en InterAcción y eLearning). Es director de la Unidad de Investigación Consolidada de la Junta de Castilla y León (UIC 81). Ha sido Vicedecano de Innovación y Nuevas Tecnologías de la Facultad de Ciencias de la USAL entre 2004 y 2007 y Vicerrector de Innovación Tecnológica de esta Universidad entre 2007 y 2009, periodo en el que implantó el Campus Virtual, el Repositorio Institucional y el Repositorio Scopeo, entre otros hitos relacionados con la innovación. Ha editado números especiales en múltiples revistas nacionales e internacionales indexadas en los principales índices de referencia y es el editor científico de Education in the Knowledge Society (EKS) journal y de Journal of Information Technology Research (JITR). Actualmente es el Coordinador del Programa de Doctorado en Formación en la Sociedad del Conocimiento de la USAL. Correo electrónico: [fgarcia@usal.es](mailto:fgarcia@usal.es)

**Héctor Ceballos**. Director de Cienciometría e Investigador de la Escuela de Ingeniería y Ciencias del Tecnológico de Monterrey. Colabora con el grupo de investigación en Sistemas Inteligentes y es miembro de tiempo parcial de la Maestría en Sistemas Inteligentes. También es miembro adherente de la Academia Mexicana de Computación (AMEXCOMP). El Dr. Ceballos es Ingeniero en Sistemas Computacionales por el Instituto Tecnológico de Veracruz y tiene una maestría y un doctorado en Sistemas Inteligentes ambos en el Tecnológico de Monterrey, Campus Monterrey. En su producción científica se encuentran más de 30 publicaciones entre artículos de revista, memorias de congresos, capítulos de libro y reportes técnicos. Correo: [ceballos@tec.mx](mailto:ceballos@tec.mx)

**Erika Juárez**. Es Licenciada en Ciencias de la Computación por la Universidad Autónoma de Nuevo León y Maestra en Administración de Tecnologías de Información por el Tecnológico de Monterrey. Responsable de la Coordinación de Datos para Investigación en la Dirección de Investigación del Tecnológico de Monterrey. Tiene más de 10 años de experiencia en el desarrollo y administración de Sistemas CRIS, es actualmente responsable del desarrollo y de la administración del CRIS-Tec así como de las colecciones de Producción Científica en el Repositorio Institucional del Tec (RITEC). Colabora con el Grupo de Investigación e Innovación en Educación del Tecnológico de Monterrey. Correo: [ujuarez@itesm.mx](mailto:ujuarez@itesm.mx)