

17

Fecha de presentación: abril, 2024
Fecha de aceptación: junio, 2024
Fecha de publicación: septiembre, 2024

GESTIÓN DE DATOS

DE INVESTIGACIÓN EN LA UNIVERSIDAD DE CIENFUEGOS

RESEARCH DATA MANAGEMENT AT THE UNIVERSITY OF CIENFUEGOS

Arelexys Antonia García Cartaya ^{1*}

E-mail: aagarcia@ucf.edu.cu

ORCID: <https://orcid.org/0000-0003-3885-6907>

Dayme García Valdes ¹

E-mail: dgvaldes@ucf.edu.cu

ORCID: <https://orcid.org/0000-0001-8542-7318>

Geidy Medina Ruíz ¹

E-mail: gmedina@ucf.edu.cu

ORCID: <https://orcid.org/0000-0003-0608-2273>

¹ Universidad de Cienfuegos "Carlos Rafael Rodríguez" de Cienfuegos, Cuba.

*Autor para correspondencia

Cita sugerida (APA, séptima edición):

García Cartaya, A. A., García Valdes, D. & Medina Ruíz, G. (2024). Gestión de Datos de Investigación en la Universidad de Cienfuegos. *Universidad y Sociedad*, 16(5), 173-182.

RESUMEN

El enfoque de Ciencia Abierta se ha ganado protagonismo a nivel mundial desde mediados del siglo XX en aras de lograr una mayor visibilidad de la investigación científica. En el escenario cubano, ha ganado espacio de desarrollo a partir de la evolución de las tecnologías de la información y de la alta actividad científico investigativa desplegada en las instituciones de la Educación Superior. En este sentido, se han realizados diversos proyectos y estrategias para la implementación de la Ciencia y la Educación Abierta de aras de potenciar el acceso al conocimiento científico. La presente investigación tiene como principales objetivos analizar diversas definiciones de Ciencia Abierta, y el estado de conocimiento sobre la Gestión de Datos de Investigación en la Universidad de Cienfuegos Carlos Rafael Rodríguez. La investigación se sustenta en una metodología de enfoque mixto, ya que se presenta un análisis cualitativo del tema objeto de estudio, así como, los resultados cuantitativos de un diagnóstico aplicado en esta institución académica. Dichos resultados constituyen una parte de una amplia investigación realizada en 4 universidades latinoamericanas para el diseño de políticas institucionales para la Gestión de Datos de Investigación y con ello fomentar el desarrollo de la Ciencia Abierta en este contexto regional. La investigación concluye con la impronta de realizar actividades de capacitación para el manejo y uso de la Gestión de Datos de Investigación por parte de los investigadores.

Palabras clave: Ciencia Abierta, Datos Abiertos, Gestión de Datos de Investigación, Universidad de Cienfuegos.

ABSTRACT

The Open Science approach has gained prominence worldwide since the mid-20th century in order to achieve greater visibility of scientific research. In the Cuban scenario, it has gained space for development based on the evolution of information technologies and the high scientific research activity deployed in Higher Education institutions. In this sense, various projects and strategies have been carried out for the implementation of Science and Open Education in order to enhance access to scientific knowledge. The main objectives of this research are to analyze various definitions of Open Science, and the state of knowledge about Research Data Management at the Carlos Rafael Rodríguez University of Cienfuegos. The research is based on a mixed approach methodology, since a qualitative analysis of the topic under study is presented, as well as the quantitative results of a diagnosis applied in this academic institution. These results constitute part of extensive research carried out in 4 Latin American universities for the design of institutional policies for Research Data Management and thereby promote the development of Open Science in this regional context. The research concludes with the imprint of carrying out training activities for the management and use of Research Data Management by researchers.

Keywords: Open Science, Open Data, Research Data Management, University of Cienfuegos.

INTRODUCCIÓN

A partir del modelo de las revoluciones científicas de Kuhn (1962), la Ciencia Abierta (CA) representa en la actualidad un cambio de paradigma en la producción de la investigación, y por tanto supone transformaciones sustanciales en sus procedimientos. Dicho proceso pretende impulsar prácticas que permiten compartir los datos, las publicaciones, las metodologías, las herramientas de análisis y, en algunos casos, también las agendas de investigación. Tal como mencionan algunos autores:

La ciencia abierta es un cambio de paradigma en la manera de hacer ciencia. Ésta no cambia sustancialmente con respecto a sus motivaciones y objetivos, pero sí lo hace (sustancialmente) en cuanto a sus métodos. El cambio no está en lo que se hace, sino en cómo se hace. (Anglada y Abadal, 2018).

Fry et al., (2009) destacaban una doble dimensión de esta «apertura» en la ciencia: en primer lugar, apertura de las fases de investigación (desde compartir datos hasta la difusión de los resultados en abierto) y, en segundo lugar, apertura en el grado de colaboración entre los agentes implicados.

Más tarde, Vicente-Sáez y Martínez-Fuentes en el 2018 realizan una revisión bibliográfica de setenta y cinco estudios sobre CA publicados entre 1985 y 2016 para encontrar las características fundamentales de este movimiento (apertura, transparencia y trabajo colaborativo) (Abadal, 2021). En esta revisión se evidencian importantes aportaciones de organismos, que han sido especialmente relevantes para perfilar y divulgar el concepto de CA. Tal es el caso de la Comisión Europea, la cual ha tenido un papel fundamental, tanto desde el punto de vista teórico como en la práctica, promoviendo el desarrollo de políticas de apoyo a este nuevo modelo de investigación y comunicación social de la ciencia.

Digital Science in European Commission Decision (2017) expone cómo Internet y otras tecnologías estaban cambiando la creación y difusión del conocimiento y, por lo tanto, creando una nueva forma de hacer ciencia. Los informes de Abadal (2021) situaban a la CA y la innovación como los motores del crecimiento y del bienestar económico de Europa, y detallaba las acciones que se estaban llevando a cabo y las que se deberían realizar para caminar hacia el nuevo horizonte establecido. Algunas de las definiciones emitidas por organizaciones impulsoras de la CA son:

La Ciencia Abierta se refiere a un nuevo enfoque del proceso científico basado en el trabajo cooperativo y nuevas formas de difundir el conocimiento, mejorando la accesibilidad y la reutilización de los resultados de la investigación mediante el uso de tecnologías digitales y nuevas herramientas colaborativas (Comisión Europea, 2018, p.12).

La Ciencia Abierta es la práctica de la ciencia de tal manera que otros puedan colaborar y contribuir, donde los datos de la investigación, las notas de laboratorio y otros

procesos de investigación están disponibles gratuitamente, bajo términos que permitan la reutilización, redistribución y reproducción de la investigación y sus datos y métodos subyacentes (Foster, 2019).

Según indica Ortiz-Núñez (2021) la CA “abarca todo el ciclo de producción del conocimiento científico, incluida la concepción, la recogida de datos, su procesamiento, publicación y distribución o la reutilización y la evaluación de los resultados” (p.6). De ahí que contemple varios elementos de las prácticas científicas (Figura 1) como los códigos, los recursos, los canales para la recuperación de la información y los datos abiertos.

Dentro de los componentes de la CA se encuentran los Datos Abiertos (DA), una práctica que persigue que determinados datos estén disponibles de forma libre y pública sin restricción de copyright, patentes u otros mecanismos de control (Junta de Castilla y León, 2012). No obstante, su nivel de desarrollo es aún bajo, en especial en comparación con el denominado Open Access (OA) del sistema de publicaciones.

De manera particular, la literatura asegura que las primeras iniciativas para compartir datos resultan dictadas por Cook-Deegan y McGuire (2017). Sin embargo, no fue hasta el surgimiento del programa European Commission Decision (2020) que se establece la obligatoriedad de poner en abierto los datos subyacentes a las publicaciones de los proyectos financiados por la Comisión Europea, con excepción de aquellos que estuvieran afectados por requisitos de privacidad y seguridad.

Fig 1: Componentes de la Ciencia Abierta.



Fuente: Repositorio Institucional CONICET Digital.

La apertura de los datos tiene como objetivo fundamental facilitar su reutilización por parte de otros investigadores, por lo que no se limita a la difusión del contenido de la publicación, sino que además ofrece mayor transparencia en el proceso investigativo, se evita el fraude y se fundamenta la integridad de la ciencia. Para ello se han desarrollado los principios FAIR (siglas en inglés), cuyo objetivo fundamental es hacer que los datos científicos sean libremente localizables, accesibles, interoperables y reutilizables, promoviendo así: la colaboración entre investigadores, la génesis de nuevas investigaciones y la socialización del conocimiento (Ortiz y Nuñez., 2021). Esto apunta a que los DA de carácter científicos deben ser etiquetados con metadatos y ser almacenados bajo licencias abiertas para su posterior recuperación y reutilización.

Otra de las ventajas de tener fuentes de DA es que se convierten en fuentes de información para desarrollar aplicaciones web y software que faciliten el trabajo de los investigadores desde el punto de vista organizativo. También sirven como fuentes externas de consulta para iniciativas Big Data.

Varios autores defienden las ventajas de compartir los datos, pero sigue existiendo temor al hacerlo por no disponer de incentivos suficientes para estimular esta colaboración, y no se realizan de forma adecuada la citación de los ficheros de datos utilizados. En este sentido, la mayoría de los repositorios ya facilitan un Digital Object Identifier (DOI) a cada conjunto de datos archivados y, además, añaden un modelo de citación para que pueda recibir la justa referencia (Abadal, 2021).

Lo anterior es importante porque las referencias bibliográficas se convierten también en datos de investigación, pero ¿Qué se entiende por datos de investigación?

Los datos incluyen: cuadernos de laboratorio, cuadernos de campo, datos de investigación primaria, cuestionarios, cintas de audio, videos, desarrollo de modelos, fotografías, películas, objetos digitales, algoritmos, scripts, bases de datos, metadatos y esquemas de metadatos, configuraciones de software y las comprobaciones y las respuestas de la prueba (Abadal, 2012). No son considerados datos finales de investigación: las notas de laboratorio, los sets de datos parciales, análisis preliminares, borradores de trabajos, planes para investigaciones futuras, comunicaciones con colegas, objetos físicos y ejemplares de laboratorio. Los tipos de datos pueden clasificarse en:

- Numéricos, descriptivos o visuales.
- Según su naturaleza: cualitativos o cuantitativos.
- Según su nivel de procesamiento: En estado bruto (datos primarios), procesados o analizados.
- Según la fuente de la que provienen pueden ser experimentales (ej. cromatografías), observacionales (ej. encuestas) y computacionales (obtenidos mediante simulación).
- Según su formato: Textuales (Word, PDF, RTF, etc.), Numéricos (Excel, CSV, etc.), Multimedia (JPEG, MPEG, WAV, etc.), Estructurados (XML, MySQL, etc.), Código de software (Java, C, etc.), específicos de un software (Mesh, 3D CAD, modelo estadístico, etc.), específicos de una disciplina o instrumento.

Gestión de Datos de Investigación (GDI) o Research Data Management (RDM)

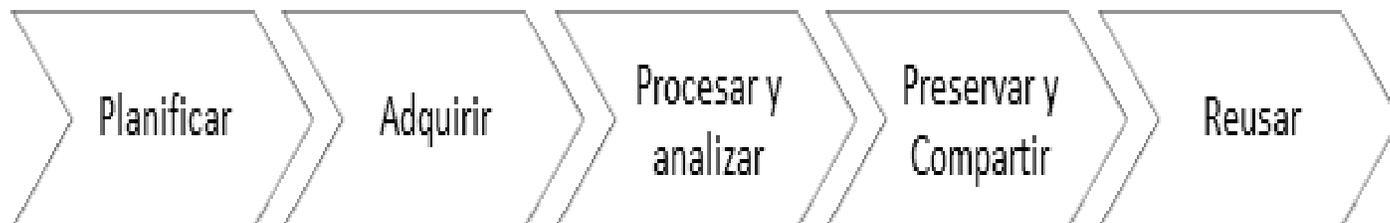
La GDI está presente en todas las fases de la investigación y engloba la recopilación, organización, documentación, almacenamiento y preservación de los datos utilizados o generados durante un proyecto de investigación. Una gestión de datos adecuada ayuda a los investigadores a llevar a cabo una mejor investigación y supone:

- Cumplir con los requerimientos de los organismos financiadores.
- Mayor transparencia para la validación de los resultados de la investigación.
- Mejorar la protección de datos y minimizar el riesgo de pérdida de los mismos.
- Asegurar que los datos sean localizables, accesibles, interoperables y reutilizables: datos FAIR.
- Ahorrar tiempo evitando duplicaciones y haciendo un uso eficiente de los recursos disponibles.
- Mejorar el perfil del investigador, el impacto y la visibilidad de los proyectos.

La GDI involucra la organización y mantenimiento activo de los datos durante todo el desarrollo de la investigación, enfocándose en la confiabilidad y replicabilidad de los productos finales. Se puede decir que la **GDI es una forma de implementar mejores prácticas en todo el ciclo de vida de un proyecto de investigación** e incluye 5 etapas (Figura 2). Los beneficios incluyen mayor transparencia de los resultados, la optimización del tiempo y los recursos

disponibles, la reutilización de los datos en otros proyectos, y el impacto y visibilidad del proyecto final (Marín-Arraiza et al., 2019).

Fig 2: Etapas del ciclo de la Gestión de Datos de Investigación.



Fuente: Elaboración propia a partir de Marín-Arraiza et al. (2019)

En este primer paso se recomienda preparar un Plan de Gestión de Datos, el cual debe estar compuesto por un grupo de preguntas de orden lógico para ser respondidas o solucionadas durante el estudio. Dichas preguntas están enfocadas a la tipología de datos y metadatos, la localización de los mismos, y nivel de privatización para las cuestiones legales.

En este paso se procede a la obtención de los datos necesarios a través de la experimentación, simulación, medición u observación; todo dependiendo de la disciplina o área del conocimiento. También es buen momento para preparar los metadatos y el material complementario para capturar la información de manera consistente durante el proyecto. Para el almacenamiento, deberán tenerse en cuenta los requerimientos de espacio específicos del proyecto, así como el tiempo que permanecen almacenados.

Con los datos almacenados se realiza el correspondiente procesamiento, lo que implica que los datos deben ser validados, limpiados, recodificados, y de ser necesario, anonimizados. Estos aspectos son esenciales para asegurar la calidad de los resultados. Posteriormente se analizan los datos para producir nuevos hallazgos o conocimientos y presentarlos ya sea como una publicación, artículo, reporte, tesis, y otros.

Inmediatamente después de realizar el análisis, es muy importante guardar los datos en un formato abierto, crear un diccionario de datos y actualizar los metadatos para su publicación. Después de tener en cuenta cuestiones de privacidad de los datos, permisos para compartir la información y los derechos de propiedad intelectual, es momento de publicar los datos en un servidor público o un repositorio.

Completados los pasos anteriores puede decirse que el proyecto de investigación cuenta ya con una GDI implementada, documentada y abierta al público que puede ser utilizado por el propio investigador u otros académicos en análisis secundarios, o en proyectos de investigación completamente nuevos.

Estado de la CA en la Educación Superior en Cuba

Aunque algunas de las herramientas que se muestran en la Figura 1 son usadas en los procesos investigativos de las universidades en Cuba, es de reconocer que aún no se tiene una cultura de educación y ciencia abierta (ECA). Dentro del marco de los convenios de colaboración VLIR-UOS con instituciones universitarias cubanas se han desarrollado iniciativas para implementar aplicaciones y bases de datos OA. En este sentido, resulta importante destacar que la Universidad Central de Las Villas (UCLV) en trabajo colaborativo con la Universidad de Ciencias Informáticas son líderes en cuestiones de CA a través del proyecto iniciado en 2005 con el Sistema Automatizado de Gestión Bibliotecaria (Meneses et al., 2022).

Al tiempo se han realizado acciones para el desarrollo de la web semántica, los metadatos, los recursos educativos abiertos a través de la plataforma MOODLE y los repositorios digitales institucionales con Dspace. Todos estos proyectos anteceden la necesidad de elevar el conocimiento sobre las licencias legales asociadas al movimiento Open

Source para el depósito de datos de investigación, así como, la fomentar el diseño de políticas públicas e institucionales para la gestión de datos de investigación

La Universidad de Cienfuegos “Carlos Rafael Rodríguez” (UCf) en los últimos cinco años ha llevado a cabo algunas acciones en pos de fomentar la Educación y Ciencia Abierta, sobre todo a través de publicaciones y cursos online masivos abiertos (MOOCs). La Editorial Universo Sur dispone su sistema de revistas en Open Access y recientemente ha iniciado *Open Peer Review* de acuerdo con los principios FAST. Paralelamente se potencia desde el Centro de Recursos para el Aprendizaje y la Investigación (CRAI) la alfabetización de docentes e investigadores en el manejo de sus redes sociales académico, lo que permitirá incrementar el número de investigadores con perfiles fundamentalmente en ORCID, Google Académico y ResearchGate.

MATERIALES Y MÉTODOS

Teniendo en consideración estos elementos, el presente estudio tiene como objetivo describir el estado de conocimiento y aplicación de la GDI en la UCf. Se basa en una metodología de investigación de carácter mixto pues se realizan análisis cualitativos a partir de resultados cuantitativos.

Para darle cumplimiento a los objetivos trazados se emplean varios métodos y técnicas de investigación del nivel teórico y empírico. Los métodos analítico-sintético e inductivo-deductivo permitieron penetrar en la esencia del fenómeno objeto de estudio y arribar a generalizaciones en torno a la CA, los DA como uno de sus componentes, y la GDI. Dentro de las técnicas empíricas, el análisis de contenido permitió identificar el conocimiento existente sobre el objeto de estudio a partir de la revisión sistemática de fuentes documentales.

Para el desarrollo del diagnóstico el CRAI aplicó en el periodo de octubre-noviembre de 2023 un cuestionario sobre GDI a una muestra de los directivos de la actividad de Ciencia, Tecnología e Innovación en la UCf. Esta iniciativa forma parte del proyecto “Desarrollo de una Estrategia para la Gestión de Datos de Investigación en la Educación Superior e Instituciones Científicas dentro del contexto Latinoamericano (RDMS-LatAm)” que coordina la Universiteit Hasselt de Bélgica y la Universidad Central “Marta Abreu” de Las Villas (UCLV), y donde la UCf actúa como socio participante.

Dicho cuestionario fue formulado por un equipo multidisciplinario de especialistas del ámbito de las Ciencias de la Información de las 4 universidades miembros del proyecto mencionado anteriormente (UCLV, Universidad Nacional Mayor de San Marcos de Perú, Universidad Católica Boliviana San Pablo y la UCf) y se tomó como referente una encuesta diseñada por la Universiteit Hasselt para medir el conocimiento de la GDI en el entorno europeo universitario.

El cuestionario estuvo compuesto por 16 preguntas, segmentado en tres dimensiones relacionadas con el manejo, uso y almacenamiento de los datos de investigación: Tipos y formatos; Almacenamiento, archivado, copias de respaldo y pérdida; Infraestructura y servicios.

Como objeto de análisis se tomó una muestra de 25 directivos de la UCF, correspondientes a las 7 áreas del conocimiento que existen en esta institución académica (Figura 3). El total de la muestra ostenta el grado científico de Doctor en Ciencias y la categoría docente de Profesor Titular. También se tuvo en cuenta la experiencia en la actividad investigativa y cierto equilibrio de género (Figura 4). En esta selección se tuvo en cuenta principalmente al personal directivo del Vicerrectorado de Investigación y Posgrado, Vicerrectorado de Formación, de algunos Centros de Estudios¹ y las diferentes facultades, incluyéndose al director de un Centro Universitario Municipal.

Por su parte los investigadores se representan en una muestra de 169 personificadas en todas las áreas del conocimiento y en diferentes categorías científicas que ostentan los investigadores de la Universidad de Cienfuegos Carlos Rafael Rodríguez.

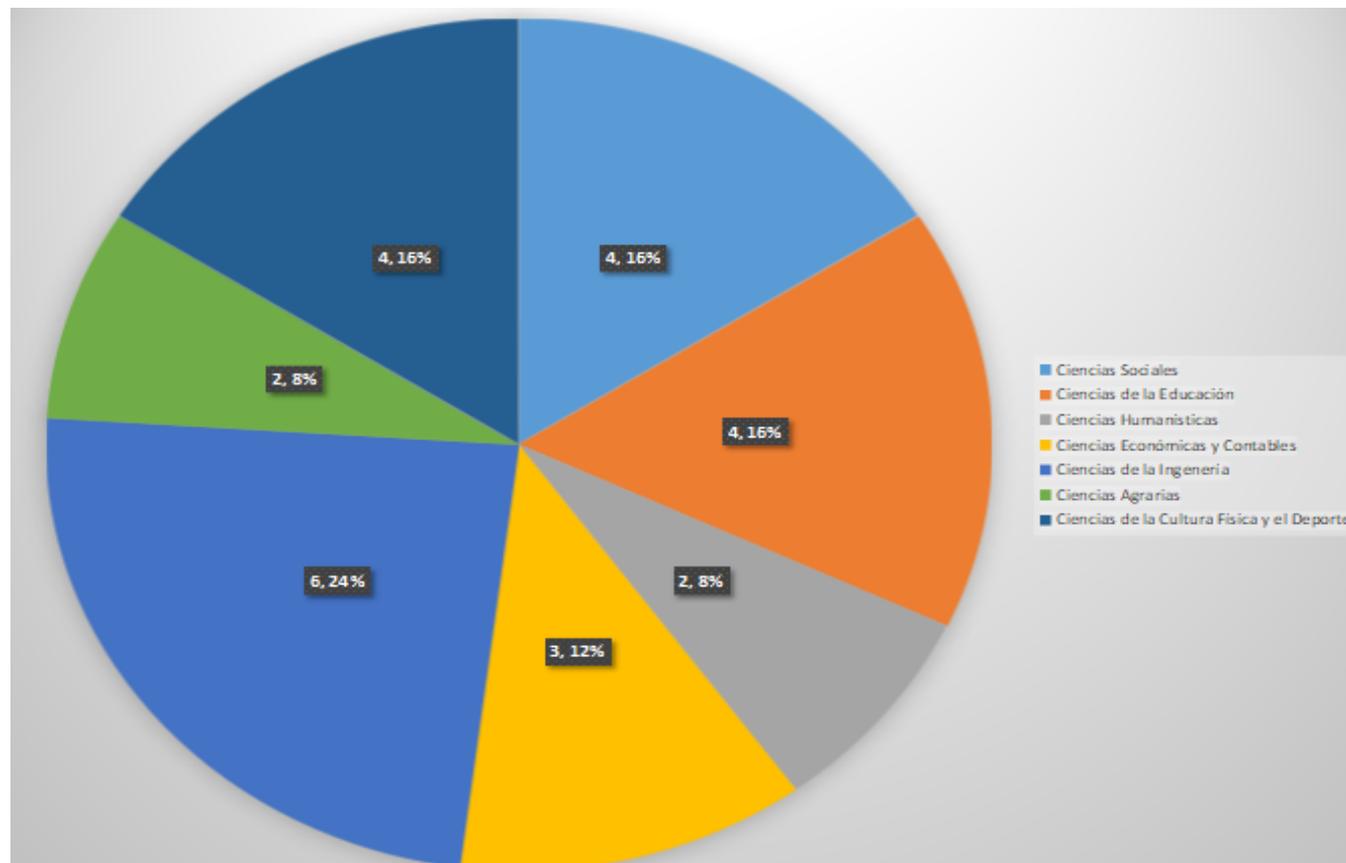
1 Centro de Estudios de Energía y Medio Ambiente (CEEMA)

Centro de Estudio de la Didáctica y Dirección de la Educación Superior (CEDDES)

Centro de Estudio para la Transformación Agraria Sostenible (CETAS)

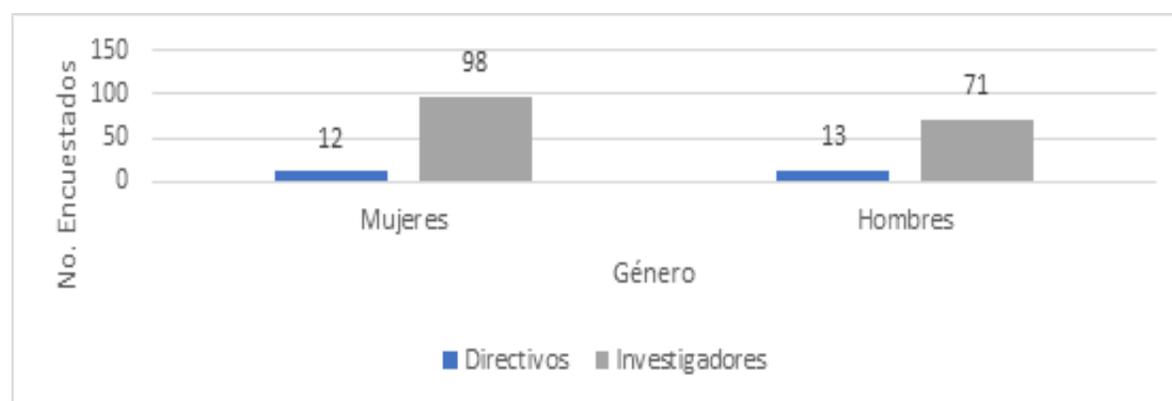
La representación de la mujer en la actividad científica de la universidad es de vital importancia por lo que en la Figura 4 se refleja la participación de la mujer investigadora y directiva, con un 12% de participación en las encuestas aplicadas para el desarrollo de las actividades de la Ciencia Abierta que se expone en la presente indagación respecto a las directivas y 13 % a las investigadoras.

Fig 3: Representatividad de la Muestra según Áreas del Conocimiento de la UCf (n=25).



Fuente: elaboración propia.

Fig 4: Representatividad de la Muestra según Género (n=25).



Fuente: elaboración propia.

RESULTADOS Y DISCUSIÓN

Del total de 16 preguntas formuladas en el cuestionario se presentan a continuación los resultados de aquellas con mayor valor agregado para los fines de este estudio.

La investigación científica utiliza una gran cantidad de datos de fuentes diversas, a la vez que es generadora de sus propios datos por la aplicación sistemática de técnicas y metodologías para la recopilación, el procesamiento y el almacenamiento de datos.

Según el criterio de los encuestados, la tipología de los datos usados en sus estudios depende del contexto particular de la investigación, de los objetivos propuestos, la metodología y por supuesto de área del conocimiento a la cual responden. No obstante, existe una prominencia de hacer uso de datos o conjunto de datos generados previamente por el propio investigador (Tabla 1) y en menor medida aquellos que son aportados por entidades que solicitan el pago por el uso de los datos de investigación que ofrecen (proveedores comerciales). En tanto, la consulta y empleo de los datos en abierto disponibles en repositorios, como por ejemplo ZENODO, aún es escasa debido al desconocimiento de estos recursos y a la falta de confiabilidad en cuanto a la preservación de los datos.

Tabla 1: Fuente de obtención de Datos para la Investigación (n>25 por opción múltiple de respuesta).

Fuentes de obtención de datos	No.	%
De organizaciones públicas	3	12
De proveedores comerciales	1	4
Datos abiertos disponibles públicamente	6	24
Datos generados por usted mismo	15	60

Fuente: elaboración propia.

Por otra parte, existe un balance en cuanto al tipo de contenido de los datos de investigación generados y utilizados en el proceso investigativo por académicos de la UCf. Como se observa en la Tabla 2 la tendencia ha sido el manejo de datos cualitativos y cuantitativos, motivado significativamente por el empleo de un enfoque metodológico mixto y por auge que han tenido las perspectivas multi y transdisciplinarias en la actividad científica.

Tabla 2: Tipología de Datos de Investigación generados por los investigadores (n>25 por opción múltiple de respuesta).

Tipos de Datos de Investigación	No.	%
Datos cuantitativos	25	100
Datos cualitativos	23	92
Datos primarios o datos en bruto	10	40

Datos procesados	11	44
Datos clínicos	1	4
Material audiovisual	5	20
Fotografías	6	24
Cuadernos de laboratorio	2	8

Fuente: elaboración propia.

En menor medida se mencionan los datos clínicos y los cuadernos de laboratorio, pues en la UCf no es muy común el desarrollo de investigaciones que impliquen la experimentación animal o la utilización de muestras biológicas de seres humanos. Al tiempo que los datos procesados representan una importante fuente de información para las especialidades de las Ingenierías y las Ciencias Agrarias.

La mayor parte de los encuestados indica que los datos o conjunto de datos se generan en formato digital, siendo preponderantes el documento de texto, las hojas de cálculo en Excel y los gráficos (Tabla 3). El primero (DOC, ODF, PDF, TXT, etc.) sigue siendo el formato de mayor uso y aunque no se muestra en la tabla, se emplea en combinación con otros, como la imagen y el software. Debido a su propia área del conocimiento y a la experticia obtenida en la formación investigativa, alegan la utilidad de los formatos para presentar datos contables y numéricos que posibiliten contratar o agregar valor a los propios resultados cualitativos. Por ende, también son referidas las distintas aplicaciones informáticas (8, 32%) para la captura, organización y procesamiento de datos.

Tabla 3: Tipología de los Formatos Digitales de los datos generados en la investigación (opción múltiple de respuesta).

Formatos digitales de los datos	No.	%
Documentos de texto	15	60
Texto estructurado (HTML, XML, etc.)	5	20
Hoja de Cálculo	11	44
Base de datos	7	28
Gráficos / Imagen	10	40
Audio	4	16
Video / Película	4	16
Aplicación de Software Código fuente	4	16
Datos de Configuración	2	8
Aplicación de Software	8	32
Código de Fuente / Código de programas	4	16

Fuente: elaboración propia.

El segundo bloque de preguntas del cuestionario tenía como objetivo conocer los criterios de los directivos en

relación al depósito y conservación de los datos, ya fuera a nivel personal o en sistemas de información creados por la institución. En tal sentido, la tendencia ha sido la salvaguarda central en las instalaciones de almacenamiento de la unidad de investigación (15, 60%), y se promueve la utilidad del Servicio en la Nube de la universidad (14, 56%) por la posibilidad de acceder a estos datos desde cualquier equipo electrónico conectado a la red (Tabla 4). También resulta prioridad el almacenamiento en portátiles, así como en otros dispositivos tecnológicos de carácter personal.

Tabla 4: Almacenamiento de los Datos de Investigación (opción múltiple de respuesta).

Almacén Habitual	No.	%
En la unidad de Investigación	15	60
En la Universidad	9	36
A nivel regional /Nacional	0	0
Instalaciones específicas de la Disciplina	3	12
Un centro de datos externo	0	0
Servicio en la nube de la Universidad	14	56
Servicio en la nube externa	3	12
Dispositivos individuales	8	32
PC, computadora portátil	12	48

Fuente: elaboración propia.

Sin embargo, tal como se aprecia en la Tabla 4 no se mencionan los centros de datos externo ni aquellas instalaciones de almacenamiento a nivel regional/nacional. Al tiempo que se muestra un desconocimiento de sistemas propios para las disciplinas o áreas específicas del conocimiento, así como, temor a la pérdida de los datos al colocarlos en servicios en la nube externos a la institución.

Paralelamente el cuestionario incluía una pregunta sobre la preferencia en cuanto a la disponibilidad de opciones para el almacenamiento de los datos de investigación (Tabla 5). En este caso, las respuestas varían un poco con respecto al resultado anterior, evidenciando la importancia que le otorgan los directivos a tener un servicio propio en la UCf (18, 72%) que cumpla con este principio. Igualmente, el Servicio en la Nube de la Universidad (15, 60%) se visualiza como una necesaria opción ante el déficit de recursos tecnológicos en los departamentos docentes y los cortes de electricidad por la situación energética del país.

Tabla 5: Preferencia para el Almacenamiento de los Datos de Investigación (opción múltiple de respuesta).

Servicio de almacenamiento de preferencia por el investigador	No.	%
En la Universidad	18	72
En la unidad de Investigación	6	24
Un centro de datos externo	2	8
Servicio en la Nube de la Universidad	15	60
Dispositivos individuales	8	32
PC, computadora portátil	3	12
Otros	8	32

Fuente: elaboración propia.

Ante la interrogante sobre la responsabilidad para el archivo de los datos de investigación de su unidad, los resultados muestran una correspondencia con lo señalado respecto a los servicios para el almacenamiento (Tabla 6). Para los encuestados este criterio forma parte de su modo de actuación como director de un área o grupo de investigadores (14, 56%). Aunque se reconoce como tendencia que cada investigador preserva sus datos (14, 56%) y en aquellos estudios que son realizados por equipos de trabajo, es el investigador principal (8, 32%) quien debe controlar el ciclo de vida de los datos, desde su origen hasta su posible destrucción en caso de que así se decida.

Tabla 6: Responsabilidad para el Archivo de los Datos de Investigación (opción múltiple de respuesta).

Responsable	No.	%.
Yo, como jefe de la unidad de investigación	14	56
Investigador principal	8	32
Cada investigador archiva sus propios datos	14	56
Un investigador asignado	2	8
Secretaría/personal administrativo	2	8
Personal de IT/Informatización	5	20
Administrador de datos de la universidad	3	12
Administrador de datos de la unidad	2	8
Proveedor de servicios externo	0	0

Fuente: elaboración propia.

De manera similar, el proveedor del servicio externo no se asume como una figura responsable para tal objetivo ante el temor de la pérdida de los datos, el plagio, el uso indebido y el incumplimiento de las directrices para el derecho de autor y la protección de la propiedad intelectual.

En cuanto al bloque de preguntas sobre la Infraestructura y los servicios, los investigadores de la UCf reconocen la necesidad de que la UCf posea directrices o políticas

para la gestión y tratamiento de datos de investigación, lo que contribuye a mejorar la gestión de la actividad de ciencia, tecnología e innovación y aportaría mayor visibilidad a los propios resultados de la investigación. De ahí que proporcionar infraestructura tecnológica y soporte técnico haya sido de los más demandados con un 101% de aceptación, seguido por la posibilidad de considerar la creación/reutilización/citación de datos como resultado científico relevante para la evaluación representado por el 78 % (Tabla 7).

Esta interrogante evidencia la impronta de diseñar e implementar en la UCf en un primer momento un repositorio institucional (Dspace) para proveer el acceso a la producción científica de la institución y posteriormente crear un repositorio de Datos de Investigación para asegurar la preservación a largo plazo de los mismos. No obstante, ninguno de los directivos señaló el Asesoramiento Legal como una acción importante a realizar en la UCf, lo que contradice los criterios respecto a las licencias para la protección de los derechos de autor.

Tabla 7. Apoyo de la institución académica para la GDI (opción múltiple de respuesta).

Acciones adicionales espera de la institución	Investigadores
Directrices o políticas para la gestión y tratamiento de datos de investigación	121
Proporcionar infraestructura tecnológica y soporte técnico	101
Considerar la creación/reutilización/citación de datos como resultado científico relevante para la evaluación	78
Incluir la gestión de datos de investigación en las responsabilidades del trabajo	87
Incluir la gestión de datos de investigación en el plan de estudios	67

Fuente: elaboración propia.

Finalmente, ante la pregunta para crear acciones que favorezca la GDI en la UCf se propuso lo siguiente (Tabla 8). Sin embargo, la redacción e implementación de Directrices o Políticas para el tratamiento de datos de investigación no recibió ningún puntaje, siendo un requisito fundamental para poder desarrollar adecuadamente la GDI.

Tabla 8: Propuesta de acciones de apoyo de la institución académica para la GDI (opción múltiple de respuesta).

Acciones	No.	%
Incluir la gestión de datos de investigación en las responsabilidades de trabajo	7	28
Emplear personal calificado para tareas de gestión de datos de investigación	7	28
Incluir la gestión de datos de investigación en el plan de estudio	3	12
Creación de datos para su reutilización/citación como resultado científico.	7	28
Proporcionar infraestructura técnica/ soporte técnico	5	20

Fuente: elaboración propia.

En función de este resultado, una de las finalidades de este instrumento es precisamente confeccionar las políticas que permitan orientar al investigador y a la universidad sobre el manejo y uso de datos, así como las cuestiones legales asociadas a la apertura y compartición de datos.

La CA es una visión transformadora de la investigación científica, basada en la implementación generalizada de los valores de la colaboración científica, la accesibilidad y la transparencia del proceso. Implica la transformación en la cultura científica que afecta una serie de prácticas en la producción y socialización del conocimiento científico para beneficio de toda la sociedad.

La GDI constituye en la actualidad una herramienta necesaria para la integridad y comunicabilidad de los datos dentro del contexto de la CA. Para que el modelo sea funcional es imprescindible contar con una política explícita y clara en cuanto a qué, cómo, para qué conservar y compartir los datos.

El diagnóstico sobre el estado de conocimiento de la GDI en la UCf se realizó con una muestra selectiva de los directivos de la actividad científica. Dicha acción representa la primera fase del proyecto citado anteriormente, en aras de identificar las oportunidades de la GDI para la mejora de los procesos de ciencia, tecnología e innovación de esta institución académica, y también, las dificultades existentes en cuanto a la infraestructura técnica que limita tanto el archivo como la curaduría de los datos.

La preocupación en relación a los derechos de propiedad intelectual y los intereses legítimos de los propietarios de los datos, es una realidad tangible que atenta contra la transparencia e integridad de los procesos y resultados científicos.

La aplicación del cuestionario representó para muchos directivos el descubrimiento de un nuevo conocimiento,

pero sobre todo, de un enfoque metodológico que amplía las perspectivas de la UCf en torno a la CA.

Existe la necesidad de fomentar el manejo y uso adecuado de los recursos y datos abiertos, así como, de una cultura de trabajo sobre la base de los Planes de Gestión de Datos. También es necesario promover el empleo de herramientas informáticas que permitan reducir el riesgo de pérdida de datos, y aumentar el impacto de la investigación al reutilizar los datos y recibir citaciones por el trabajo realizado.

CONCLUSIONES

En sentido general, los resultados evidenciaron que los datos de investigación siguen estando en unidades particulares, ya sea por falta de infraestructura o por desconfianza hacia servicios de archivo fuera del ámbito institucional. Los números reflejan la urgencia de disponer de recursos tecnológicos adecuados para desarrollar la GDI, y de manera muy particular para implementar un repositorio en la UCf que posibilite la descentralización de la información y evitar la duplicidad de posibles temas de investigación. Por otra parte, los resultados de este diagnóstico le otorgan un rol protagónico al CRAI de la UCf ante la perspectiva de convertirse en un centro de apoyo para las diferentes tareas de la GDI. En tal sentido se prevé a corto y mediano plazo:

- Diseño de una política institucional para la GDI que regule todo el ciclo de vida de los datos en la UCf.
- Capacitación de los investigadores o la denominada alfabetización en datos para el periodo 2023-2024 (confección de guías, tutoriales, encuentros presenciales).
- Realización de talleres y conferencias sobre la Ciencia Abierta.
- Capacitación de los bibliotecarios en cuanto a las licencias abiertas disponibles para los conjuntos de datos y sobre los aspectos éticos de su reutilización con su correspondiente reconocimiento y citación.

AGRADECIMIENTOS

- Al Proyecto Research Data Management Strategy in the Latin-Americans Context (#RDMSLatAm) Financiado por VLIR-UOS.

REFERENCIAS BIBLIOGRÁFICAS

- Abadal, E. (2021). Ciencia abierta: un modelo con piezas por encajar. *Arbor*, 197(799), a588. <https://doi.org/10.3989/arbor.2021.799003>
- Anglada, L.I. & Abadal, E. (2018). Qué es la Ciencia Abierta. *Anuario ThinkEPI*, 12, 292-298. <https://doi.org/10.3145/thinkepi.2018.43>

Cook-Deegan, R., & McGuire, A. (2017). Moving beyond Bermuda: sharing data to build a medical information common. *Genome research*, 27(6): 897–901. <https://doi.org/10.1101/gr.216911.116>

European Commission. (2018). *Commission Recommendation of 25.4.2018 on access to and preservation of scientific information*. European Commission. <https://ec.europa.eu/digital-single-market/en/news/recommendation-access-and-reservation-scientific-information>

European Commission Decision C. (2017). Horizon 2020: Work Programme 2016 – 2017: 16. Science with and for Society. 2468 of 24 April 2017. Brussels, 2017. http://ec.europa.eu/research/participants/data/ref/h2020/wp/2016_2017/main/h2020-wp1617-swfs_en.pdf

Foster. (2019). *The future of science is open*. <https://www.fosteropenscience.eu/>

Fry, J., Schroeder, R., & Den Besten, M. (2009). Open science in e-science: contingency or policy. *Journal of Documentation*, 65(1): 6-32. <https://doi.org/10.1108/00220410910926103>

Junta de Castilla y León. (2012). Qué son los datos abiertos. <http://datosabiertos.jcyl.es/web/es/iniciativa-datos-abiertos>

Kuhn, Thomas S. (1962). *The structure of scientific revolutions*. University of Chicago Press.

Marín-Arraiza, P., Puerta-Díaz, M., & Vidotti, S. G. (2019). Gestión de datos de investigación y bibliotecas: preservando los nuevos bienes científicos. *Hipertext.net*, (19), 13-31. <https://doi.org/10.31009/hipertext.net.2019.i19.02>

Meneses Placeres, G; Álvarez Reinaldo, L.A., Machado Rivero, M. O. (2022). Revisión de las prácticas de Ciencia Abierta en América Latina y el Caribe. *Revista Cubana de Transformación Digital*. 3(1). <https://rctd.uic.cu/rctd/article/view/159>

Ortiz-Núñez, R. (2021). Acceso abierto a la producción científica: ciencia y saber compartidos. *Boletín OCTI*, 12.