

**EL ECOSISTEMA DIGITAL DE
CIENCIA ABIERTA DE
AMÉRICA LATINA Y EL CARIBE**
NOVIEMBRE 2021

El Ecosistema Digital de Ciencia Abierta de América Latina y el Caribe

Agosto 2022

Autores:

Bianca Amaro, LA Referencia

Carlos Barrios, SCALAC

Luis Eliécer Cadenas, RedCLARA

Federico Cetrangolo, LA Referencia

Leandro Ciuffo, RNP

Rafael Mayo-García, CIEMAT

Lautaro Matas, LA Referencia

Phillipe Navaux, SCALAC

Luis Núñez, UIS

Washington L. Ribeiro de Carvalho Segundo, IBICT

Edición y diseño: María José López Pourailly, RedCLARA.

Santiago de Chile

26-08-2022

Copyright de la presente edición:

“El Ecosistema Digital de Ciencia Abierta de América Latina y el Caribe”,

elaborado por LA Referencia, RedCLARA y SCALAC,

se encuentra bajo una Licencia Creative Commons

Atribución-NoComercial-SinDerivadas 3.0



Índice

1.	Introducción	3
2.	El ecosistema digital de la ciencia	4
2.1.	La Cooperación Latino Americana de Redes Avanzadas.....	5
2.2	La red de repositorios de acceso abierto de Latinoamérica (LA Referencia).....	7
2.3	Red Ecuatoriana de Investigación e Innovación - REDI.....	9
2.4	Las Revistas Científicas	9
2.4.1	Plataforma SciELO	10
2.4.2	Plataforma Redalyc	11
2.4.3	Latindex.....	12
2.5	La infraestructura de supercómputo de Latinoamérica y el Caribe (SCALAC).....	13
2.6	Las plataformas de perfiles de investigadores (CV)	16
2.7	Los sistemas de gestión de la investigación (CRIS)	16
2.7.1	El proyecto BrCris	17
2.7.2	El proyecto PerúCRIS	18
2.8	Servicios y plataformas para análisis y experimentación.....	19
3.	Algunas conclusiones preliminares	21
3.1	El Impacto de la transformación digital	21
3.2	La cooperación internacional en Ciencia Abierta.....	22
3.3	La infraestructura de datos, un gran ausente	22
4.	Una hoja de ruta posible con algunas recomendaciones	23
4.1	Ampliar la conectividad de las organizaciones que hacen parte del ecosistema Latinoamericano de Ciencia Abierta	23
4.2	Avanzar integrando los aprendizajes de proyectos exitosos que han conseguido construir capacidades importante en la región.....	24
4.3	Aprovechar los servicios desarrollados por los grandes experimentos científicos de la región ...	24
4.4	Ampliación y fortalecimiento de los servicios de LA Referencia	24
4.5	Selección e implementación de casos de uso	24
4.6	Recursos Humanos.....	25
4.7	Apoyo de la banca multilateral de desarrollo	25
4.8	Profundizar el trabajo cooperativo con Europa y sus ecosistemas digitales.....	25
	Referencias.....	26
	Agradecimientos	26

1. Introducción

La Ciencia Abierta es el movimiento para hacer que la investigación científica (incluidas las publicaciones, los datos, las muestras físicas y el software) y su difusión sean accesibles a todos los niveles de una sociedad, sean investigadores, aficionados o profesionales [1].

La Ciencia Abierta, como modelo de funcionamiento de la actividad científica, tiene el potencial de promover un mayor desarrollo económico y social en Latinoamérica, favoreciendo el uso eficiente y seguro de los recursos que se dedican a financiar la ciencia y facilitando los procesos de transferencia y divulgación del conocimiento en una región llena de retos e inequidades. La incorporación de prácticas científicas transparentes, colaborativas y abiertas, también mejora la calidad de los resultados, su impacto y su reproducibilidad.

La capacidad de traducir los avances científicos en innovaciones se ve fortalecida a través de esta transparencia y accesibilidad y permite la interacción del mundo científico con los sistemas de innovación abierta que necesitan ser fortalecidos para facilitar la creación de riqueza, los aumentos de productividad y la acumulación de capital humano en nuestra sociedad.

El proyecto de recomendación sobre Ciencia Abierta elaborado por la UNESCO destaca cinco pilares que la sustentan: los conocimientos científicos abiertos, las infraestructuras de la Ciencia Abierta, la comunicación científica, la participación abierta de los actores de la sociedad y el diálogo abierto con otros sistemas de conocimiento. [2]

La actividad científica ha alcanzado dimensiones y características que supeditan su productividad a algunos factores claves como el acceso a redes de colaboración global, la gestión efectiva de datos y publicaciones científicas (p.e. la posibilidad de buscar, acceder, interoperar y reutilizar esos datos y las publicaciones vinculadas a ellos), el poder de cómputo, la velocidad de conexión, y la disponibilidad de infraestructuras de investigación compartidas.

Este conjunto de infraestructuras requeridas forman las bases del ecosistema digital necesario para dar sustento e impulso a la ciencia abierta. Latinoamérica cuenta con una larga tradición en este ámbito que puede darnos las bases para desarrollar un liderazgo importante en el despliegue y fortalecimiento de estos ecosistemas, incrementando nuestra capacidad para incorporarnos de forma beneficiosa en colaboraciones ya existentes y para promover el desarrollo de nuevas colaboraciones científicas.

En este documento proponemos una visión de ese ecosistema digital de la Ciencia Abierta que debemos terminar de construir de manera ordenada, analizando sus componentes, interacciones y sugiriendo algunas posibles vías de acción para alcanzarla.

2. El ecosistema digital de la ciencia

La actividad científica se ha digitalizado al igual que casi todas las actividades humanas. Los investigadores requieren, para su labor, de un conjunto de servicios básicos que les permitan:

- Recolectar, curar, almacenar, publicar, mantener y asegurar los datos producidos durante el desarrollo de su investigación.
- Acceder a recursos de cómputo de todo tipo, incluyendo capacidades de cómputo de altas prestaciones tradicionales y no tradicionales, arquitecturas computacionales heterogéneas, ciber infraestructuras de gran escala o cuánticas.
- Publicar y compartir los artículos o trabajos desarrollados y vincularlos de forma permanente con los datos y los algoritmos usados para procesarlos y que son parte integral de su reproducibilidad (para ello son necesarias, también, infraestructuras que integren capacidades masivas de análisis de datos, simulación e inteligencia artificial).
- Contar con herramientas de colaboración flexibles que les permitan integrar su trabajo con otros investigadores y dejen un registro apropiado del proceso a través del cual se extrajeron sus conclusiones.
- Utilizar plataformas y ambientes de desarrollo colaborativos que permitan el uso, evaluación y contribución de recursos abiertos, de manera trazada, federada y segura, que permitan la construcción y difusión del conocimiento por parte de diferentes comunidades.

La necesidad de infraestructuras y organizaciones que den soporte a estas capacidades es evidente a nivel mundial y los países que lideran en productividad cuentan con ellas en abundancia. En el caso europeo, por ejemplo, ese ecosistema está constituido por una amplia multiplicidad de organizaciones que coordinan, desarrollan, administran y hacen accesibles esas infraestructuras a la comunidad científica de Europa. Redes de altas capacidades y conectividad, centros de computación avanzada (proveyendo recursos de supercomputación o cuánticos), centros de datos masivos, centros de excelencia, *hubs* de innovación digital, grandes infraestructuras compartidas de investigación y una nube europea para la Ciencia Abierta, constituyen el ecosistema que da sustento a su actividad científica. Lo mismo, con pequeñas variantes, ocurre entre otros en Canadá, Estados Unidos de América, Japón y China. En esos casos existe un claro alineamiento entre las instituciones que financian la ciencia, la comunidad científica y el conjunto de instituciones que le dan soporte para promover y desarrollar estas infraestructuras.

En Latinoamérica, a lo largo de varias décadas, se han creado organizaciones que han aportado a la construcción de ese ecosistema digital, aún incompleto. Muchas de estas iniciativas surgieron de iniciativas científicas colaborativas en áreas como la Física de Altas Energías, la Astrofísica o las ciencias del medio ambiente. Estas organizaciones, en la actualidad, están en condiciones de ofrecer el pilar de infraestructura antes mencionado e integrarlo a nivel nacional, en cada uno de los países, facilitando la interconexión regional de los sistemas nacionales de ciencia, tecnología e innovación, y su vinculación con otros ecosistemas globales. Entre las organizaciones referidas, se encuentran:

- Las redes nacionales de investigación y educación.
- CLARA, la Cooperación Latino Americana de Redes Avanzadas.
- LA Referencia, la red de repositorios de acceso abierto a la ciencia.
- SCALAC, Sistema de Computación Avanzada para América Latina y el Caribe.
- Infraestructuras de investigación de grandes experimentos internacionales generadores de datos, tales como European Southern Observatory (ESO, <https://www.eso.org/>), Pierre Auger Observatory (PAO, <https://www.auger.org/>) o Latin American Giant Observatory (LAGO, <http://lagoproject.net>).

En este ecosistema faltan esfuerzos y capacidades regionales para contar con una infraestructura integrada de datos abiertos que pueda ser usada para completar las herramientas requeridas por los investigadores para su labor. En este documento proponemos una arquitectura de referencia de ese ecosistema de infraestructuras y organizaciones, y planteamos algunas alternativas para construir las capacidades faltantes.

2.1. La Cooperación Latino Americana de Redes Avanzadas

CLARA nace en el año 2003, en Valle de Bravo, México. Allí se reunieron 13 redes nacionales de investigación y educación de Latinoamérica, para constituir una organización que ayudara a crear la infraestructura de interconexión entre ellas y con el mundo. Interconexión que para ese momento ocurría por defecto en los puntos de acceso a red (NAP, sigla del inglés: *Network Access Points*) ubicados en Florida. Las redes nacionales de investigación y educación, para esa fecha, ya tenían alrededor de una década de existencia, a lo menos en los países que iniciaron estos esfuerzos.

Estas redes nacionales nacieron con el objetivo de proveer servicios de interconexión y acceso a internet a universidades y centros de investigación de cada país. Tras más de un lustro de trabajo cooperativo en el marco del Foro Latinoamericanos de Redes, la creación de CLARA fue un paso natural a la formalización de una cooperación internacional que, en la práctica, ya venía operando muchos años antes.

A partir de su creación, y gracias al apoyo sostenido de la Comisión Europea, CLARA logró desplegar RedCLARA, una infraestructura única en su naturaleza y características, que cuenta con grandes capacidades de conectividad, eslabón fundamental para la integración de un ecosistema para la Ciencia Abierta. Los montos de inversión han permitido construir un bien público regional con activos e inversiones que superan los cien millones de dólares.

Durante los últimos cuatro años, el proyecto BELLA (*Building the Europe Link to Latin America*), cofinanciado por cuatro redes nacionales Latinoamericanas y por la Comisión Europea, desplegó una conexión de alta velocidad directa entre Latinoamérica y Europa, a través de un nuevo cable submarino que conecta Sines, en Portugal, con Fortaleza, en Brasil. Este proyecto, en su primera fase, desplegó varios enlaces terrestres de alta capacidad, conectando a Brasil, Argentina, Chile, Panamá y Ecuador. Durante el año 2020 se estableció una conexión directa y de alta capacidad con África, creando oportunidades de cooperación con las redes africanas. En la figura 1 se muestra la topología actual de la RedCLARA.



Fig. 1: Infraestructura de RedCLARA.

Como se observa en la Fig. 1 hay varios países de Latinoamérica que o bien no están conectados o tienen conexiones de baja capacidad. RedCLARA aspira a completar la conectividad de estos países de forma que tengan igualdad de oportunidades para integrarse al ecosistema regional y global de ciencia, tecnología e innovación [3]. Estos planes de conectividad se complementan y fortalecen con otros proyectos Latinoamericanos como el cable Humboldt, que busca conectar Chile con Oceanía y Asia, y el cable a la Antártica, que permitiría conectar ese continente a través de la Patagonia (Fig. 3).



Fig. 2: Topología BELLÁ 2030.

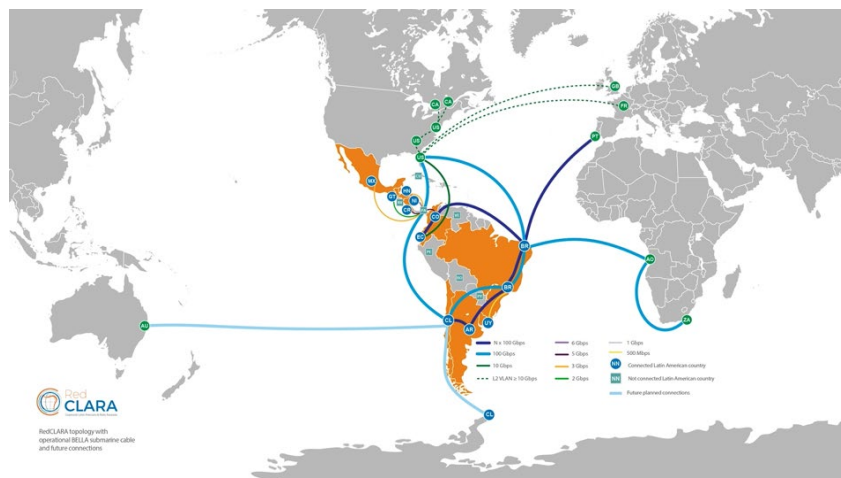


Fig. 3: Cable Humboldt y cable a la Antártica.

2.2 La red de repositorios de acceso abierto de Latinoamérica (LA Referencia)

La Red Federada de Repositorios Institucionales de Publicaciones Científicas, LA Referencia, es la red latinoamericana de repositorios de Ciencia Abierta creada y sostenida por los gobiernos de la región. Por medio de acuerdos, apoya las estrategias nacionales de Ciencia Abierta en América Latina, implementando, desde sus inicios, una plataforma de servicios tecnológicos con estándares de interoperabilidad, que han permitido compartir y visibilizar la producción científica generada en las instituciones de educación superior y de investigación científica. En este contexto, se entiende como producción científica no solo a la literatura, sino también a los datos y metadatos de los insumos, productos y procesos del trabajo de investigación científica.

LA Referencia nace del acuerdo de cooperación -firmado en Buenos Aires en el año 2012- que refleja la decisión política de ofrecer en acceso abierto la producción científica de América Latina como un bien público regional, con énfasis en los resultados financiados con fondos públicos.

La iniciativa opera bajo el principio de una red federada regional donde cada país participante es un nodo nacional que decide cuáles metadatos de su producción científica aporta. El nodo central LA Referencia actúa como agregador regional, construyendo servicios de valor agregado sobre la colección, e interoperando con otras redes regionales y globales, direccionando los accesos a las fuentes que originan el contenido: los repositorios.

A partir de los nodos nacionales, se integran metadatos de artículos científicos, tesis doctorales y de maestría y otros documentos académicos y científicos, provenientes de universidades e instituciones de investigación de los países que hoy conforman LA Referencia: Argentina, Brasil, Chile, Colombia, Costa Rica, Ecuador, El Salvador, España, México, Panamá, Perú y Uruguay.

La red está gobernada por un consejo directivo, conformado por los representantes de organismos gubernamentales de ciencia y tecnología de los países integrantes, junto a RedCLARA. Es un ámbito de discusión internacional con alta legitimidad política. Los integrantes del consejo de LA Referencia son

elegidos por los gobiernos de la región, teniendo, por lo general, el papel de gerentes y tomadores de decisiones en sus países. En ese ámbito se discuten y definen acuerdos en el área de acceso abierto y Ciencia Abierta, promoviendo normas comunes y desarrollando tecnologías de servicios transferibles e interoperables entre sus socios. Como resultado, se trabaja en forma coordinada entre organismos públicos de ciencia y tecnología (Ministerios y ONCyTs) de los países miembros, junto a RedCLARA que participa activamente aportando infraestructura tecnológica, gestión de recursos y experiencia en la gestión de bienes públicos regionales.

LA Referencia ha desarrollado y consolidado una plataforma propia de software LRHarvester, resultado de más de ocho años de desarrollo, es una solución integral que se transfiere a cada nodo de la red. No sólo facilita la agregación, validación y enriquecimiento de metadatos, sino que ha avanzado para ofrecer servicios de valor agregado, responder a los desafíos de interoperabilidad con otras iniciativas globales tales como OpenAIRE. Su nueva versión apunta a atender la creciente necesidad de integración con los sistemas de gestión de la investigación (CRIS) y la posibilidad de elaborar estadísticas que sirvan a los niveles de toma de decisión.

En sus casi 10 años de vida, LA Referencia (ver Fig. 4), nacida como un proyecto regional, se ha consolidado a nivel internacional como el ámbito y referente natural para la discusión sobre políticas nacionales y regionales en temas de Ciencia Abierta, siendo además citado frecuentemente como un caso exitoso de red colaborativa sostenida en el tiempo a pesar de las diferencias y cambios políticos.



Fig. 4: Miembros de LA Referencia.

2.3 Red Ecuatoriana de Investigación e Innovación - REDI

El proyecto Red Ecuatoriana de Investigación e Innovación - REDI, nace en el 2015 como un repositorio de publicaciones científicas y perfiles de investigadores ecuatorianos. En el 2021 se evidenció la necesidad de extender su alcance para apoyar al desarrollo del ecosistema de investigación e innovación del país. Es así que CEDIA desarrolla la plataforma utilizando una arquitectura basada en sistemas CRIS (figura 1), que permite visibilizar globalmente la información de publicaciones científicas, perfiles de investigadores, patentes, grupos de colaboración, tendencia de temáticas de investigación, eventos científicos y servicios tecnológicos ofertados por instituciones ecuatorianas.



Fig. 5: Red Ecuatoriana de Investigación e Innovación.

REDI dispone de un registro de 170349 publicaciones científicas, de 38918 autores dentro de 50 instituciones de educación superior, se cosechan desde repositorios DSpace, revistas OJS y bases de datos bibliográficas como SCIELO, Springer, DBLP, etc.

La plataforma cumple con los principios FAIR (encontrabilidad, accesibilidad, interoperabilidad y reutilización) utilizando tecnologías de web semántica facilitando de manera estandarizada el acceso, visualización y reutilización de metadatos registrado en la plataforma.

Como siguiente paso para el proyecto REDI se apoyará a las instituciones de educación superior a adoptar mayoritariamente los sistemas CRIS, que soporten la investigación institucional y a su vez provean información a REDI, así la generación de estadísticas e indicadores de investigación apoyen a la toma de decisiones a diferentes niveles institucionales, y optimicen la articulación de los actores en el ecosistema de investigación.

2.4 Las Revistas Científicas

Las revistas científicas son el medio de publicación más importante para la comunicación científica, por su capacidad para integrar la rapidez en la publicación de artículos con la evaluación de estos artículos

por parte de especialistas en la materia. Ellas permiten vincular autores y contenidos, ofreciendo una garantía de la científicidad de lo publicado en el menor tiempo posible, al considerar el trabajo generado a través de la revisión por pares. En el contexto latinoamericano, se caracterizan por una alta adhesión al movimiento de Acceso Abierto a la información científica, y también por una práctica de Acceso Abierto de vanguardia que es vista como un modelo de gran influencia dentro del mismo.

El movimiento de Acceso Abierto a la información científica se fortaleció, a partir de finales de la década de 1980, como reacción a los precios exorbitantes cobrados por los oligopolios de las editoriales científicas comerciales. Los precios de las revistas científicas se convirtieron en un impedimento incluso para las bibliotecas de grandes instituciones, restringiendo el acceso a las publicaciones científicas a los propios productores científicos. En este escenario surge la Vía Dorada del Acceso Abierto que promueve la creación de revistas científicas que publican artículos sin restricciones de acceso y uso.

Bajo la influencia de este movimiento global, y ante la imposibilidad de pagar por revistas científicas comerciales, se pensaron muchas acciones prácticas para resolver el problema. En los países del norte global, con el objetivo de abolir las restricciones de acceso y uso, se difundieron revistas de Acceso Abierto que cobran honorarios a los investigadores por procesar o enviar artículos. Esta práctica elimina las restricciones de acceso y uso de los artículos, pero restringe la publicación de artículos a investigadores sin financiación para la investigación. También crea un problema al abrir la posibilidad para la creación de revistas depredadoras, que aprovechan la perspectiva del Acceso Abierto para obtener ganancias económicas fáciles al cobrar por la publicación de artículos que no han pasado por los procesos formales de evaluación científica.

En América Latina, la práctica del Acceso Abierto se construyó, más comúnmente, a través de financiamiento institucional (principalmente público) para la producción y mantenimiento de estas publicaciones. Este modelo elimina las restricciones de acceso y uso de los artículos, así como el costo de publicación que recae sobre el investigador, ya que las publicaciones son mantenidas por las propias instituciones de educación superior e investigación. A pesar de las ventajas, el modelo adolece de algunos contratiempos que se relacionan con la falta de incentivos por parte de los agentes nacionales de financiamiento de la ciencia, que priorizan los modelos tradicionales restrictivos en detrimento de la Ciencia Abierta. Esta falta de financiación acaba dificultando la adaptación de las revistas, lo que genera revistas con un equipo reducido, sin el personal necesario para su sostenibilidad y protagonismo en la comunidad científica. En general, las revistas latinoamericanas, a pesar de cumplir comúnmente con los procesos formales de evaluación científica, se ven afectadas por una baja calidad editorial, lo que impide la indexación de un gran número de revistas en indexadores y plataformas con alcance internacional y limita el impacto de la ciencia latinoamericana en el mundo, socavando sus calificaciones internacionales.

Con el fin de solucionar estos problemas, surgieron iniciativas como SciELO y Redalyc que buscan una mayor adecuación y sustentabilidad del modelo latinoamericano de Acceso Abierto y la expansión de sus capacidades para promover la internacionalización y el consecuente incremento en el impacto de la ciencia latinoamericana.

2.4.1 Plataforma SciELO

El modelo SciELO fue desarrollado en 1997 a través de un proyecto de colaboración entre la Fundación de Apoyo a la Investigación del Estado de São Paulo (FAPESP) y el Centro Latinoamericano y del Caribe de Información en Ciencias de la Salud (BIREME/OPS/OMS). El proyecto fue desarrollado

en el marco de un programa especial de FAPESP, con el apoyo de CAPES y CNPq. Lanzado en Brasil en 1998, fue adoptado en el mismo año por CONICYT Chile, iniciando la Red SciELO de colecciones nacionales de revistas que incluía, desde sus inicios, una importante colección en salud pública. En el año 2021, a veintitrés años de su creación, SciELO está presente en 17 países (Sudáfrica, Argentina, Bolivia, Brasil, Chile, Colombia, Costa Rica, Cuba, Ecuador, España, Antillas, México, Paraguay, Perú, Portugal, Uruguay y Venezuela). Las colecciones de revistas son administradas, financiadas y desarrolladas a nivel nacional, en general, con el apoyo de una organización de ciencia y tecnología (consejo de ciencia, ministerio de ciencias o similares). Las colecciones son coordinadas por una institución reconocida y de alto prestigio y desarrolladas con el apoyo de un comité científico representativo de la comunidad de publicaciones periódicas del país

Los servicios ofrecidos por SciELO incluyen la indexación de revistas y otros objetos de comunicación con evaluación del desempeño bibliométrico, almacenamiento y preservación digital, publicación de acceso abierto, difusión e interoperabilidad. El objetivo permanente es promover la calidad de los artículos y otros documentos operados en colecciones de revistas accesibles a través del portal <https://www.scielo.org/>. SciELO gestiona preimpresos a través del servidor SciELO Preprints - <https://preprints.scielo.org/> y en el futuro a través de SciELO PostPrint, archivos de datos, y otros materiales a través del repositorio de datos SciELO - <https://data.scielo.org/> , y SciELO Books - <https://books.scielo.org/>.

Al mismo tiempo, las revistas SciELO están ajustando sus políticas editoriales en favor de prácticas de comunicación de Ciencia Abierta. Los servicios ofrecidos por la plataforma de SciELO siguen estándares nacionales e internacionales que son implementados por políticas, procedimientos y tecnologías. Entre los principales estándares adoptados se encuentran el uso del sistema de licencias de acceso abierto Creative Commons, SciELO Publishing Schema para el marcado de documentos en XML que sigue el estándar NISO Journal Article Tag Suite (JATS), los códigos ISO de idiomas y países, los identificadores comunes DOI (Digital Object Identifier) para objetos de comunicación de investigación, ORCID para investigadores, CRediT (Contributor Roles Taxonomy) para identificar la contribución de autores de artículos colaborativos, sistema LOCKSS de preservación digital a través de la Red Carriana y el código de práctica COUNTER para el conteo de accesos. La aplicación del Modelo Editorial SciELO está documentada por una serie de manuales y guías en español, inglés y portugués ampliamente accesibles desde su página web.

Actualmente la Red SciELO está gestionando alrededor de 1.300 revistas activas, que publican más de 50.000 documentos anualmente. La Red SciELO acumula un total de más de 950.000 documentos, que atienden un promedio diario de más de 1,5 millones de descargas medidas por el protocolo COUNTER. Todos ellos están indexados por Google Scholar, LA Referencia, SciELO Citation Index en la plataforma Web of Science y catalogados en Latindex. En 2020 se publicaron 50.470 trabajos de investigación, de los cuales el 70% está indexado por Scopus y el 50% en WoS, que son los índices bibliométricos más utilizados por los sistemas nacionales de evaluación. En el índice Scimago, en 2020, el 30% de las revistas y artículos de la Red SciELO se encuentran en el primer y segundo cuartil del indicador SCImago Journal Rank (SJR).

2.4.2 Plataforma Redalyc

Redalyc es una infraestructura que nace desde el modelo de ciencia como bien común con un enfoque de comunicación científica centrado en la región de América Latina, con el objetivo de ofrecer

visibilidad, interoperabilidad y servicios para la producción digital y marcación XML-JATS de las revistas científicas.

Actualmente, Redalyc indexa y provee servicios no solamente a revistas de América Latina sino de cualquier región del mundo, siempre que éstas cumplan con el requisito de ser revistas revisadas por pares de acceso abierto no comercial (sin APC, acceso abierto diamante), y que aprueben el proceso de evaluación de calidad científica y editorial. Actualmente Redalyc indexa más de 1.400 revistas científicas de 31 países, con más de 600 instituciones, y una colección que cuenta con más de 800.000 artículos a texto completo.

La visión de Redalyc contempla un ecosistema de circulación libre de conocimiento manejado por la comunidad académica y fuera del mercado comercial, a través de la cooperación intrínseca entre la vía verde y la vía diamante del acceso abierto. Redalyc ha buscado contribuir a regresar a las instituciones el conocimiento que ha sido producido por sus investigadores.

Redalyc ha desarrollado procesos de normalización de la información correspondiente a los datos de afiliación de los autores de los artículos, lo cual permite agrupar y recuperar la producción científica publicada en las revistas por criterios de institución o país del autor. Sobre ello Redalyc ha desarrollado un servicio para contribuir al fortalecimiento de los repositorios institucionales y las redes nacionales de repositorios, a través de la creación de un punto de acceso OAI-PMH para cada una de las instituciones con producción científica publicada en las revistas indexadas por Redalyc.

Es así como Redalyc ofrece información de más de 10.000 instituciones, disponible para ser cosechada por sus repositorios. Esto permite a las instituciones conocer y recuperar la producción científica de sus investigadores publicadas en revistas de Acceso Abierto para poblar los repositorios institucionales o nacionales. Los puntos OAI-PMH ofrecen, por ejemplo, más de 92.000 artículos científicos de 2.743 instituciones de Colombia, 274.000 de 7,254 instituciones de Brasil, 122.000 de 2.788 instituciones mexicanas o 45.000 artículos de 1.516 instituciones de Argentina, por mencionar algunos.

2.4.3 Latindex

Latindex es producto de la cooperación de una red de instituciones que funcionan de manera coordinada para reunir y diseminar información sobre las publicaciones científicas seriadas producidas en Iberoamérica. La idea de creación de Latindex surgió en 1995 en la Universidad Nacional Autónoma de México (UNAM) y se convirtió en una red de cooperación regional a partir de 1997.

Latindex cuenta con dos productos de información: Directorio, ofrece datos bibliográficos y de contacto de todas las revistas registradas y Catálogo 2.0, compuesto por las revistas que cumplen con los más altos estándares de calidad según la metodología de Latindex. Se incluyen revistas de investigación científica, técnico-profesionales y de divulgación científica y cultural que se editan en América Latina, el Caribe, España y Portugal. Además, ofrece información sobre revistas con contenidos iberoamericanistas editadas en cualquier parte del mundo. Las revistas pueden ser impresas o en línea, de todas las disciplinas científicas.

2.4.4 Los repositorios de preimpresos dedicados a revistas científicas

2.4.4.1 Emerging Research Information (EmeRI)

El repositorio de preimpresos Emerging Research Information (EmeRI) se implementó en 2020 a través de una cooperación entre la Asociación Brasileña de Editores Científicos (ABEC) y el Instituto Brasileño de Información en Ciencia y Tecnología (IBICT), con el objetivo de proporcionar servicios a revistas y editores que les ayuden a acelerar la difusión de los resultados de investigaciones científicas emergentes. EmeRI basa estos servicios en la disponibilidad de archivos de preimpresos.

La propuesta de EmeRI surgió de demandas de algunos editores científicos brasileños que vieron la necesidad de acelerar la disponibilidad de artículos enviados a sus revistas, especialmente frente a la pandemia de Coronavirus.

EmeRI promueve el acceso a la información científica libre de barreras y apoya la apertura y velocidad del proceso científico.

2.4.4.2 SciELO Preimpresos

La Colección SciELO Preprints es una parte integral de SciELO. Es una colección de manuscritos, no revisados por pares, dentro de la Red SciELO de colecciones nacionales y temáticas de revistas. Como tal, los *preprints* se benefician de los mismos servicios que los artículos de revistas, incluida la indexación y la preservación. Se ejecuta sobre el *software* Open Preprint Systems (OPS) que es un servidor en línea gratuito y sin fines de lucro de archivos y distribución de preimpresos desarrollado y mantenido por el Public Knowledge Project (PKP). El servidor SciELO Preprints tiene una interfaz directa con las revistas SciELO para el envío de *preprints*, así como con cualquier sistema de envío que ejecute el protocolo SWORD.

2.5 La infraestructura de supercómputo de Latinoamérica y el Caribe (SCALAC)

El Sistema de Cómputo Avanzado para América Latina y el Caribe (SCALAC, <https://scalac.redclara.net>), es una iniciativa que surge directamente de las comunidades de usuarios de RedCLARA, que, habiendo participado durante más de una década en diferentes proyectos de infraestructuras para el cómputo distribuido tipo *Grid*, reúnen toda esa experiencia para formalizar esta colaboración. Proyectos como EELA, EELA-2 y GISELA permitieron no solo ofrecer infraestructura sino también formar recursos humanos a diferentes niveles y hacia diferentes roles: desde usuarios científicos hasta administradores de arquitecturas computacionales, desarrolladores de aplicaciones e investigadores de ciencias e ingenierías computacionales e informática.

Progresivamente, sistemas nacionales como el Sistema Nacional de Procesamiento de Alto Desempeño, SINAPAD, en Brasil, así como otros centros nacionales y universitarios, se fueron integrando como una comunidad abierta de computación de alto rendimiento y distribuida. También se incorporaron proyectos de computación como es el caso de la Red Iberoamericana de Supercomputación (RISC), la Red Iberoamericana de Computación de Altas Prestaciones (RICAP) o la colaboración Francia-Brasil, que permitió la integración de Brasil al ecosistema de computación avanzada Grid5000 francés, y a PlanetLab.

SCALAC tiene como principal objetivo organizar e integrar en la región, a través de la conectividad provista por RedCLARA y las redes nacionales de investigación y educación, las capacidades de cómputo avanzado y las competencias en conocimiento regionales para proveer servicios dirigidos a intereses comunes y sin ánimo de lucro, que requieran computación de alto rendimiento. Es importante mencionar que se hace mucho énfasis en la formación, divulgación de resultados y la colaboración entre países tanto de la región como con el resto del mundo. SCALAC fue lanzado oficialmente en Colombia en el 2013, durante la finalización del programa de comunidad Grid y Computación Científica de RedCLARA y aunque estratégicamente está muy enlazado con RedCLARA (no solo en la interconectividad sino en la actividad), es una entidad civil independiente, sin ánimo de lucro, legalmente constituida en Costa Rica.

Actualmente, los centros y sistemas nacionales que hacen parte de SCALAC proveen capacidades de cómputo para los investigadores de la región cuyas instituciones hacen parte de las redes nacionales de investigación y educación (ver Cuadro 1).

Otros centros e iniciativas nacionales se han integrado al sistema gracias al apoyo de las redes nacionales de investigación y educación (como es el caso de la Red Colombiana de Computación Avanzada, con el soporte y apoyo de la red nacional de investigación y educación de Colombia, RENATA). SCALAC, igualmente, integra colaboraciones internacionales como RISC2 (A network for supporting the coordination of High-Performance Computing research between Europe and Latin America), liderada por el Centro Nacional de Supercomputación de España (BSC-CNS), la colaboración continental de supercomputación de las Américas (Americas HPC Collaboration), en la cual participan actores del Departamento de Energía de los Estados Unidos de América, y el sistema de computación avanzada de Canadá (ComputeCanada), entre otras.

Cuadro 1: Centros de supercómputo de SCALAC interconectados a través de RedCLARA

Centro de Cómputo	Capacidad Total de Cómputo	País	Página Web
Centro de simulación computacional para aplicaciones tecnológicas	48 TFlops	Argentina	https://csc.conicet.gov.ar/
Laboratorio Nacional de Computación Científica (incluye el Núcleo de Cómputo Avanzado Santos Dumont)	5.4 PFlops	Brasil	https://www.gov.br/mcti/pt-br/rede-mcti/lnce y https://sdumont.lnce.br
Sistema Nacional de Computación de Alto Desempeño (incluye las plataformas del sistema federal de universidades y algunas otras instituciones)	1.3 PFlops	Brasil	https://www.lnce.br/sinapad/
Laboratorio Nacional para computación de alto desempeño	266	Chile	https://www.nlhpc.cl/
Supercomputación y Cálculo Científico- Universidad Industrial de Santander	105	Colombia	https://www.sc3.uis.edu.co/
Universidad de los Andes	48	Colombia	https://www.uniandes.edu.co
Colaboratorio nacional de computación avanzada	85	Costa Rica	https://kabre.cenat.ac.cr/
Abacus: Laboratorio de matemática aplicada y cómputo de alto rendimiento	429	México	https://www.abacus.cinvestav.mx
Red Mexicana de Supercomputación (RedMex via CUDI)	200	México	http://www.redmexsu.mx/
Universidad de la República	98	Uruguay	https://www.udr.uy
Servicio Ecuatoriano de Supercomputación (Via CEDIA)	230	Ecuador	https://cedia.edu.ec/

Fuente: SCALAC.

2.6 Las plataformas de perfiles de investigadores (CV)

2.6.1 La plataforma Lattes

Latinoamérica, a través de esta iniciativa originada en Brasil, ha sido pionera en este campo, creando, en 1999, la plataforma Lattes (<https://lattes.cnpq.br/>) de perfiles de investigadores que actualmente (2021) presenta más de 7,2 millones de perfiles, siendo 377 mil doctores (conteniendo 7 mil investigadores declarados con libre docencia) y 653 mil exclusivamente a nivel de maestría. Este sistema de información es desarrollado y mantenido por el Consejo Nacional de Desarrollo Científico y Tecnológico (CNPq).

El contenido descrito por la plataforma Lattes es muy rico y se convirtió en objeto de estudios con el apoyo de diversas herramientas de análisis, como el ScriptLattes y el LattesDataExplorer. Dichas herramientas realizan “recuperación de datos” y procesamiento de los CVs albergados en la plataforma. En la plataforma hay más de 4,5 millones de artículos de revistas; 10,3 millones de artículos de conferencias; 1,4 millón de tesis y disertaciones; 1,2 millón de capítulos de libros; 520 mil libros; 750 mil textos en periódicos y magazines; 75 mil patentes y 920 trabajos técnicos.

2.6.2 Plataforma CVAR

Creada en 2011 por el Ministerio de Ciencia, Tecnología e Innovación Productiva (MINCYT) de Argentina, como un relevamiento exhaustivo del sistema científico tecnológico de la nación, la plataforma CVAR hoy reúne más de 60.000 currículum vitae (CV).

CVAR es un registro unificado y normalizado de datos curriculares de los científicos y tecnólogos de todas las instituciones del país, con actualización permanente y en línea. El sitio www.sicytar.mincyt.gob.ar permite realizar búsquedas por temas de investigación, nombre del científico; por áreas, por provincia e instituciones, entre otras variables.

Entre sus funciones, la plataforma provee información estadística para la planificación y la toma de decisiones, sirve como herramienta para convocatorias institucionales, publica datos curriculares, lograr la redacción de un CV estándar aceptado por todas las instituciones y promover el intercambio de información.

2.7 Los sistemas de gestión de la investigación (CRIS)

Los CRIS (Current Research Information Systems) son sistemas que sirven para recoger y difundir toda la información relacionada con las actividades de investigación de una institución y/o país, es decir, cuáles son sus investigadores/as, las publicaciones, las patentes y los conjuntos de datos que han generado, los proyectos de investigación y sus fuentes de financiamiento (De Castro, 2019). Se empezaron a crear en Europa durante la década del '90 y en el año 2002 se fundó euroCRIS9 (European Organisation for International Research Information). Una infraestructura de estilo CRIS en Latinoamérica ayudaría a construir indicadores adecuados para evaluar la producción científica yendo más allá de las publicaciones tradicionales.

En las Recomendaciones de Ciencia Abierta de la UNESCO se plantea la necesidad de avanzar en inversiones de largo plazo vinculadas a la infraestructura digital. Dentro de las prioridades se encuentra el desarrollo de los sistemas de gestión de la investigación CRIS. Para su implementación, la UNESCO recomienda que la gobernanza de dichas plataformas esté a cargo de la comunidad científica mientras que el financiamiento, a cargo de los gobiernos.

2.7.1 El proyecto BrCris

En Brasil IBICT comenzó en 2014 el proceso de estudio para crear un Sistema Nacional de Información de Investigación (CRIS), este estudio culminó en 2020 con el lanzamiento de un proyecto de Investigación para el desarrollo de un sistema CRIS nacional brasileño, llamado BrCris (ver Fig. 5).

Por otro lado, existe una red en Brasil de más de 110 repositorios institucionales de publicaciones científicas, 125 bibliotecas de tesis y disertaciones y 1.000 revistas científicas de acceso abierto. Esta red de fuentes de acceso abierto está reunida en el nodo brasileño de LA Referencia, el Portal oasisbr < <https://oasisbr.ibict.br/> >, que actualmente agrega más de 2 millones y 670 documentos en texto completo. Este portal fue creado y coordinado por el Instituto Brasileño de Información sobre Ciencia y Tecnología (IBICT).

También hay otras bases importantes en el contexto del ecosistema brasileño de investigación científica. Un ejemplo importante es la Plataforma Sucupira < <https://sucupira.capes.gov.br/> >, que fue desarrollada y mantenida por la Coordinación para la Mejora del Personal de Educación Superior (CAPES), y con datos recopilados para la evaluación de la producción técnica y científica generada en el marco de los programas académicos brasileños de maestría y doctorado.

La arquitectura elegida para BrCris fue la de agregar diferentes fuentes, organizadas centralmente en un formato de red de conocimiento. Al igual que Oasisbr, el software elegido para almacenar y organizar el contenido recopilado para BrCris fue la Plataforma de LA Referencia. La plataforma permite exportar los datos a diferentes plataformas de servicios y visualización.



Fig. 6: Proyecto BrCRIS.

2.7.2 El proyecto PerúCRIS

PerúCRIS es el nombre del proyecto que viene realizando el Concytec Perú, el cual busca establecer, desarrollar y operar la Red Nacional de Información en Ciencia, Tecnología e Innovación Tecnológica (CTI), a fin de consolidar y gestionar la información científica y académica de todo el Perú, generar estadísticas para apoyar la toma de decisiones a nivel institucional, regional, sectorial y nacional; así como visibilizar globalmente las actividades, capacidades y producción científica de los investigadores peruanos.

El proyecto #PerúCRIS, va más allá de la construcción de una plataforma informática, está conformado por una serie de medidas y acciones para la operación de la Red Nacional de Información en CTI.

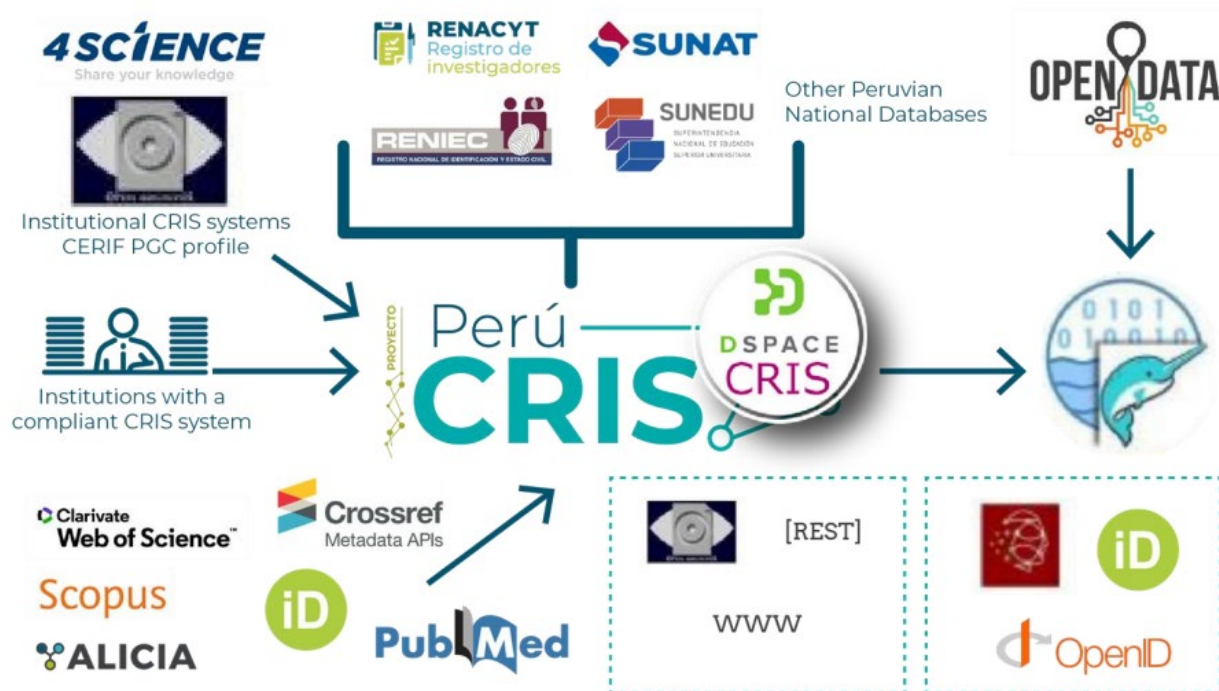


Fig. 7: Proyecto PerúCRIS.

2.8 Servicios y plataformas para análisis y experimentación

Entre los servicios que un ecosistema digital de Ciencia Abierta debe proveer a científicos e investigadores se encuentra el acceso abierto a plataformas, infraestructuras e instrumentos para experimentación. Estas infraestructuras o instrumentos van desde grandes observatorios astronómicos, aceleradores de partículas, pasando por recursos que uno o varios grupos de investigación tienen a su disposición en un laboratorio local, hasta sensores y dispositivos desplegados a los efectos de obtener datos del entorno.

En el caso de las grandes infraestructuras, el acceso abierto y compartido es una necesidad impuesta por el costo asociado a construirlas. Son infraestructuras demasiado costosas para no aprovecharlas colectivamente. En el caso de otros tipos de instrumentos, el compartirlos puede conducir a un uso más eficiente de la inversión. En ambos casos es indispensable el soporte digital apropiado para su uso.

Por otra parte, los procesos de transformación digital están creando otras fuentes de datos que son insumos valiosos para apoyar la investigación en diversas disciplinas científicas. Con la introducción de nuevas tecnologías como el IoT, la inteligencia artificial, el *blockchain* y otros, se está produciendo una transformación profunda en las oportunidades y modos de desarrollar el conocimiento científico.

El IoT, por ejemplo, puede servir para generar datos muy detallados y granulares de gran diversidad de eventos. En un futuro la cantidad de dispositivos que estarán conectados alcanza dimensiones enormes

y requieren de una infraestructura adecuada para poder desplegarla y aprovecharla. El volumen de datos y el modelo de procesamiento requerido ha llevado al desarrollo de nuevos paradigmas computacionales como es la computación en el borde (*edge computing*) que ofrece la posibilidad de alcanzar las escalas de cómputo requeridas para recibir, procesar y actuar en función de los datos colectados, dentro de los tiempos de respuesta requeridos.

Esta capacidad de recolección de datos a través de dispositivos IoT está, a su vez, directamente vinculada al despliegue de servicios 5G y Wifi 6, ambas tecnologías que cambian la arquitectura de conectividad para llevar las capacidades de procesamiento hacia los bordes, acercando la capacidad de cómputo a los dispositivos que se desea manejar o de los cuales se debe coleccionar datos. Así, si pensamos un entorno muy extendido tanto en superficie como en número de sensores, necesitaremos una infraestructura computacional que nos vaya permitiendo realizar distintos análisis en distintas capas, por ejemplo, en un entorno IoT:

- Análisis en el sensor y *edge computing* (ARM, FPGA...) para, por ejemplo, un primer calibrado de datos.
- *Fog computing*, véase clústeres de computación para un primer análisis de los datos previamente filtrados.
- Supercomputadores más grandes para realizar análisis basados en cálculos más precisos y demandantes de cálculo y/o memoria.
- Esta capa ya debería de contar con la infraestructura de almacenamiento adicional que para los casos anteriores pudiera no ser tan necesaria porque los datos se borran tras el proceso y posterior envío a servidores siguientes.
- Implementación de contenedores adaptados para cada tipo de análisis que se levanten sobre entornos virtualizados.

De manera complementaria, podríamos pensar en grandes experimentos que proveen datos de manera continua, tanto experimentales como sintéticos, a partir de simulaciones. Para ellos, se necesitaría, igualmente, de una infraestructura que integre distintas capacidades:

- Herramientas de procesado y limpieza de datos.
- Conexión con infraestructuras paralelas y distribuidas para almacenamiento y simulación.
- Este tipo de plataformas también podrían contar con entornos virtualizados *ad hoc*.
- Categorización de datos en distintos niveles para su posterior explotación y embargo temporal por parte de la colaboración.

Como puede deducirse, fácilmente, tanto la infraestructura IoT como la de los grandes experimentos no son conjuntos disjuntos y ambas pueden requerir de capacidades inicialmente pensadas para la otra por motivos del entorno de generación de datos específicos para cada caso.

Por tanto, la inteligencia artificial, el *big data* y las capacidades de cómputo, están impactando significativamente la productividad del investigador (e incluso de los ciudadanos, quienes cada vez tienen más acceso a datos de cualquier ámbito científico, tecnológico y social). En condiciones ideales un investigador debería contar con conectividad con grandes anchos de banda a infraestructuras y dispositivos, desde los más grandes y escasos, a los más pequeños y abundantes. Estas conexiones combinan tecnologías de comunicación que van desde la fibra óptica, para las conexiones de gran

capacidad, hasta tecnologías inalámbricas como 5G o Wifi 6. Esta conexión debe permitir al investigador el acceso y manipulación de los instrumentos asociados al experimento al igual que la colección, depuración y almacenamiento de los datos producidos en repositorios destinados para tal fin. Las herramientas de cómputo y el acceso a algoritmos y técnicas de procesamiento, incorporando nuevas tecnologías como la inteligencia artificial que facilitan el procesamiento de volúmenes cada vez más grandes de datos. Un ejemplo en funcionamiento de estos entornos son los servicios de lechos de prueba ofrecidos por la Red Nacional de Pesquisa (RNP) en Brasil (ver <https://www.rnp.br/servicios/testbeds>).

3. Algunas conclusiones preliminares

3.1 El Impacto de la transformación digital

Como indica el informe sobre la Economía Digital publicado por UNCTAD en 2021, los datos digitales son la base de las tecnologías digitales emergentes. Se destacan: la analítica de datos, la inteligencia artificial, la tecnología de cadena de bloques (*blockchain*), el internet de las cosas, la computación en la nube y todos los servicios basados en internet.

La relevancia que comienzan a tomar dichas tecnologías, en términos de balanzas comerciales, innovación y desarrollo económico, en su conjunto, hace indispensable que los distintos países discutan sus implicancias normativas. Del mismo modo, existen implicancias en términos de equidad y seguridad nacional.

Adicionalmente, la pandemia del COVID-19 aceleró considerablemente los procesos de transformación digital, en 2020 el ancho de banda global de internet se incrementó en un 35%, y se estima que esta tendencia va a continuar en los próximos años. Dicha aceleración acentúa una serie de desequilibrios en términos de equidad a nivel global. En este sentido, el informe sobre la Economía Digital (UNCTAD, 2020) indica que el 20% de los habitantes de países menos desarrollados son usuarios de internet y adicionalmente se enfrentan a altas dificultades en términos de calidad de servicio a costos altos en términos relativos con los que asumen los usuarios de países desarrollados. Adicionalmente, se observan mayores niveles de exclusión en poblaciones rurales y en mujeres.

Dentro de los países con mayores indicadores de utilización de tecnologías digitales se encuentran los Estados Unidos y China. Entre ambos países poseen el 50% de los centros de datos a nivel global, las mayores tasas de adopción de 5G, el 94% de la financiación de empresas emergentes de Inteligencia Artificial, el 70% de los mejores investigadores en Inteligencia Artificial y el 90% de la capitalización bursátil de las plataformas digitales (UNCTAD, 2020).

En lo que respecta a la región de América Latina y el Caribe, si bien ha aumentado considerablemente el uso de los datos digitales, como en el resto del mundo, enfrenta hoy importantes desafíos para disminuir la brecha digital. Según datos del Banco Mundial, en la región hay aproximadamente 200 millones de personas que carecen de infraestructura digital básica y otras tantas poseen cierta infraestructura básica, pero la calidad de los servicios es insuficiente y los costos son elevados con respecto a otras regiones. Como sucede en el resto del planeta, la pandemia aceleró los procesos de transformación digital pero este proceso avanzó a la par de la profundización de las brechas en términos de igualdad de acceso.

Existen varios motivos por los cuales se profundizan dichas brechas: elevado costo de conectividad en relación a los ingresos de los hogares que se relacionan directamente con falta de competitividad de las empresas proveedoras de servicios de comunicación, marcos regulatorios complejos, etc., falta de infraestructura en el sistema de telecomunicaciones, e ineficiencia en la prestación de servicios.

La inequidad en la obtención de los beneficios que emergen de los procesos de transformación digital hace indispensable el desarrollo de políticas regionales vinculadas a aumentar la infraestructura digital, centros de datos, etc. El rol de los organismos multilaterales de financiamiento puede tener un rol clave, siendo un socio estratégico para que el sector privado pueda asumir el riesgo de las inversiones necesarias para mejorar los servicios asociados a la transformación digital, sobre todo en territorios rurales o de difícil acceso por las complejas condiciones geográficas.

3.2 La cooperación internacional en Ciencia Abierta

A lo largo del documento se han mencionado distintas colaboraciones articuladas a través de distintos proyectos, véase BELLA en la sección 2.1, OpenAIRE en la sección 2.2, EELA, CHAIN o RICAP en la sección 2.4, simplemente por mencionar unos cuantos en los cuales la cooperación con entidades europeas, norteamericanas, africanas o asiáticas han supuesto un gran avance para el desarrollo de la Ciencia Abierta en Latinoamérica.

Es obvio constatar, por tanto, que si se quiere desarrollar un sistema de Ciencia Abierta en la región, es imprescindible la cooperación con entes financiadores e instituciones claves en otras regiones del mundo. Dentro del ámbito de las redes avanzadas (Internet2, Geánt, Ubuntunet Alliance, RedCLARA) siempre ha habido una colaboración fructífera y duradera que ahora ha de ampliarse a iniciativas tales como Research Data Alliance, European Open Science Cloud, EuroHPC, etc., para hacer efectivo un intercambio de conocimiento que culmine en un ecosistema de Ciencia Abierta global.

Por razones históricas y los éxitos conseguidos, la Comisión Europea, a través de sus Programas Marco (6º Programa Marco y Horizonte 2020 en el pasado y Horizonte Europa desde el 2021 en adelante), ha supuesto un motor de impulso en el desarrollo de las infraestructuras avanzadas en Latinoamérica y su papel para el futuro se antoja clave.

Más allá de potenciales convocatorias de investigación e innovación (RIA) o de acciones de coordinación y apoyo (CSA) que puedan financiar los distintos avances a los que se apunta en este documento, la propia coordinación con la Comisión a través de sus Direcciones Generales, puede ser el primer paso relevante para la consecución de los objetivos planteados. Sólo desde la definición de una hoja de ruta ambiciosa, pero a su vez realista, las distintas acciones que se promuevan podrán tener éxito.

3.3 La infraestructura de datos, un gran ausente

En Latinoamérica no contamos con una infraestructura regional para almacenar datos científicos. Cada país y, en algunos casos, cada universidad o centro de investigación, despliega sus propias soluciones. Esto conlleva a un enfoque fragmentado que impide el desarrollo e implementación de los

anteriormente mencionados principios FAIR (sigla en inglés para: *Findable, Accessible, Interoperable, Reusable*): encontrables, accesibles, interoperables y reutilizables.

¿Cuál es la ventaja de contar con una infraestructura que promueva estas propiedades en los datos? Al tener estas propiedades, los datos producidos por las comunidades científicas pasan a constituir un recurso común que puede ser utilizado por los investigadores, eliminando la duplicidad de esfuerzos y el uso ineficiente de los escasos recursos que, en general, se dedican a la ciencia en Latinoamérica.

Con una infraestructura común como la que describimos, facilitamos que las inversiones hechas en la recolección y procesamiento de datos científicos sume regionalmente, ampliando indirectamente el tamaño de las inversiones nacionales al darles acceso al mismo costo a una mayor cantidad de evidencia y recursos.

La propuesta es, por tanto, implementar una arquitectura como la propuesta en distintos ámbitos científico-tecnológicos, identificando aquellos que en una primera instancia estén más preparados para ello para que, de ese modo, sirvan como ejemplo para que posteriormente otras comunidades adopten un sistema similar para su arquitectura particular.

4. Una hoja de ruta posible con algunas recomendaciones

Teniendo en cuenta todos los puntos anteriores, se describe, de manera sucinta, a continuación, una serie de recomendaciones que esperamos contribuyan a la implementación efectiva de un ecosistema digital de Ciencia Abierta de América Latina y el Caribe.

4.1 Ampliar la conectividad de las organizaciones que hacen parte del ecosistema Latinoamericano de Ciencia Abierta

Universidades, centros de investigación, organismos de ciencia y tecnología, y otros actores del ecosistema científico, deben estar bien conectados entre sí y con sus pares. Una buena conexión de datos para soporte a la ciencia requiere de grandes anchos de banda, baja latencia, niveles elevados de ciberseguridad e interconexión con las mismas características con redes de investigación y educación a nivel mundial. La región latinoamericana adolece de grandes brechas digitales que profundizan una brecha de conocimiento con los países líderes en investigación científica. La conectividad apropiada solo está disponible para una parte del ecosistema y deja por fuera un conjunto importante de organizaciones que no tienen posibilidad de conectarse con la capacidad requerida. Aunar esfuerzos entre los gobiernos, las agencias internacionales de cooperación, la banca de desarrollo y otros, para enfatizar la creación de infraestructuras dedicadas a la ciencia, es un paso muy importante. El ecosistema de organizaciones en la región cuenta con activos muy relevantes que deben ser integrados coherentemente en una estrategia del continente en este ámbito. RedCLARA ha propuesto una visión de conectividad para el año 2030. Con apoyo de la Comisión Europea y de las redes nacionales de investigación y educación de Latinoamérica, se aspira a conectar a los países aun no conectados a la infraestructura de BELLA. Es importante

complementar estos esfuerzos y generar sinergias para lograr el máximo de efectividad y eficiencia en el uso de los recursos disponibles para estos proyectos.

4.2 Avanzar integrando los aprendizajes de proyectos exitosos que han conseguido construir capacidades importante en la región

Es necesario integrar los aprendizajes de proyectos que ha sido exitosos en desarrollar capacidades. El ejemplo más reciente son los logros obtenidos a través de proyecto BELLA, que ha desplegado un cable submarino nuevo entre Europa y Latinoamérica donde se ha reservado una capacidad muy importante dedicada para la ciencia y la educación y que representa un medio de integración con los ecosistemas digitales de ciencia y educación europeos. Otros proyectos, como el caso de Amlight, han construido capacidades integrando una conectividad de avanzada con Estados Unidos y Canadá. En un futuro estos proyectos pueden servir de referencia para los planes de desplegar el cable Humboldt, entre Chile y Australia, integrando al continente con los ecosistemas de ciencia y educación de Asia, o los planes de desplegar una conectividad directa a la Antártica, logro que además puede representar una gran contribución para la comunidad científica mundial.

4.3 Aprovechar los servicios desarrollados por los grandes experimentos científicos de la región

Análisis de las capacidades que puedan ofrecer los grandes experimentos de la región (ESO, PAO, LAGO, etc.) más allá de la mera explotación de la conectividad del ancho de banda del cable entre Latinoamérica y Europa, pues es probable que ya tengan integradas capacidades integradas FAIR y de simulación en entornos virtualizados.

4.4 Ampliación y fortalecimiento de los servicios de LA Referencia

Análisis de las capacidades que puedan ofrecer los grandes experimentos de la región (ESO, PAO, LAGO, etc.) más allá de la mera explotación de la conectividad del ancho de banda del cable entre Latinoamérica y Europa, pues es probable que ya tengan integradas capacidades integradas FAIR y de simulación en entornos virtuales.

4.5 Selección e implementación de casos de uso

Selección de casos de uso con una implantación continental que puedan servir como primeros usuarios del ecosistema digital propuesto y que complementen aquellos del punto 3 anterior. Algunos candidatos podrían ser:

- Áreas científicas relativas a la Bioinformática, tales como las actividades desarrolladas por el proyecto CABANA (<https://www.cabana.online/es>) .

- Áreas relativas a la docencia con cursos que abarquen temáticas de interés para la región, tales como los manejados por el proyecto de formación en Física de Altas Energías, LaCONGA (<https://laconga.redclara.net/>), o en energía con iniciativas como la red global de centros de energía sostenible regionales (Global Network of Regional Sustainable Energy Centres GN-SEC), de Naciones Unidas, UNIDO (<https://training.gn-sec.net/login/index.php>).

4.6 Recursos Humanos

Análisis de las capacidades en recursos humanos necesarios en los actores principales para hacer una estimación correcta de los siguientes pasos a seguir: RedCLARA, LA Referencia, SCALAC, usuarios del ecosistema.

4.7 Apoyo de la banca multilateral de desarrollo

El apoyo de la banca multilateral a proyectos y organizaciones regionales ha estado muy limitado por las características de estas organizaciones. Es importante estructurar esquemas de financiamiento con carácter internacional en los ámbitos de la ciencia para que se puedan generar las sinergias requeridas y no solo el fortalecimiento de las capacidades nacionales, que, aisladas o poco integradas, resultan insuficientes.

4.8 Profundizar el trabajo cooperativo con Europa y sus ecosistemas digitales

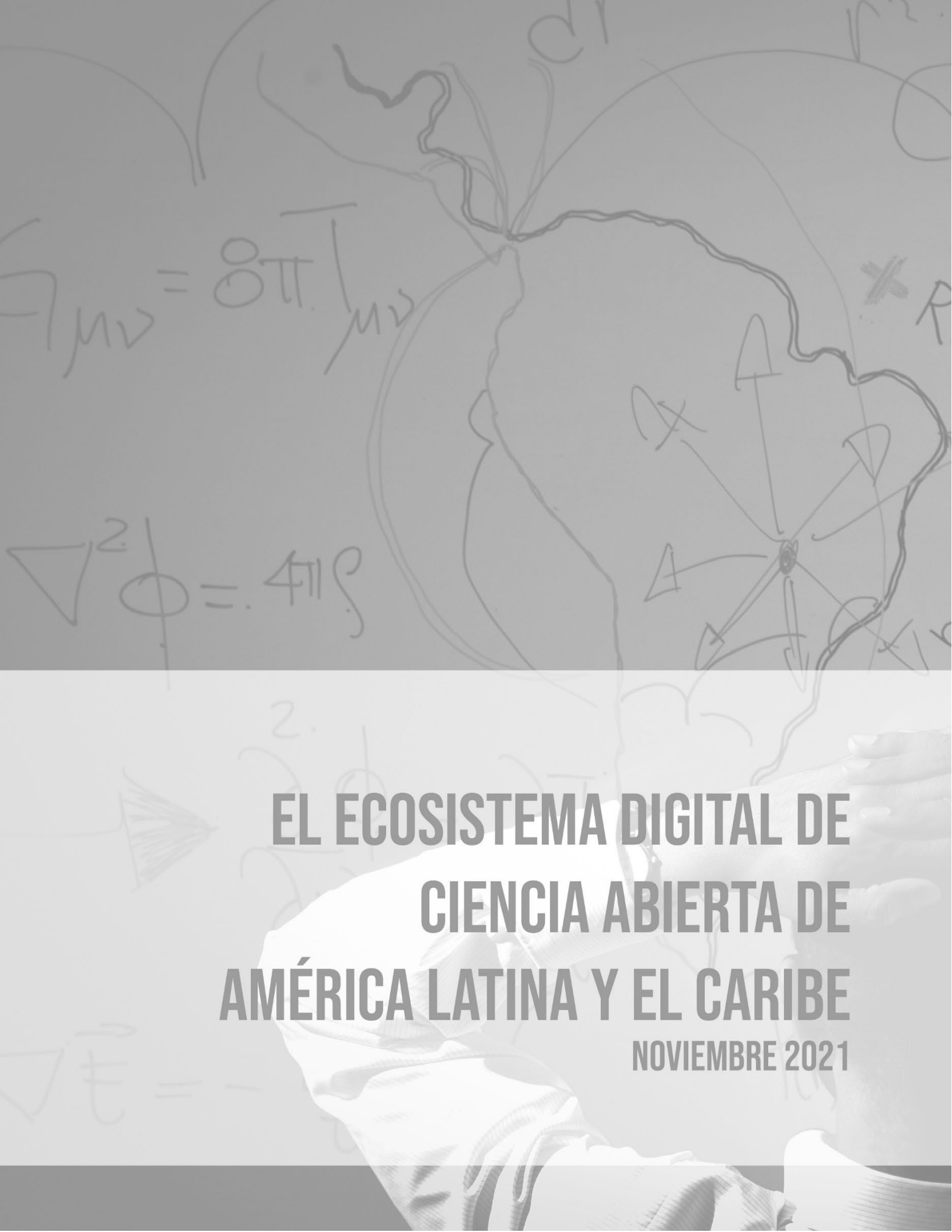
Hay una rica historia de apoyo por parte de la Comisión Europea que ha permitido construir una parte importante de la infraestructura regional requerida. Es importante ahondar estos lazos y avanzar en la búsqueda de todos los apoyos posibles, a través de acuerdos con las grandes infraestructuras y centros de investigación europeos, sus redes de cómputo, sus centros de excelencia en tecnologías de transformación digital y los programas de financiamiento y llamados en el contexto del programa Horizonte Europa. En este sentido es importante identificar proyectos de Horizonte Europa, Digital Europe, y otros, que sean susceptibles de proveer soluciones, ideas y capacidades al ecosistema digital latinoamericano de Ciencia Abierta.

Referencias

- [1] M. Woelfle, P. Olliaro, y M. H. Todd, «Open science is a research accelerator», *Nat. Chem.*, vol. 3, n.º 10, pp. 745-748, oct. 2011, doi: 10.1038/nchem.1149.
- [2] Unesco, «Proyecto de recomendación sobre la ciencia abierta». UNESCO, sep. 08, 2021. Accedido: nov. 10, 2021. [En línea]. Disponible en: https://unesdoc.unesco.org/ark:/48223/pf0000378841_spa
- [3] Carlos Seaton y Luis Cadenas, «Bella 2030: Alianza digital entre América Latina, el Caribe y Europa.» jun. 01, 2021. Accedido: sep. 06, 2021. [En línea]. Disponible en: <https://www.redclara.net/images/docs/BELLA2030-Alianza-Digital-ES.pdf>

Agradecimientos

Agradecemos a REDALYC, SciELO y Latindex, por sus importantes contribuciones textuales para la elaboración del presente documento.



**EL ECOSISTEMA DIGITAL DE
CIENCIA ABIERTA DE
AMÉRICA LATINA Y EL CARIBE**
NOVIEMBRE 2021