

## LA CIENCIA DIGITAL EN AMÉRICA LATINA: ALCANCE Y BENEFICIOS

**Martín del Castillo**

Universidad Nacional de San Martín  
ORCID ID: 0000-0002-3494-9295  
mdelcastillo@unsam.edu.ar

**Valeria Arza**

Universidad Nacional de San Martín  
ORCID ID: 0000-0003-4819-1777  
varza@unsam.edu.ar

## DIGITAL SCIENCE IN LATIN AMERICA: SCOPE AND BENEFITS

**Cómo citar este artículo/Citation:** del Castillo, Martín y Arza, Valeria (2021). La ciencia digital en América Latina: alcance y beneficios. *Arbor*, 197(799): a595. <https://doi.org/10.3989/arbor.2021.799008>

**Copyright:** © 2021 CSIC. Este es un artículo de acceso abierto distribuido bajo los términos de la licencia de uso y distribución *Creative Commons Reconocimiento 4.0 Internacional (CC BY 4.0)*.

Recibido: 25 septiembre 2020. Aceptado: 28 diciembre 2020. Publicado: 9 abril 2021

**RESUMEN:** El empleo de herramientas digitales está transformando los procesos de producción científica y sus impactos. En este trabajo evaluamos en qué medida la digitalización en ciencia ha alcanzado a la actividad científica en América Latina y cuáles han sido sus efectos. Para eso, utilizamos datos bibliométricos de tres países latinoamericanos de tamaño medio, Argentina, Chile y Colombia, que tienen una trayectoria científica importante y el análisis lo hemos centrado en cuatro disciplinas: ciencias biológicas y agrícolas, ciencias de la tierra y planetarias, ciencias ambientales y ciencias decisionales. Encontramos que la digitalización ha ido creciendo durante los últimos 25 años, pero la brecha con los países líderes sólo se ha reducido levemente. Asimismo, la incidencia de la digitalización varía por disciplina. En términos de beneficios asociados a la digitalización, nuestros resultados muestran que los artículos que utilizan prácticas o herramientas de ciencia digital tienen mayor impacto académico (tienen más citas), más colaboración (más coautorías) y más internacionalización (autorías de un mayor número de países). Es decir, en estas disciplinas y en estos países, la digitalización ha logrado aumentar la visibilidad de la investigación, potencialmente su calidad y, así, los retornos sociales de los recursos invertidos. También se han ampliado las redes, propiciando de esta forma un mejor aprovechamiento de la inteligencia colectiva y la integración de los problemas de investigación locales a una agenda global, aumentando potencialmente los recursos que se destinan a la investigación en estos países.

**PALABRAS CLAVE:** ciencia digital; difusión; impacto; colaboración; bibliometría; América Latina.

**ABSTRACT:** The use of digital tools is transforming scientific production processes and their impacts. In this article we evaluate to what extent digitalization in science has reached Latin American scientific activity and what have been its effects. For this purpose, we use bibliometric data from three medium-sized Latin American countries that have an important scientific trajectory: Argentina, Chile, and Colombia, focusing the analysis on four disciplines: biological and agricultural sciences, earth and planetary sciences, environmental sciences, and decisional sciences. We find that digitalization has been growing for the past 25 years, but the gap with leading countries has only narrowed slightly. Likewise, the incidence of digitalization varies by discipline. In terms of benefits associated with digitalization, our results show that articles that use digital science practices or tools have greater academic impact (have more citations), more collaboration (more co-authors), and more internationalization (authors from a larger number of countries). In other words, in these disciplines and in these countries, digitalization has managed to increase the visibility of research, potentially its quality and, thus, the social returns of the resources invested. Networks have also been expanded, thus promoting a better use of collective intelligence and the integration of local research problems into a global agenda, potentially increasing the resources devoted to research in these countries.

**KEYWORDS:** Digital science, diffusion, impact, collaboration, bibliometrics, Latin America.

## 1. INTRODUCCIÓN

El empleo de herramientas digitales tiene efectos sobre todas las instancias de la producción de conocimiento científico, desde el diseño y organización de la investigación, la recolección de datos, su análisis y difusión hasta la evaluación de sus resultados. La digitalización incide incluso en la cultura académica (haciendo de la ciencia una práctica más abierta, más global y colaborativa) porque puede convertir los lugares de trabajo en espacios virtuales y globales que facilitan la colaboración entre científicos y otros actores que no provienen del mundo académico. El concepto «ciencia digital» busca dar cuenta de todas estas aristas de la digitalización de los procesos, espacios y contenidos de la producción de conocimiento científico.

Si bien se trata de un fenómeno en expansión que concita el interés de quienes hacen políticas públicas y de la propia comunidad académica, no hay evidencia concreta de su alcance ni de la situación en los países de América Latina.

En este artículo nos proponemos dos objetivos que abordamos con métodos bibliométricos: analizar cuál es y cómo ha evolucionado la penetración de la digitalización en la producción de conocimiento científico en la región. Ponemos el foco en tres países de tamaño medio que tienen una trayectoria científica importante, Argentina, Chile y Colombia en cuatro disciplinas de diferentes áreas de conocimiento en las cuales estos países tienen también un recorrido notable. Las elegidas fueron ciencias biológicas y agrícolas, ciencias de la tierra y planetarias, ciencias ambientales y ciencias decisionales.

En esta elección contemplamos que hubiera diversidad de áreas de conocimiento, dado que el potencial de la digitalización puede ser distinto, así como también sus impactos, que es lo que nos interesa evaluar en el segundo objetivo. Más específicamente, estudiamos la relación de la ciencia digital con dos cuestiones de interés: la formación de redes de colaboración y el impacto o visibilidad académica. Los documentos reconocen que la digitalización puede dar lugar a un aumento de las colaboraciones al habilitar una mayor flexibilidad geográfica y temporal. Estos efectos, si bien suelen ser citados no han sido evaluados empíricamente en el contexto de América Latina.

En lo que sigue, presentamos una breve revisión de la literatura para contextualizar conceptualmente el artículo. A esto sigue una descripción de la fuente de

datos y de los métodos empleados para cumplir con los dos objetivos. Después pasamos a la sección de resultados donde se presenta y analiza la evidencia recogida. El artículo concluye con una sección destinada a la discusión y síntesis de los resultados obtenidos.

## 2. REVISIÓN DE LA LITERATURA

El neologismo «ciencia digital» busca dar cuenta de los cambios y transformaciones en los procesos y prácticas de la investigación científica a partir del uso extendido de internet y las tecnologías de la información y las comunicaciones (TIC) o, más ampliamente, de las tecnologías digitales en todas sus etapas. A lo largo del tiempo, han surgido una gran variedad de términos que, de una u otra forma, hablan o hacen alusión al papel transformador de las TIC y las tecnologías de redes en la producción de conocimiento. Tsatsou (2016) identifica, por ejemplo, el uso muchas veces intercambiable de los términos *digital research*, *internet research*, *online research*, y *e-research*. Pacheco, Nascimiento y Weber (2018) parten de los conceptos de ciberinfraestructura, *e-science* (ambos muy presentes en las discusiones de política pública en Estados Unidos y la Unión Europea) para definir el alcance de la digitalización en la ciencia. Ciber ciencia, ciencia 2.0, ciencia de la web son otros términos que han aparecido en distintos documentos de política en ciencia y tecnología (Hey y Trefethen, 2002; Atkins *et al.*, 2003; Sargent, 2006; Comisión Europea, 2013).

Otro término emparentado, que es anterior a los de *e-science* y de ciberinfraestructura, es el de «colaboratorios». Esta palabra fue utilizada, en primer lugar, para designar a «un centro sin paredes en el cual los investigadores pueden trabajar de manera conjunta sin importar la ubicación física» (Wulf, 1993: 854). Buscaba caracterizar el uso de Internet para intercambiar grandes cantidades de datos, el acceso a recursos computacionales y el uso remoto de instrumentos. Con el tiempo su significado se amplió hasta cubrir todo el conjunto de actividades requeridas para realizar ciencia e incluyó toda la variedad de interacciones humanas que son un elemento de la colaboración científica. En un taller realizado en la Universidad de Michigan en 2001 se utilizó el término para nombrar a:

«una entidad organizativa que extiende la distancia, da lugar a una interacción humana rica y recurrente orientada a un área de investigación común y que proporciona acceso a fuentes de datos, artefactos y herramientas necesarias para realizar tareas de investigación» (Olson, Zimmerman y Bos, 2008: 3).

La magnitud y alcance de las transformaciones en métodos y procesos de producción de conocimiento es tal que algunos autores hablan de la emergencia de nuevas maneras de resolver problemas y avanzar en el conocimiento científico (Hey y Trefethen, 2002 y 2008; Atkins *et al.*, 2003; Nentwich, 2003). Por un lado, y en primer lugar, la ciencia computacional, que implica la utilización de modelos matemáticos y técnicas de análisis cuantitativos en computadoras, se ubica como una metodología científica complementaria a los pilares tradicionales de la ciencia experimental/observacional y de la teórica/analítica (Atkins *et al.*, 2003).<sup>1</sup> Por otro, y más recientemente, tendríamos a la *e-science* o ciencia basada en la exploración o uso intensivo de datos (*datacentric science*) (Hey y Trefethen, 2002 y 2008; Hey, Tansley y Tolle, 2009). Aquí, la generación y captura de datos a través de dispositivos digitales y su procesamiento mediante software específicos junto con la complejidad de los problemas científicos estaría dando lugar a la «próxima generación de problemas científicos, y a las herramientas y tecnologías colaborativas que serán requeridas para solucionarlos» (Hey y Trefethen, 2008: 5).

La producción de conocimiento se apoya, en definitiva y de manera creciente, en herramientas de carácter digital para, por ejemplo, mediar la comunicación entre la comunidad científica, recolectar y analizar datos (digitales) o difundir y comunicar los resultados de las investigaciones. Muchas de las metodologías de investigación convencionales tienen ahora una versión virtual con características o funcionalidades nuevas: como encuestas online, entrevistas por mail o Skype, grupos focales virtuales u online, *web text analysis*, *digital conversation analysis*, etc. Al mismo tiempo aparecen otros enfoques metodológicos que deben su existencia a las nuevas posibilidades que abre el dominio tecnológico digital: *web crawling*, *social media data* y *text scraping* y *text mining*, *sentiment analysis*, *hyperlink analysis*, y *recommendation systems* (Rogers, 2013 y 2015; Kennedy *et al.*, 2015). Estos cambios permiten abordar los temas y preguntas de investigación clásicos con nuevas herramientas metodológicas y también dan lugar a nuevas líneas de investigación, y hacen, por supuesto, a la digitalización misma objeto de investigación. Asimismo, junto

a lo que se ha dado en llamar las tecnologías de trabajo cooperativo asistido por computadora (*Computer Supported Cooperative Work*, CSCW)<sup>2</sup>, la digitalización habilita la extensión de las redes de colaboración en cantidad y diversidad (Olson y Olson, 2012).

## 2.1. El alcance de la digitalización: prácticas y políticas

Aunque no parece haber estudios que dimensionen de manera cuantitativa el alcance y la evolución del fenómeno en América Latina, hay una evidencia, al menos de carácter circunstancial o anecdótico, que indica el crecimiento del interés que concita en ámbitos académicos y políticos, tanto dentro como fuera de la región, y otra evidencia, no siempre sistemática, de una incidencia variable según disciplina.

En relación al interés que concita el tema en ámbitos académicos, Pacheco, Nascimento y Weber (2018) identifican, en su revisión de la bibliografía, que en los últimos años ha ido en aumento el número de artículos publicados en revistas científicas que responden a algunas de las palabras que refieren a la digitalización de la ciencia. En base a encuestas a investigadores, Arcila, Piñuel y Calderín (2013) y Dutton y Meyer (2008) encuentran una predisposición o actitud positiva hacia la utilización de herramientas digitales en el área de ciencias de la comunicación en Iberoamérica y en el de ciencias sociales en Reino Unido, respectivamente. Pero Arcila, Piñuel y Calderín (2013) también encuentran que el uso intensivo y avanzado de las TICs es más bien limitado en el área de conocimiento que analizan.

El estudio de Galindo-Rueda (2020) ofrece un intento de medir la evolución e incidencia del fenómeno a partir de un análisis bibliométrico. Identifica que a nivel global la investigación científica que utiliza o busca desarrollar herramientas de inteligencia artificial (IA) creció entre el 2006 y el 2016 a un ritmo más acelerado que el del total de las publicaciones, siendo los lugares en donde está mayoritariamente concentrada China (con el 27% de las publicaciones relacionadas con IA en 2016), Estados Unidos (con el 19%), la Unión Europea (con el 12%) y la India (con el 11%). En el período mencionado, el volumen anual de publicaciones relacionadas con la IA creció en el mundo el 150% mientras que el total de publicaciones indexadas aumentó el 50%.

1 Dougherty y Dunne (2012), por ejemplo, exploran la ciencia digital en el contexto del descubrimiento de nuevas drogas. Concluyen que la digitalización aporta nuevas formas de medir, analizar y modelizar compuestos químicos y el comportamiento de enfermedades que complementan a las ciencias terapéuticas (*therapy sciences*), al tiempo que expanden los límites del conocimiento.

2 CSCW emerge como campo de estudio a mediados de los años ochenta con la aparición simultánea de conferencias, revistas, libros y cursos de universidades (Grudin, 1994).

En relación a la incidencia variable según disciplina, Tsatsou (2016), en un estudio cualitativo que realizó en el Reino Unido acerca del uso de herramientas, recursos y servicios digitales en varios campos de las ciencias sociales (negocios, educación, historia, ciencias políticas y literatura), concluye que los principios disciplinares y las tradiciones de investigación influyen en cómo y cuántas tecnologías digitales son empleadas con fines de investigación. Holmberg y Thelwall (2014) alcanzan resultados similares, aunque concentrando la atención en la esfera de la difusión y la comunicación y en un número más variado de campos. Analizan cómo especialistas de diez disciplinas distintas (entre ellas, astrofísica, bioquímica, humanidades digitales, economía, historia de la ciencia y descubrimiento de fármacos) utilizan el sitio de *microblogging* Twitter. Encuentran diferencias entre las distintas disciplinas en el uso de la plataforma para la comunicación académica, siendo los investigadores en bioquímica y astrofísica quienes más la utilizan con esa finalidad.

Esta incidencia variable según disciplina se expresa no solo en términos del uso de herramientas digitales para producir o comunicar conocimientos, sino también en la cantidad de aplicaciones desarrolladas y montadas sobre las e-infraestructuras. Andronico *et al.* (2011), que analizan el estado de las e-infraestructuras y las aplicaciones asociadas que dan soporte a la investigación, encuentran que el número de aplicaciones disponibles para cada disciplina es distinto, sugiriendo que la penetración es desigual según el campo del conocimiento. Ciencias de la vida, seguido por ciencia de la computación y matemáticas, y luego por astronomía y astrofísica son los dominios científicos con mayor cantidad de aplicaciones.

El estudio más completo que encontramos en términos de variabilidad de prácticas y cobertura disciplinar fue realizado por la OECD a partir de una encuesta realizada a doce mil investigadoras e investigadores de todo el mundo durante 2018, cuyos resultados se sintetizan en Galindo-Rueda (2020). Con el fin de identificar patrones, determinantes y efectos de la digitalización en la investigación científica en un amplio rango de disciplinas, estudiaron: 1) el uso de herramientas de productividad (*productivity tools*) para llevar a cabo tareas regulares, como recuperar información y colaborar con colegas, 2) la puesta a disposición de otros los datos y códigos ob-

tenidos, 3) la utilización o desarrollo de datos y métodos computacionales no convencionales como el *big data*, y 4) la comunicación científico digital. Encontraron que ciencias de la computación es la disciplina que, relativamente, más uso hace de herramientas digitales avanzadas (*big data*). En el otro extremo se ubicaron ciencias sociales y psicología. En ciencias de la tierra y planetarias la intensidad de uso se ubicó por encima de ciencias biológicas y agrarias y ésta a su vez se colocó por arriba de la categoría energía y ciencias ambientales.

El crecimiento del fenómeno de la ciencia digital se puede apreciar también en el interés de los gobiernos de las principales economías desarrolladas que, desde al menos los últimos años de los 90 y los primeros de la década siguiente, vienen elaborando políticas públicas específicas para promover la digitalización en un número creciente de áreas del conocimiento.

En 1999, por ejemplo, el gobierno del Reino Unido lanzó el programa *e-science*, una iniciativa que buscó desarrollar tecnologías genéricas, conocidas como *middleware*, para permitir que diferentes recursos puedan estar disponibles en red y creen cuadrículas o redes (*grids*) informáticas (Hey y Trefethen, 2002 y 2008). El gobierno de Estados Unidos tiene un programa llamado *Advanced Cyberinfrastructure* que apoya y coordina el desarrollo, adquisición y provisión de recursos, herramientas y servicios de ciberinfraestructura para el avance y la transformación de la ciencia (Atkins *et al.*, 2003). Australia lanzó en torno a 2005 dos programas en estas mismas direcciones: el *eResearch Framework*, una estrategia nacional de inversión en la creación y fortalecimiento de e-infraestructura para la investigación colaborativa, y el proyecto MAPS (*Middleware Action Plan and Strategy*) (Sargent, 2006). Y la Unión Europea, dentro del programa *Horizon 2020*, ha estado promoviendo la ciencia digital con distintos instrumentos a través de la Dirección General de Redes de Comunicación, Contenido y Tecnología (DG *Connect*) (Comisión Europea, 2013).

A nivel regional, en Latinoamérica, encontramos que se adoptaron medidas, unos años más tarde, para favorecer la digitalización de la ciencia, pese a que no parece haber una política que la promueva de manera explícita.<sup>3</sup> Es decir, no hay una política que exprese la necesidad de reconocer a la ciencia digital, que establezca metas

3 América Latina sí tiene una tradición más antigua en términos de digitalización de las revistas académicas y bases de datos bibliográficas latinoamericanas, que se remonta a los años ochenta y que creció considerablemente en los noventa con la consolidación de internet (Alonso-Gamboia y Russell, 2012). Existen varias colecciones digitales de revistas (*Latindex* en 1995, *SciELO* en 1997, *Redalyc* en 2003) que han contribuido a visibilizar y mejorar su calidad (Packer y Meneghini, 2007).

y prioridades de desarrollo y que traduzca los objetivos en planes de acción concretos y transversales a todas las áreas de la producción de conocimiento, como se señala en Arza *et al.* (2019). Dicho estudio releva veinticuatro iniciativas en los tres países que analizamos en este artículo de las cuales cinco corresponden a marcos legales, tres a estructuras administrativas y dieciséis a instrumentos operacionales<sup>4</sup>. Entre los instrumentos operacionales se destacaron los esfuerzos de los organismos públicos de ciencia y tecnología para proveer infraestructura (ciberinfraestructuras, plataformas y portales, algunas de alcance nacional e insertas en otras de alcance regional)<sup>5</sup>. Los marcos legales identificados estuvieron asociados al acceso abierto de los resultados de las investigaciones financiadas con fondos públicos y a la promoción de la ciencia abierta.<sup>6</sup>

## 2.2. Los beneficios de la digitalización

Existe en la literatura una serie de factores que anticipan beneficios asociados a la digitalización, que de alguna forma justifican las políticas de promoción de la ciencia digital. Por ejemplo, es de esperar que la digitalización permita cada vez más trascender las fronteras, habilitando una mayor flexibilidad en términos de distancia y tiempo (Olson, Zimmerman y Bos, 2008). Mientras que en el pasado hubiera sido considerado necesario juntar al personal investigador en un único laboratorio, hoy en día muchas de estas asociaciones se pueden llevar a cabo a distancia gracias a las tecnologías digitales (e-mail, videoconferencias, pizarras y bases de datos compartidas). Hay software diseñado para brindar apoyo a actividades que se desarrollan en cualquiera de las combinaciones de tiempo y lugar: mismo/diferente lugar y mismo/diferente tiempo (Olson y Olson, 2012)<sup>7</sup> y para todas las etapas del proceso de producción de conocimiento.

El aumento de la colaboración está a su vez identificado con un mayor impacto o visibilidad académica (citas) de las publicaciones científicas (Persson, Glänzel y Danell, 2004) y podría dar lugar a una mayor eficiencia y eficacia. Mayor eficiencia, si al aumentar los incentivos para compartir información (por la facilidad de acceso a infraestructuras) se reduce la probabilidad de realizar inversiones redundantes (evitar la duplicación de esfuerzos) y si la ciberinfraestructura es utilizada para compartir recursos que son luego empleados en responder distintas preguntas de investigación. Y mayor eficacia, al permitir la intervención de nuevos actores, pertenecientes o no a la academia, y aumentar por tanto la inteligencia colectiva: las contribuciones a la resolución de interrogantes científicos pueden ser más inter y transdisciplinarias.

Además de facilitar la colaboración, la digitalización hace posible generar, juntar, almacenar, compartir y analizar grandes cantidades de datos con instrumentos/herramientas crecientemente especializados y sofisticados (Olson, Zimmerman y Bos, 2008). Esto se traduce en una mayor centralidad de los datos en el proceso de investigación que puede dar lugar a nuevas líneas o posibilidades de investigación, así como cambios en los métodos de investigación (Hey y Trefethen, 2008). A su vez, el crecimiento exponencial de la cantidad y variedad de datos puede exigir la colaboración de varias experticias para poder procesarlos, analizarlos, visualizarlos e interpretarlos.

Asimismo, la ciencia digital puede aumentar la transparencia y la confianza en los resultados de una investigación al facilitar que diferentes grupos de investigación tengan acceso a los datos utilizados (Comisión Europea, 2013)<sup>8</sup>. A su vez, esto puede redundar en una mayor inclusión cuando las ciberinfraestructuras facilitan que grupos de investigación tradicional-

- 4 Las categorías se basan en la clasificación propuesta por Sagasti y Araoz (1976) y retomada por Emiliozzi, Lemarchand y Gordon (2009).
- 5 En Chile la red se llama REUNA (<https://www.reuna.cl/>), en Colombia, RENATA (<https://www.renata.edu.co/>) y, de forma más limitada, con menos servicios disponibles, en Argentina está INNOVA RED (<https://www.innova-red.net/>). A nivel regional, los tres países participan de plataformas, como la Red Clara (<https://www.redclara.net/index.php/es/>) y LAReferencia (<http://www.lareferencia.info/es/>), que aglutinan diversos esfuerzos económicos en torno a aspectos asociados a la digitalización (Arza *et al.*, 2019). Estas iniciativas no están sustentadas por entidades públicas, sino por los aportes de sus socios (instituciones de educación superior e investigación), pero los principales organismos de CyT participan de las mismas.
- 6 Nos referimos a la Ley de Repositorios Digitales Institucionales de Acceso Abierto sancionada en 2013 en Argentina (<http://servicios.infoleg.gob.ar/infolegInternet/anexos/220000-224999/223459/norma.htm>), al artículo vinculado al acceso abierto en la Ley N° 21.105 de 2018 de creación del Ministerio de Ciencia en Chile (<https://www.bcn.cl/leychile/navegar?idNorma=1121682>), y a la resolución 0167 de 2019 por la cual se adoptan los lineamientos para una Política de Ciencia Abierta en Colombia (<https://www.colciencias.gov.co/normatividad/resolucion-0167-2019>).
- 7 Olson y Olson (2012) reseñan variedad de software/tecnologías diseñado para ejecutarse en red para apoyar actividades de grupos u organizaciones.
- 8 La digitalización de la ciencia habilitó o alimentó el movimiento hacia el acceso abierto, el principio según el cual los datos y resultados de la investigación financiada con recursos públicos deberían ser accesibles a los contribuyentes que pagaron a la investigación (Hey y Trefethen 2008).



mente inhabilitados por falta de recursos o acceso a equipamientos, puedan colaborar remotamente.

### 3. OBJETIVOS DE LA INVESTIGACIÓN

La falta de una política integral sobre digitalización de la ciencia en América Latina en contraposición con el interés público que el tema despierta en las economías desarrolladas, sus esquemas de promoción de la digitalización y, más en general, su liderazgo científico y tecnológico, nos permiten suponer que la región se encuentra rezagada en términos del alcance de la digitalización de la ciencia. Y, sin embargo, son numerosos y variados los efectos positivos que los documentos de política pública asocian con la digitalización de la producción de conocimiento. En este sentido, el objetivo del presente estudio es doble:

1. dimensionar la penetración del uso de herramientas digitales en la producción de conocimiento científico en algunos países latinoamericanos que sirvan como casos testigo de lo que acontece en la región;
2. identificar la relación entre la digitalización de la ciencia en la región y la formación de redes de colaboración y el impacto académico (en términos de citas).

Para ello realizaremos un análisis bibliométrico de la producción académica en cuatro disciplinas en tres países latinoamericanos de tamaño medio.

### 4. DATOS Y METODOLOGÍA

Empleamos técnicas cuantitativas para realizar un análisis bibliométrico que nos muestre la incidencia de la digitalización de la ciencia en el contexto de América Latina y que nos permita luego analizar la relación entre la ciencia digital y la formación de redes de colaboración y el impacto académico.

Para llevar adelante el estudio recurrimos a la base bibliográfica *Scopus*. En la actualidad, *Scopus*, propiedad de Elsevier y disponible desde el 2002, es la base interdisciplinaria de citas y resúmenes curados más grande que existe en el mundo (Schotten *et al.*, 2017).

No obstante, una limitación de este estudio que comparte con otros estudios bibliométricos es que

*Scopus*, al igual que cualquiera de las otras bases de datos bibliográficas existentes como la *Web of Science* (WoS), no es exhaustiva, pues no existen bases bibliográficas capaces de cubrir la totalidad de la producción científica de un país. Más aún, las bases de datos disponibles difieren en la cobertura de disciplinas y países (ver el Capítulo 5 en Chavarro Bohórquez, 2016). Esto es particularmente problemático cuando se quiere analizar la actividad científica en países en desarrollo.

Sin embargo, cabe resaltar, que esta limitación no ha impedido el uso de los metadatos de los artículos científicos de bases determinadas para estudiar las dinámicas de la producción de conocimiento en América Latina (Confraria y Vargas, 2017; Morán-Mariños *et al.*, 2020; Ruiz-Patiño *et al.*, 2020). En todo caso, debe tenerse presente cuando se analizan los resultados y las conclusiones deben limitarse a las publicaciones indexadas en *Scopus*.<sup>9</sup>

A continuación, explicitamos los criterios de selección del espacio temporal, de países y de disciplinas, los pasos seguidos para la identificación de los artículos de ciencia digital y la manera en que estimamos el efecto de la digitalización sobre las variables de interés junto con los indicadores construidos.

#### 4.1. Selección del recorte temporal

Para analizar la evolución de la incidencia de la ciencia digital tomamos los artículos publicados entre 1990 y 2017. Elegimos una ventana de tiempo extensa que permitiera visualizar tendencias. En el momento del análisis, como la cantidad de publicaciones al año puede variar (particularmente en disciplinas con una menor cantidad de publicaciones) y dado que el interés está en el hecho estilizado (la tendencia), en lugar de tomar la cantidad de publicaciones en un año dado calculamos la media móvil de cuatro años para visualizar el comportamiento de las series.

Para el segundo de los objetivos tomamos los artículos publicados durante los últimos diez años desde el momento en que diseñamos la investigación (diciembre de 2018). El período analizado va, entonces, de 2008 a 2017.

<sup>9</sup> Recurrimos a *Scopus* porque teníamos acceso a través de la institución de la cual formamos parte. No contábamos con acceso a la base de datos bibliográfica *Web of Science*. No complementamos los registros obtenidos en *Scopus* con *SciELO*, que tiene una mayor cobertura de revistas de la región, porque la información que se puede descargar no incluye los resúmenes de los artículos que, como se puede ver más abajo, los utilizamos para desagregar las disciplinas seleccionadas en subdisciplinas a través de un ejercicio de *topic modeling*.

#### 4.2. Selección de países

El análisis se focaliza en tres países latinoamericanos que tienen una trayectoria científica importante: Argentina, Chile y Colombia. Los elegimos por ser países de tamaño medio con excelente desempeño científico en la región según los indicadores convencionales.<sup>10</sup>

#### 4.3. Selección de disciplinas

Con el fin de evaluar en qué medida la ciencia digital ha penetrado en la producción científica en estos países, decidimos focalizar el análisis en algunas disciplinas, sabiendo que las prácticas digitales son variadas y pueden tener incidencia muy diferente en distintos campos disciplinarios (Holmberg y Thelwall, 2014; Tsatsou, 2016; Galindo-Rueda, 2020). Para elegir las disciplinas, utilizamos distintos criterios. Por un lado, buscamos que hubiera cierta diversidad para poder captar distintas prácticas. Así elegimos disciplinas de tres grandes áreas, tal como están definidas en *Scopus*: ciencias sociales, ciencias físicas, y ciencias de la vida. Por otro lado, buscamos disciplinas en las cuales al menos dos de los tres países elegidos evidenciaran cierto grado de especialización relativa. Esto podría sugerir que en estos casos las prácticas científicas se encuentran mejor emparentadas con lo que serían las prácticas de vanguardia a nivel internacional.

Para analizar la especialización relativa en cada disciplina por parte de cada uno de los tres países recurrimos al índice de ventajas comparativas reveladas (VCR) propuesto por Balassa (1965). El valor que toma este índice resulta de comparar el peso medido en publicaciones de una disciplina dada en un país concreto con el peso de la misma disciplina a nivel mundial:

$$VCR_{ij} = \frac{P_{ij}/P_i}{P_j/P}$$

Donde  $P_{ij}$  es la cantidad de publicaciones del país  $i$  en la disciplina  $j$ ,  $P_i$  es el número total de publicaciones del país  $i$ ,  $P_j$  es el número total de

publicaciones en la disciplina  $j$  a nivel mundial y, finalmente,  $P$  es el número total de publicaciones a nivel mundial.

La tabla A1, que presentamos en el anexo 1, muestra los valores del índice de VCR según país y disciplina, utilizando todos los documentos publicados entre 2013 y 2017. Se puede observar que la producción científica de los tres países es particularmente fuerte, evidencia ventajas comparativas a nivel internacional, en ciencias biológicas y agrícolas. Al tener esta disciplina relevancia económica en términos de las estructuras productivas, decidimos seleccionarla.

Para la elección de las otras tres disciplinas, buscamos que al menos dos de los tres países tuvieran ventajas comparativas procurando además que pertenecieran a distintas grandes áreas para asegurar diversidad de prácticas como mencionamos antes. Las disciplinas bajo estos criterios fueron: ciencias de la tierra y planetarias, ciencias ambientales y ciencias decisionales.

Seleccionados los países y las disciplinas, procedimos a descargar las referencias bibliográficas de *Scopus*. Descargamos los metadatos de todos los artículos (*Document type: article*) publicados en *journals* (*Source type: journals*) en cada disciplina (*Subject area: Agricultural and Biological Sciences*) y país (*Country/territory = Chile*) entre el 2008 y el 2017 (descargamos los datos para Argentina y ciencias ambientales, luego para Argentina y ciencias decisionales, y así con todos los países y disciplinas, ver tabla 1).

De esta manera los datos a analizar para el segundo objetivo provienen de alrededor de 87.000 artículos publicados en revistas científicas indexadas en la base bibliográfica *Scopus* publicados entre el 2008 y el 2017. De este conjunto de publicaciones, el 54% pertenece a ciencias biológicas y agrícolas, el 25% a ciencias de la tierra y planetarias, el 18% a ciencias ambientales y el restante 3% a ciencias decisionales. Según país la distribución es la siguiente: 48% están firmados por una persona afiliada a una institución de Argentina, el 34% a una de Chile y el 18% a una de Colombia (tabla 1).<sup>11</sup>

10 Excluyendo a Brasil y a México que por su tamaño y lugar geoestratégico tienen características muy particulares, los tres países son los que tienen un ranking más alto en la región en términos de cantidad de publicaciones e índice h. Estos países tienen más publicaciones que los tres elegidos, como es de esperar por sus tamaños, y sus índice h son también mayores (datos de SCIMAGOJR: <https://www.scimagojr.com/>).

11 Estos porcentajes se obtuvieron siguiendo el método de conteo completo. Los datos bibliográficos se obtuvieron descargando los artículos publicados en un año, país y disciplina determinados. Un artículo firmado por un investigador afiliado a una institución de Argentina y otro a una de Chile y que pertenece a dos disciplinas distintas es contado cuatro veces. Los artículos se descargaron de esta manera por la limitación de descarga de 2000 registros por vez que impone *Scopus*.

**Tabla 1.** Artículos publicados en revistas académicas entre el 2008 y el 2017 según país y disciplina.

	Ciencias Biológicas y Agrícolas	Ciencias de la Tierra y Planetarias	Ciencias Ambientales	Ciencias Decisionales	Total (conteo completo) #	Total (conteo completo) %
<b>Argentina</b>	24,686	8,895	7,453	509	41,543	48%
<b>Chile</b>	12,875	11,067	5,070	1,007	30,019	34%
<b>Colombia</b>	9,917	2,134	2,950	845	15,846	18%
<b>Total (conteo completo) #</b>	47,478	22,096	15,473	2,361	87,408	100%
<b>Total (conteo completo) %</b>	54%	25%	18%	3%	100%	

Fuente: elaboración propia a partir de datos de *Scopus*.

#### 4.4. Identificación del alcance de la ciencia digital

Para evaluar el alcance de la ciencia digital en el conjunto de artículos seleccionado, identificamos primero palabras asociadas al uso de herramientas digitales en la producción de conocimiento científico a partir de los documentos que formaron parte de la revisión de la literatura.

Luego, con este primer conjunto de palabras realizamos una nueva búsqueda de artículos científicos publicados en revistas en la base de datos bibliográfica *Scopus* para todas las áreas de conocimiento en el campo «título, resumen y palabras clave»<sup>12</sup>. De esta manera identificamos artículos que utilizan herramientas digitales —además de aquellos que las tienen como objeto de estudio— y, a partir de las palabras claves de estos artículos que están asociadas a herramientas digitales (*Algorithm*, *Support Vector Machine*, *Remote Sensing*) y que se presentan con mayor frecuencia,<sup>13</sup> construimos la lista de términos que finalmente utilizamos para clasificar los artículos.<sup>14</sup> En la tabla A2, que figura en el anexo 2, presentamos esta lista.

La incidencia de la ciencia digital en cada país y año se mide como el cociente entre los artículos que ha-

cen uso de herramientas digitales y el total de artículos publicados.

Para evaluar la incidencia de la ciencia digital en las disciplinas seleccionadas se siguió un criterio similar al descrito en los párrafos anteriores. Pero en este caso refinamos las palabras claves para poder dar cuenta con mayor precisión de las herramientas digitales específicas a cada disciplina. Para ello, a partir del primer conjunto de palabras realizamos otra búsqueda de artículos en el campo de «título, resumen y palabras clave» pero ahora filtramos el resultado por las disciplinas seleccionadas; repetimos por ello la búsqueda cuatro veces<sup>15</sup>. Nuevamente, a partir de las palabras claves de los artículos así obtenidos que están asociadas a herramientas digitales y que se presentan con mayor frecuencia<sup>16</sup> construimos el conjunto de palabras que utilizamos para identificar a los artículos de ciencia digital. La tabla A2 (en el anexo 2) también presenta estos conjuntos de palabras por disciplina. Al igual que antes, el grupo de palabras de cada disciplina se utilizó para identificar en *Scopus*, a partir del campo de búsqueda «título, resumen y palabras claves», la producción de conocimiento científico (artículos en revistas) que hace uso de herramientas digitales en cada disciplina y país (tabla 2).

12 El comando de búsqueda utilizado fue: (TITLE-ABS-KEY(((digital OR virtual) AND (collaboration OR research) AND (plataform OR communit\* OR infrastructure OR technolog\*)) OR ((“web 2.0” OR “web2.0”) AND (infrastructure OR technolog\*)) OR “crowdsourcing” OR “big data” OR “text mining” OR “machine learning”))

13 Ordenamos las palabras claves en orden descendente en función de la cantidad de veces que se repiten. Luego de las cincuenta palabras claves que más se repiten seleccionamos aquellas que están asociadas a herramientas digitales.

14 Esta búsqueda no está exenta de limitaciones. Algunas disciplinas además de ser usuarias de herramientas de ciencia digital, también las tienen por objeto de estudio. En nuestro estudio a nivel de disciplinas, esto puede suceder exclusivamente en ciencias decisionales. La estimación de la evolución general de la producción académica en ciencia digital en ese caso debe ser tomada con cautela. Sin embargo, para el caso del gráfico 5, donde mostramos que el uso de la digitalización en general, puede estar sobrevalorada ya que la ciencia digital puede ser objeto de estudios en diferentes disciplinas en el área de ciencias sociales y de la computación. No esperamos que este sesgo difiera mucho por país.

15 Utilizamos el comando de búsqueda que aparece en la nota 12 y acotamos la búsqueda de artículos a los de una disciplina particular (*Subjetc area*: ciencias decisionales). Luego, a partir de las palabras claves de los artículos así obtenidos construimos la tabla que se presenta en el anexo 2. Repetimos este paso para cada disciplina seleccionada.

16 Procedimos según lo descrito en la nota 13.



**Tabla 2.** Artículos de ciencia digital publicados en revistas académicas entre el 2008 y el 2017 según país y disciplina.

	Ciencias Biológicas y Agrícolas	Ciencias de la Tierra y Planetarias	Ciencias Ambientales	Ciencias Decisionales	Total (conteo completo) #	Total (conteo completo) %
Argentina	759	543	342	90	1,734	39%
Chile	530	724	268	223	1,745	39%
Colombia	432	254	159	100	945	21%
<b>Total (conteo completo)</b>	1,721	1,521	769	413	4,424	100%
<b>Total (conteo completo)</b>	39%	34%	17%	9%	100%	

Fuente: elaboración propia a partir de datos de Scopus.

#### 4.5. Digitalización, redes de colaboración e impacto académico

Realizamos un ejercicio econométrico sencillo utilizando cada artículo como unidad de observación.

Elegimos tres variables dependientes, una que da cuenta del impacto académico (citas al año) y otras dos asociadas a la formación de redes de colaboración (cantidad de países distintos de afiliación y cantidad de firmas).

La variable explicativa de interés es una variable dicotómica que toma el valor 1 si el artículo hace uso de herramientas digitales (es de ciencia digital) y 0 en el caso contrario.

Incluimos como variables de control la edad del artículo, tres variables dicotómicas que toman el valor 1 si el país de afiliación es Argentina, Chile o Colombia, y otras cuatro variables dicotómicas que identifican la disciplina a la que pertenece el artículo.

Adicionalmente, con el fin de desgranar y controlar la heterogeneidad dentro de cada disciplina, clasificamos en cuatro grupos (o tópicos) el contenido de cada una de ellas. Hicimos para ello un análisis de modelado de tópicos o *topic modeling* que es una técnica de *machine learning* que permite clasificar de manera automática temas a partir de colecciones de palabras presentes en documentos de texto. En nuestro caso, utilizamos las descripciones o *abstracts* de los artículos publicados en revistas científicas (*journals*) entre el 2008 y el 2017.

Empleamos, dentro del modelado de tópicos, el modelo/ algoritmo conocido con el nombre de *Latent Dirichlet Allocation* (LDA) que trata a cada artículo científico como una mezcla de tópicos y a cada tópico como una composición de palabras. De esta manera, en un modelo de cuatro tópicos podríamos encontrarlos con artículos que son, por ejemplo, 70% tópico 1, 5% tópico 2, 15% tópico 3 y 10% tópico 4. Ese porcentaje que también puede leerse como probabilidad, es la proporción de palabras de un determinado docu-

mento (*abstract*, en nuestro caso) que son generadas por un tópico.

Incluimos como controles adicionales, entonces, variables que nos muestran la probabilidad que tiene un artículo de pertenecer a cada uno los tópicos dejando siempre uno como base, ya que la colinealidad entre los cuatro es perfecta. En la tabla A4, que figura en el anexo 4, presentamos una descripción del contenido de cada uno de los tópicos.

En suma, estimamos un modelo de regresión lineal para las ecuaciones [1] y [2]:

Estimación general

[1]

$$Dep. Var_i = \alpha + \beta_1 CD_i + \beta_2 AR_i + \beta_3 CH_i + \beta_4 CO_i + \beta_5 EDAD_i + \beta_6 AGRI_i + \beta_7 EART_i + \beta_8 ENVI_i + \beta_9 DECI_i + \sum_{j=1}^m \beta_{9+j} T_{j,i} + \mu_i$$

Según disciplinas

[2]

$$Dep. Var_i = \alpha + \beta_1 CD_i + \beta_2 AR_i + \beta_3 CH_i + \beta_4 CO_i + \beta_5 EDAD_i + \sum_{j=1}^m \beta_{5+j} T_{j,i} + \mu_i$$

Donde,  $m$ , en la sumatoria, es igual a doce tópicos para la estimación general [1] (tres tópicos por cada una de las cuatro disciplinas) y a tres tópicos para las estimaciones por disciplina [2].

En la tabla 3 están definidas las variables y en el anexo 3 se presentan las estadísticas descriptivas.

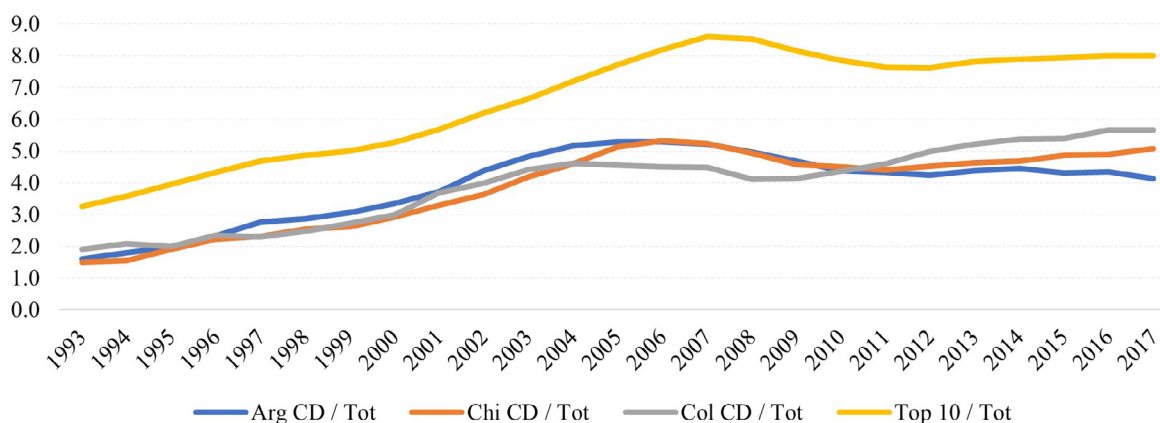
## 5. RESULTADOS

### 5.1. El avance de la ciencia digital en América Latina

El gráfico 1 muestra la evolución de la incidencia de la ciencia digital en los tres países seleccionados y en un conjunto de diez países líderes en producción académica a nivel mundial. Aquí todavía no acotamos el análisis a las disciplinas seleccionadas, el gráfico incluye el conjunto de publicaciones de cada uno de los países. Se puede observar que los tres países tienen un patrón de comportamiento similar a lo largo de

**Tabla 3.** Variables incluidas en el modelo de regresión lineal

VARIABLES DEPENDIENTES	DESCRIPCIÓN
TCxAño	Promedio anual de citas de un artículo
colab_int	Número de países distintos que aparece en la afiliación institucional de los autores de un artículo
AU cantidad	Número de autores de un artículo
VARIABLES INDEPENDIENTES	DESCRIPCIÓN
CD	Dummy = 1 si el artículo es de ciencia digital
T1	Probabilidad de que el artículo pertenezca al tópico 1, estandarizado con media 0 y desvío estándar 1
T2	Probabilidad de que el artículo pertenezca al tópico 2, estandarizado con media 0 y desvío estándar 1
T3	Probabilidad de que el artículo pertenezca al tópico 3, estandarizado con media 0 y desvío estándar 1
AR	Dummy = 1 si el artículo tiene algún autor afiliado a una institución argentina
CH	Dummy = 1 si el artículo tiene algún autor afiliado a una institución chilena
CO	Dummy = 1 si el artículo tiene algún autor afiliado a una institución colombiana
EDAD	Número de años desde la publicación del artículo hasta el 2019
AGRI	Dummy = 1 si el artículo es de Ciencias Biológicas y Agrícolas
EART	Dummy = 1 si el artículo es de Ciencias de la Tierra y Planetarias
ENVI	Dummy = 1 si el artículo es de Ciencias Ambientales
DECI	Dummy = 1 si el artículo es de Ciencias Decisionales

**Gráfico 1.** Evolución de la incidencia de la ciencia digital, según país\* (media móvil de los últimos cuatro años, en % de publicaciones), 1993-2017. \*


Top 10 = Principales 10 países en términos de artículos de ciencia digital: Estados Unidos, China, Reino Unido, Alemania, Japón, Francia, Canadá, Italia, India y Corea del Sur. Fuente: elaboración propia a partir de datos de Scopus.

toda la serie. La introducción de la ciencia digital se incrementa entre los extremos del periodo considerado a una tasa anual acumulativa de 4,1% en Argentina, 5,2% en Chile y 4,7% en Colombia, en tanto en el conjunto de economías desarrolladas seleccionadas este

ritmo, algo más lento, es de 3,8%. De manera que si bien la región achicó la brecha relativa, pues en los tres países el crecimiento es mayor al de las economías desarrolladas, ésta aún persiste: al final del período la incidencia de la ciencia digital es de 3,9% en

Argentina, 5,4% en Chile, 5,7% en Colombia y de 8% en los países tomados como referencia.

## 5.2. La incidencia de la ciencia digital según disciplinas

Los gráficos 2 a 5 presentan la incidencia de la ciencia digital según disciplinas en cada uno de los tres países. Muestran que la misma está creciendo en todas las disciplinas analizadas con la excepción, tal vez, de ciencias decisionales y con un ritmo que no ha sido el mismo en todos los casos. La tasa de crecimiento anual acumulativa en ciencias biológicas y agrícolas fue, en promedio para los tres países, de 6,8%, muy similar a la de ciencias de la tierra y planetarias (6,4%), pero bastante por encima de ciencias ambientales (3,6%). En tanto, la de ciencias decisionales, una disciplina donde la digitalización es de las más altas, el crecimiento de esta estuvo apenas por encima de cero: 0,4%.

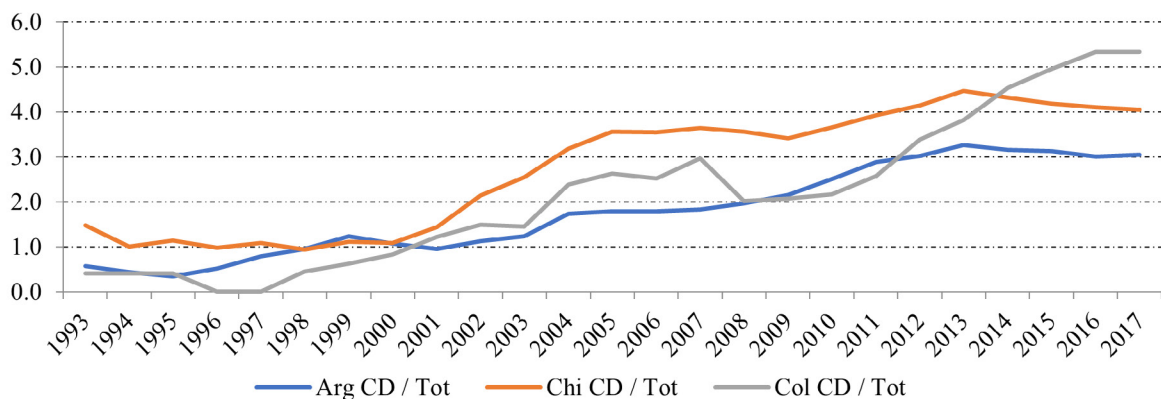
Si centramos la mirada en los países, observamos que los puntos de partida son distintos en cada una de las disciplinas y también lo son las tasas de crecimiento que evidencian. En Argentina la incidencia de la ciencia digital creció en promedio al 6,2% anual acumulativo, alcanzando el 7,2% en ciencias biológicas y agrícolas, el 6,9% en ciencias de la tierra y planetarias y el 2% tanto en ciencias ambientales como en ciencias decisionales. En Chile, por su parte, que exhibe una tasa promedio de 5,6%, el ritmo más alto de crecimiento aparece en ciencias de la tierra y planetarias, 7,2%, seguida por ciencias ambientales y ciencias biológicas y agrícolas con el 6,2 y el 4,3%, respectivamente. En este país, ciencias decisionales

exhibe una pequeña caída, que se explica en parte por el alto nivel de incidencia en el año de inicio de la serie (el 25% de los artículos fueron identificados como de ciencia digital). Finalmente, en Colombia, que muestra la tasa promedio más alta, 8%, la mayor tasa se observa en ciencias biológicas y agrícolas, 11,2%, seguida de ciencias decisionales, 6,0%, y luego por ciencias de la tierra y planetarias y ciencias ambientales, cada una con tasas de 2,6% y 2,4% respectivamente.

Los gráficos 2 a 5 también muestran que la incidencia varía según la disciplina. En 2017, por ejemplo, llegaba al 4,5% en ciencias ambientales con máximo de 5,1% en Chile. En ciencias biológicas y agrícolas la incidencia promedio fue de 3,9% con un máximo de 5,3% en Colombia. En ciencias de la tierra y planetarias fue de 7,7% alcanzando el 13% también en Colombia, en tanto en ciencias decisionales, como dijimos, se registran los valores más altos: una incidencia promedio de 17% y un máximo de 23% en Chile.

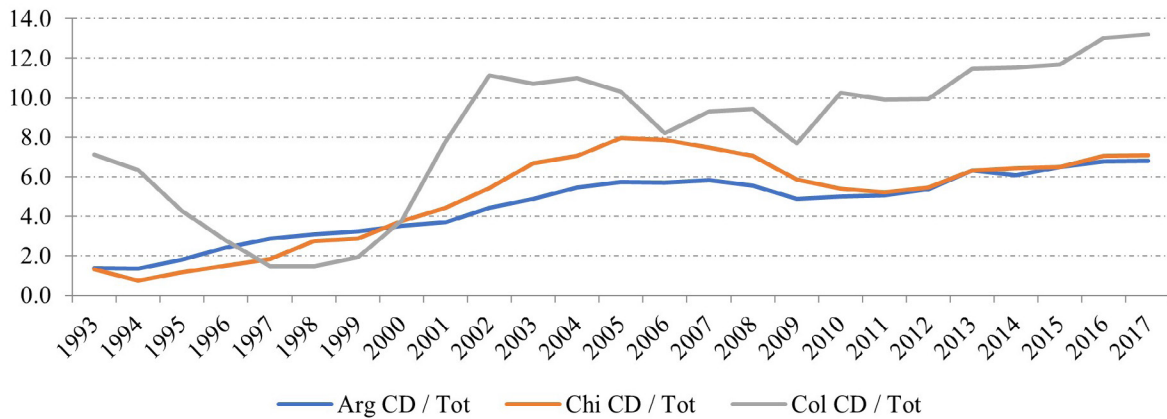
El gráfico 6 pone en perspectiva la situación de la región. Podemos observar que, en general, y en línea con el gráfico 1, la incidencia de la digitalización en estos países es menor que la mundial, con excepción de ciencias decisionales para Chile y ciencias de la tierra y planetarias para Colombia. Aunque esto podría sugerir que son áreas de amplio potencial en estos países, hay que leer los resultados con cautela porque se trata de disciplinas en las que la producción total de estos países es baja (el denominador del ratio es pequeño).

**Gráfico 2.** Evolución de la introducción de la ciencia digital en la ciencias biológicas y agrícolas, según país (media móvil de los últimos cuatro años, en % de publicaciones), 1993-2017.



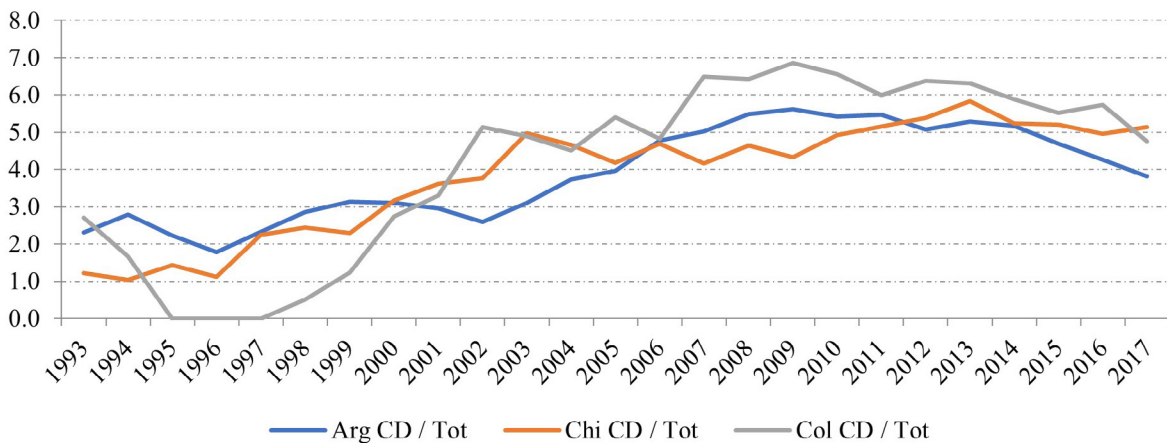
Fuente: elaboración propia a partir de datos de Scopus.

**Gráfico 3.** Evolución de la introducción de la ciencia digital en ciencias de la tierra y planetarias, según país (media móvil de los últimos cuatro años, en % de publicaciones), 1993-2017.



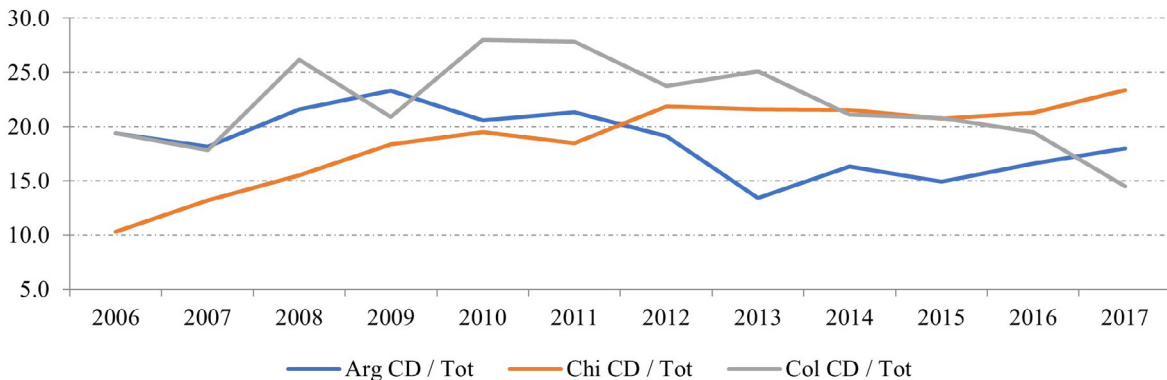
Fuente: elaboración propia a partir de datos de Scopus.

**Gráfico 4.** Evolución de la introducción de la ciencia digital en ciencias ambientales, según país (media móvil de los últimos cuatro años, en % de publicaciones), 1993-2017.



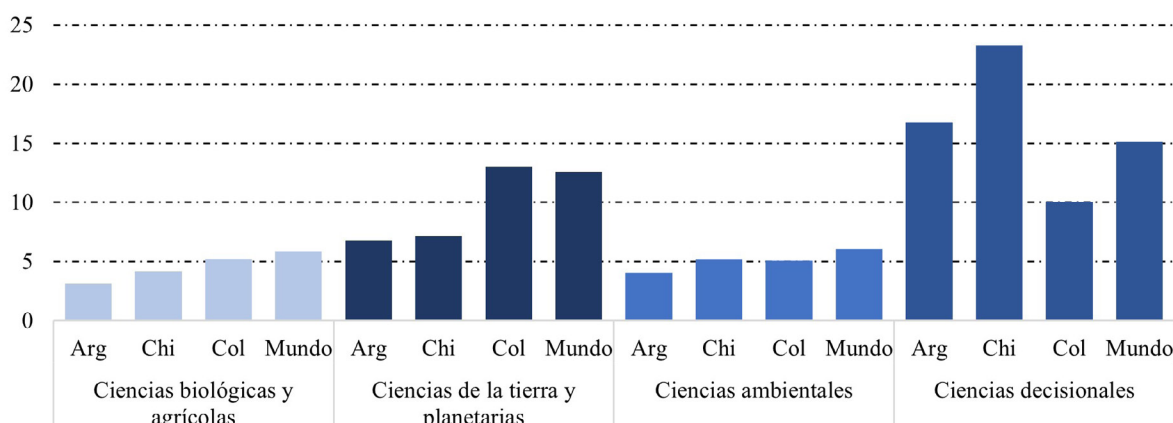
Fuente: elaboración propia a partir de datos de Scopus.

**Gráfico 5.** Evolución de la penetración de la ciencia digital en ciencias decisionales, según país (media móvil de los últimos cuatro años, en % de publicaciones), 2006-2017. Fuente: elaboración propia en base a datos de Scopus.



Fuente: elaboración propia a partir de datos de Scopus.

**Gráfico 6.** Comparación internacional de la penetración (%) de la ciencia digital. Promedio para el período 2013-2017.



Fuente: elaboración propia a partir de datos de *Scopus*.

### 5.3. Digitalización, impacto académico y redes

En esta sección presentamos hallazgos respecto de la relación entre digitalización de la ciencia, impacto académico y su potencial para ampliar redes.

Antes de pasar a las salidas de la regresión lineal, miremos las estadísticas descriptivas en el anexo 3 que nos permiten resaltar algunas características del conjunto de artículos con los que trabajamos y adelantar resultados esperados. La tabla A3 presenta para cada disciplina los valores promedio de las variables que van a integrar la regresión según el grupo al que pertenecen los artículos (los que hacen uso de herramientas digitales y lo que no) y también para el total y presenta el resultado de un *test* de diferencias de medias.

Observamos ahí que entre ambos grupos hay valores distintos en nuestras variables de interés, dentro algunas disciplinas. Por ejemplo, si centramos la mirada en ciencias biológicas y agrarias vemos que los artículos de ciencia digital tienen en promedio más citas por año, autorías de un mayor número de países y más coautorías. Estas diferencias no son significativas en el caso ciencias de la tierra y planetarias, con excepción del número de autorías de distintos países que colaboran, en donde la diferencia es levemente significativa (5%).

Si miramos los tópicos, cuyo contenido está descrito en la tabla A4, que presentamos en el anexo 4, observamos que los artículos de ciencia digital y los restantes presentan una distribución distinta que es necesario tener en cuenta al mirar los efectos de la digitalización. En el caso de ciencias biológicas y agrícola-

las, los artículos de ciencia digital están principalmente concentrados en ingeniería de alimentos, mientras que los restantes presentan una mayor aglomeración en biología sistémica. En el caso de ciencias de la tierra y planetarias, la mayor cantidad de artículos de ciencia digital entra en ciencias de la atmósfera y los restantes los encontramos en astronomía. Algo similar se observa en las otras dos disciplinas.

Ahora bien, la tabla también nos muestra diferencias en los patrones de publicación entre disciplinas, en algunos casos muy marcadas. Así, por ejemplo, la cantidad promedio de coautoría en ciencias de la tierra y planetarias es en promedio dos veces superior a la de ciencias biológicas y agrarias y ciencias ambientales y tres veces a la de ciencias decisionales. Asimismo, ciencias de la tierra y planetarias exhibe niveles más elevados de colaboración internacional que el resto: en promedio en un artículo aparecen afiliaciones institucionales de tres países en esta disciplina frente a un promedio algo más bajo de dos en las restantes.

Para poder evaluar el rol de la ciencia digital, manteniendo otros factores constantes, proponemos un análisis de regresión lineal. Las tablas 4 a 6 presentan los resultados de la estimación de los modelos [1] y [2]. Podemos ver ahí que, controlando por las variables que seleccionamos, la digitalización genera los beneficios que promete en términos de impacto y colaboración. En líneas generales, los artículos que utilizan prácticas o herramientas de ciencia digital tienen mayor impacto académico (tienen más citas), más colaboración (más coautorías) y más internacionalización (autorías de un mayor número de países).



La tabla 4 nos muestra que la digitalización se relaciona con un mayor impacto académico en todas las disciplinas. Un artículo que utiliza herramientas de ciencia digital tiene aproximadamente 0.8 citas más por año (o cuatro citas más cada cinco años) en ciencias biológicas y agrícolas, ciencias de la tierra y planetarias y en ciencias decisionales; y 0.4 citas más por año (o dos citas más cada cinco años) en ciencias ambientales.

En relación con la formación de redes (tablas 5 y 6), parece que la digitalización contribuye a colaborar con más países en ciencias de la tierra y planetarias, en ciencias biológicas y agrícolas y en ciencias decisionales. Solo en el primer caso, los artículos que usan herramientas digitales tienen a su vez más coautorías: en ciencias de la tierra y planetarias: tienen en promedio de tres coautorías más que los que no usan esas herramientas.

**Tabla 4.** Resultados de la estimación de los modelos de regresión lineal [1] y [2]. Variable dependiente: Citas por año

	Estimación de [1]			Estimación de [2]											
	est.	p.value		Cs. Biológicas y Agrícolas			Cs. Decisionales			Cs. de la Tierra y Planetarias			Cs. Ambientales		
				est.	p.value		est.	p.value		est.	p.value		est.	p.value	
(Intercepto)	3.476	0.000	***	1.187	0.000	***	3.032	0.000	***	5.039	0.000	***	2.557	0.000	***
CD	0.741	0.000	***	0.816	0.000	***	0.761	0.002	***	0.844	0.034	**	0.405	0.092	*
AGRI	-1.272	0.000	***	-	-	-	-	-	-	-	-	-	-	-	-
DECI	-1.442	0.000	***	-	-	-	-	-	-	-	-	-	-	-	-
EART	0.667	0.000	***	-	-	-	-	-	-	-	-	-	-	-	-
ENVI	0.414	0.000	***	-	-	-	-	-	-	-	-	-	-	-	-
T1_agri	0.304	0.000	***	0.274	0.000	***	-	-	-	-	-	-	-	-	-
T2_agri	0.171	0.001	***	0.194	0.000	***	-	-	-	-	-	-	-	-	-
T3_agri	-0.115	0.029	**	-0.165	0.000	***	-	-	-	-	-	-	-	-	-
T1_deci	-0.738	0.002	***	-	-	-	-0.550	0.000	***	-	-	-	-	-	-
T2_deci	-0.527	0.019	**	-	-	-	-0.587	0.000	***	-	-	-	-	-	-
T3_deci	-0.477	0.028	**	-	-	-	-0.540	0.000	***	-	-	-	-	-	-
T1_eart	-1.246	0.000	***	-	-	-	-	-	-	-1.168	0.000	***	-	-	-
T2_eart	-1.368	0.000	***	-	-	-	-	-	-	-1.443	0.000	***	-	-	-
T3_eart	-1.086	0.000	***	-	-	-	-	-	-	-0.892	0.000	***	-	-	-
T1_envi	0.109	0.286		-	-	-	-	-	-	-	-	-	0.559	0.000	***
T2_envi	0.296	0.002	***	-	-	-	-	-	-	-	-	-	0.413	0.000	***
T3_envi	-0.011	0.909		-	-	-	-	-	-	-	-	-	0.326	0.000	***
AR	0.453	0.024	**	1.299	0.000	***	-0.920	0.075	*	-0.453	0.394		1.056	0.001	***
CH	0.974	0.000	***	1.379	0.000	***	-0.181	0.723		0.954	0.078	*	1.346	0.000	***
CO	0.336	0.106		0.850	0.000	***	-1.457	0.006	***	-0.681	0.258		1.690	0.000	***
EDAD	-0.119	0.000	***	-0.045	0.000	***	-0.009	0.768		-0.255	0.000	***	-0.112	0.000	***

Fuente: elaboración propia.

**Tabla 5.** Resultados de la estimación de los modelos de regresión lineal [1] y [2]. Variable dependiente: Colaboración internacional

	Estimación de [1]			Estimación de [2]											
	est.	p.value		Cs. Biológicas y Agrícolas			Cs. Decisionales			Cs. de la Tierra y Planetarias			Cs. Ambientales		
				est.	p.value		est.	p.value		est.	p.value		est.	p.value	
(Intercepto)	0.150	0.004	***	-0.270	0.000	***	0.814	0.000	***	1.411	0.000	***	-0.872	0.000	***
CD	0.183	0.000	***	0.185	0.000	***	0.127	0.034	**	0.254	0.000	***	0.050	0.434	
AGRI	-0.252	0.000	***	-	-	-	-	-	-	-	-	-	-	-	-
DECI	-0.411	0.000	***	-	-	-	-	-	-	-	-	-	-	-	-
EART	0.961	0.000	***	-	-	-	-	-	-	-	-	-	-	-	-
ENVI	-0.001	0.958		-	-	-	-	-	-	-	-	-	-	-	-
T1_agri	0.096	0.000	***	0.081	0.000	***	-	-	-	-	-	-	-	-	-
T2_agri	0.072	0.000	***	0.087	0.000	***	-	-	-	-	-	-	-	-	-
T3_agri	0.130	0.000	***	0.120	0.000	***	-	-	-	-	-	-	-	-	-
T1_deci	-0.123	0.015	**	-	-	-	-0.078	0.013	**	-	-	-	-	-	-
T2_deci	0.133	0.006	***	-	-	-	-0.098	0.000	***	-	-	-	-	-	-
T3_deci	0.097	0.038	**	-	-	-	-0.027	0.306		-	-	-	-	-	-
T1_eart	-1.003	0.000	***	-	-	-	-	-	-	-0.976	0.000	***	-	-	-
T2_eart	-1.103	0.000	***	-	-	-	-	-	-	-1.056	0.000	***	-	-	-
T3_eart	-0.787	0.000	***	-	-	-	-	-	-	-0.712	0.000	***	-	-	-
T1_envi	-0.048	0.030	**	-	-	-	-	-	-	-	-	-	0.013	0.470	
T2_envi	0.122	0.000	***	-	-	-	-	-	-	-	-	-	0.114	0.000	***
T3_envi	-0.127	0.000	***	-	-	-	-	-	-	-	-	-	-0.099	0.000	***
AR	1.838	0.000	***	2.015	0.000	***	1.000	0.000	***	1.498	0.000	***	2.692	0.000	***
CH	2.239	0.000	***	2.262	0.000	***	1.128	0.000	***	2.271	0.000	***	2.987	0.000	***
CO	2.239	0.000	***	2.358	0.000	***	0.868	0.000	***	1.835	0.000	***	3.249	0.000	***
EDAD	-0.049	0.000	***	-0.036	0.000	***	-0.023	0.004	***	-0.087	0.000	***	-0.038	0.000	***

Fuente: elaboración propia.

**Tabla 6.** Resultados de la estimación de los modelos de regresión lineal [1] y [2]. Variable dependiente: Cantidad de autores

	Estimación de [1]			Estimación de [2]											
	est.	p.value		Cs. Biológicas y Agrícolas			Cs. Decisionales			Cs. de la Tierra y Planetarias			Cs. Ambientales		
				est.	p.value		est.	p.value		est.	p.value		est.	p.value	
(Intercepto)	-1.787	0.003	***	0.471	0.309		2.251	0.000	***	-0.819	0.557		-3.531	0.000	***
CD	1.450	0.000	***	0.498	0.165		0.095	0.644		3.186	0.001	***	0.224	0.401	
AGRI	-0.540	0.076	*	-	-	-	-	-	-	-	-	-	-	-	-
DECI	-2.706	0.000	***	-	-	-	-	-	-	-	-	-	-	-	-
EART	5.040	0.000	***	-	-	-	-	-	-	-	-	-	-	-	-
ENVI	-0.252	0.315		-	-	-	-	-	-	-	-	-	-	-	-
T1_agri	0.875	0.000	***	0.819	0.000	***	-	-	-	-	-	-	-	-	-
T2_agri	0.145	0.267		0.118	0.159		-	-	-	-	-	-	-	-	-
T3_agri	-0.117	0.375		-0.158	0.066	*	-	-	-	-	-	-	-	-	-
T1_deci	0.198	0.735		-	-	-	-0.072	0.501		-	-	-	-	-	-
T2_deci	0.188	0.737		-	-	-	-0.100	0.258		-	-	-	-	-	-
T3_deci	0.393	0.469		-	-	-	0.023	0.800		-	-	-	-	-	-
T1_eart	-5.161	0.000	***	-	-	-	-	-	-	-5.273	0.000	***	-	-	-
T2_eart	-5.590	0.000	***	-	-	-	-	-	-	-5.513	0.000	***	-	-	-
T3_eart	-4.493	0.000	***	-	-	-	-	-	-	-4.308	0.000	***	-	-	-
T1_envi	-0.515	0.044	**	-	-	-	-	-	-	-	-	-	-0.143	0.063	*
T2_envi	0.227	0.343		-	-	-	-	-	-	-	-	-	0.270	0.000	***
T3_envi	0.017	0.941		-	-	-	-	-	-	-	-	-	0.290	0.000	***
AR	8.282	0.000	***	5.190	0.000	***	1.674	0.000	***	12.875	0.000	***	8.828	0.000	***
CH	8.473	0.000	***	5.311	0.000	***	1.375	0.002	***	13.722	0.000	***	8.864	0.000	***
CO	8.465	0.000	***	5.221	0.000	***	1.329	0.003	***	13.310	0.000	***	9.497	0.000	***
EDAD	-0.286	0.000	***	-0.184	0.000	***	-0.120	0.000	***	-0.567	0.000	***	-0.154	0.000	***

Fuente: elaboración propia a partir de datos de *Scopus*.

En relación con los tópicos, en ciencias biológicas y agrícolas, ciencias de las plantas es el que suele tener más citas y más coautorías, mientras Biología sistémica es donde hay mayor internacionalización. En ciencias de la tierra y planetarias es astronomía el tópico que presenta mejor desempeño en todas estas variables: citas por año, coautorías e internacionalización. En ciencias ambientales es más variado: energías alternativas tiene más citas, biología de la conservación más internacionalización, y biología de la conservación y eco-toxicidad más coautorías. Finalmente, en ciencias decisionales, programación tiene más citas por año, mientras estadística es el que presenta mayor internacionalización.

## 6. DISCUSIÓN Y CONCLUSIONES

En este trabajo, realizamos un análisis sobre el avance de la ciencia digital en tres países de América Latina, Argentina, Chile y Colombia, analizando la intensidad de uso de prácticas digitales en cuatro disciplinas seleccionadas: ciencias biológicas y agrícolas, ciencias de la tierra y planetarias, ciencias ambientales y ciencias decisionales. Buscamos países que tuvieran una producción importante a nivel internacional y que hubiera diversidad de áreas de conocimiento, sabiendo que el potencial de la digitalización puede ser distinto. Asimismo, analizamos el impacto académico de la digitalización y sus efectos en la formación de redes de colaboración.

Encontramos que en estos tres países, Argentina, Chile y Colombia, la digitalización ha ido creciendo a una tasa anual acumulativa de 4,8% durante los últimos veinticinco años; mientras que para los países que lideran a nivel mundial la producción científica esta cifra ha sido de 3,8%. En 2017 la digitalización giró alrededor del 5% en Argentina, Chile y Colombia mientras que alcanzó el 8% en las economías líderes, con una incidencia 1,6 veces mayor en éstas últimas. En 1993 estos valores eran 1,6 y 3,3%, respectivamente. Hemos visto también, en línea con la literatura mencionada (Holmberg y Thelwall, 2014; Tsatsou, 2016; Galindo-Rueda, 2020), que la incidencia de la digitalización varía por disciplina. En ciencias ambientales y ciencias biológicas y agrarias la digitalización es parecida a los promedios nacionales de 5%, pero en ciencias de la tierra y planetarias alcanzó un máximo de 12% en Colombia, mientras que en Argentina y Chile se mantuvo en torno al 7%. En ciencias decisionales la digitalización llegó al 25% en Chile, Argentina llegó al 18% y Colombia al 15%. La mayor penetración en ciencias de la tierra y planetarias en relación con ciencias ambientales y ciencias biológicas está en concordancia con los resultados presentados por Galindo-Rueda (2020).

Las diferencias en la incidencia posiblemente se expliquen por las maneras distintas que tienen las disciplinas de producir conocimiento, tal y como concluye

Tsatsou (2016), y por diferencias en los perfiles de especialización de cada país (tabla A1, que figura en el anexo 1). Pero también por diferencias en las infraestructuras y equipamiento al alcance de los grupos de investigación. En general, hay excepciones, la incidencia de la digitalización en la región es menor que la mundial, lo que puede sugerir que todavía son prácticas incipientes para las cuales queda espacio para avanzar.

Lo que resulta muy interesante es que aún en estas etapas incipientes, la digitalización genera los beneficios que promete en términos de impacto académico y colaboración, tal como se desprende de una variedad de documentos de política pública (Hey y Trefethen, 2002; Atkins *et al.*, 2003; Sargent, 2006; Comisión Europea, 2013). En nuestro análisis bibliométrico encontramos que, en estas disciplinas y en estos países, la digitalización ha logrado aumentar la visibilidad de la investigación, potencialmente su calidad y, así, los retornos sociales de los recursos invertidos. También ha ampliado las redes (más coautorías de un número mayor de países), propiciando un mejor aprovechamiento de la inteligencia colectiva y la integración de los problemas de investigación locales a una agenda global. Esto puede haber favorecido el acceso a recursos de investigación y puede redundar en la ampliación de líneas de investigación ya existentes y en la creación de nuevas preguntas de investigación.

Estos resultados sugieren que un mayor compromiso con la digitalización en estos países podría mejorar no sólo el impacto académico de la producción de conocimiento local, sino también la participación en redes globales de conocimiento. No sabemos si esto es generalizable a otras disciplinas y otros países de la región, pero sin duda estos resultados brindan elementos de apoyo para promover la digitalización desde las políticas públicas. Como señalamos en el texto, en todos los países existen herramientas de política pública que favorecen prácticas específicas de digita-

lización y ha habido un esfuerzo en la construcción de infraestructura. Sin embargo, en ninguno de los tres casos existe hasta la fecha una política integral de promoción de la ciencia digital (Arza, *et al.*, 2019) como sí existe en algunas economías desarrolladas y como estos resultados parecieran justificar.

Una limitación que presenta este estudio es que está basado en una única base de datos bibliográfica interdisciplinar, *Scopus*, que presenta una cobertura deficiente de revistas académicas de la región. En este sentido, se presenta como una necesidad de investigación futura la incorporación de otras bases de datos bibliográficas como *SciELO* para validar los resultados. Nuestro estudio sugiere que la digitalización trae beneficios en términos de impacto académicos y creación de redes, aunque para validar esta causalidad se necesitaría un tratamiento econométrico más robusto que no descartamos realizar en estudios posteriores.

Este trabajo ofrece una manera de medir la incidencia de la digitalización de la ciencia para analizar y monitorear su desarrollo a partir de bases de datos bibliográficas. Se basa en la idea de que es posible identificar el empleo de herramientas digitales en el proceso de producción de conocimiento a través de los metadatos de los artículos científicos (resúmenes y palabras claves, principalmente) y a partir de la definición de un conjunto de palabras que designan a estas herramientas. Constituye, así, un punto de partida a partir del cual se profundizar en el análisis sistemático de la incidencia y los efectos de la ciencia digital tanto en la región como fuera de ella, refinando los criterios de búsqueda o extendiendo el análisis a otros países, o regiones, y disciplinas.

## AGRADECIMIENTOS

Este artículo se basa parcialmente en el informe *La Ciencia Digital en América Latina* elaborado en 2019 por los autores junto a otros colegas a petición del Banco Interamericano de Desarrollo (BID).

## REFERENCIAS

Alonso-Gamboa, José Octavio y Russell, Jane M. (2012). Latin American scholarly journal databases: a look back to the way forward. *Aslib Proceedings*, 64 (1): 32-45.

Andronico, Giuseppe; Ardizzone, Valeria; Barbera, Roberto; Becker, Bruce; Bruno, Riccardo; Calanducci, Antonio; Carvalho, Diego; Ciuffo, Leandro; Fargetta, Marco; Giorgio, Emidio; La Rocca, Giuseppe; Masoni, Alberto; Paganoni,

Marco; Ruggieri, Federico y Scardaci, Diego (2011). e-Infrastructures for e-Science: A Global View. *Journal of Grid Computing*, 9 (2): 155-184.

Arcila, Carlos; Piñuel, José-Luis y Calderín, Mabel (2013). La e-investigación de la Comunicación: actitudes, herramientas y prácticas en investigadores iberoamericanos. *Comunicar: Revista Científica de Comunicación y Educación*, 20 (40): 111-118.

Arza, Valeria; Fressoli, Mariano; Actis, Guillermina y del Castillo, Martín (2019). *La Ciencia Digital en América Latina. Manuscrito*. Washington D.C.: Banco Interamericano Desarrollo (BID).

Atkins, Daniel E. *et al.*, (2003). Revolutionizing science and engineering through cyberinfrastructure. *Report of the National Science Foundation blue-ribbon advisory panel on cyberinfrastructure*. Virginia: National Science Foundation.

- Disponible en: <https://www.nsf.gov/cise/sci/reports/atkins.pdf> [Consultado en junio de 2019].
- Balassa, Bela (1965). Trade Liberalisation and "Revealed" Comparative Advantage. *The Manchester School*, 33 (2): 99-123.
- Chavarro Bohórquez, Diego Andrés (2016). Universalism and Particularism: Explaining the Emergence and Growth of Regional Journal Indexing Systems. *SPRU – Science and Technology Policy Research Unit*. Brighton: University of Sussex.
- Comisión Europea (2013). Digital science in Horizon 2020: Concept paper of the Digital Science vision, and its integration in the Horizon 2020 programme: Disponible en: <https://ec.europa.eu/digital-single-market/en/news/digital-science-horizon-2020> [Consultado en junio de 2019].
- Confraria, Hugo y Vargas, Fernando (2017). Scientific systems in Latin America: performance, networks, and collaborations with industry. *The Journal of Technology Transfer*, 44: 874–915.
- Dougherty, Deborah y Dunne, Danielle D. (2012). Digital science and knowledge boundaries in complex innovation. *Organization Science*, 23 (5): 1467-1484.
- Dutton, William H. y Meyer, Erick T. (2008). E-social science as an experience technology: distance from, and attitudes toward, e-research. *4th International Conference on e-Social Science*, Manchester (UK).
- Emiliozzi, Sergio; Lemarchand, Guillermo A. y Gordon, Ariel (2009). Inventario de instrumentos y modelos de políticas de ciencia, tecnología e Innovación en América Latina y el Caribe. Working Paper 9. Washington: Banco Interamericano de Desarrollo (BID) y REDES.
- Galindo-Rueda, Fernando (2020). How are science, technology and innovation going digital? The statistical evidence. En: OECD (eds.). *The Digitalisation of Science, Technology and Innovation: Key Developments and Policies*. Paris: OECD Publishing.
- Grudin, Jonathan (1994). Computer-supported cooperative work: History and focus. *Computer*, 27 (5): 19-26.
- Hey, Tony; Tansley, Stewart y Tolle, Kristin M (eds.) (2009). *The Fourth Paradigm: Data-Intensive Scientific Discovery*. Redmond: Microsoft Research.
- Hey, Tony y Trefethen, Anne (2002). The UK e-science core programme and the grid. *Future Generation Computer Systems*, 18 (8): 1017-1031.
- Hey, Tony y Trefethen, Anne (2008). E-Science, Cyberinfrastructure, and Scholarly Communication. En: Gary M Olson, Ann Zimmerman y Nathan Bos (eds.). *Scientific Collaboration on the Internet: The MIT Press*, pp. 15 - 31.
- Holmberg, Kim y Thelwall, Mike (2014). Disciplinary differences in Twitter scholarly communication. *Scientometrics*, 101: 1027-1042.
- Kennedy, Helen; Moss, Giles; Birchall, Christopher y Moshonas, Stylianos (2015). Balancing the potential and problems of digital methods through action research: Methodological reflections. *Information, Communication & Society*, 18 (2): 172-186.
- Morán-Mariños, Cristian; Pacheco-Mendoza, Joesmel; Metcalf, Tatiana; De la Cruz Ramirez, Walter y Alva-Díaz, Carlos (2020). Collaborative scientific production of epilepsy in Latin America from 1989 to 2018: A bibliometric analysis. *Heliyon*, 6 (5): e05493.
- Nentwich, Michael (2003). *Cyberscience: Research in the Age of the Internet*. Vienna: Austrian Academy of Sciences Press.
- Olson, Gary M. y Olson, Judith S. (2012). Collaboration technologies. En: Julie A Jacko (eds.). *Human Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications*. Boca Raton: CRC press, pp. 549-564.
- Olson, Gary M.; Zimmerman, Ann y Bos, Nathan (eds.) (2008). *Scientific Collaboration on the Internet*. Cambridge: The MIT Press.
- Pacheco, Roberto C.S.; Nascimento, Everton R. y Weber, Rosina O. (2018). Digital science: cyberinfrastructure, e-Science and citizen science. En: Klaus North, Ronald Maier y Oliver Haas (eds.). *Knowledge Management in Digital Change*. Cham: Springer, pp. 377-388.
- Packer, Abel Laerte y Meneghini, Rogerio (2007). Learning to communicate science in developing countries. *Interciencia*, 32 (9): 643-647.
- Persson, Olle; Glänzel, Wolfgang y Danell, Rickard (2004). Inflationary bibliometric values: The role of scientific collaboration and the need for relative indicators in evaluative studies. *Scientometrics*, 60 (3): 421-432.
- Rogers, Richard (2013). *Digital methods*. Cambridge, MA: MIT press.
- Rogers, Richard (2015). Digital methods for web research. En: R A Scott y S M Kosslyn (eds.). *Emerging Trends in the Social and Behavioral Sciences*. London: John Wiley & Sons, pp. 1-22.
- Ruiz-Patiño, Alejandro *et al.*, (2020). Scientific publications in cancer: in Latin America, strong scientific networks increase productivity (the TENJIN study). *Journal of Clinical Epidemiology*, 126: 1-8.
- Sagasti, Francisco y Araoz, Alberto (1976). *Science and technology policy implementation in less - developed countries: methodological guidelines for the STPI project*. Ottawa, Canada: IDRC.
- Sargent, Mike (2006). An Australian e-research strategy and implementation framework. *Final Report of the e-Research Coordinating Committee*. Canberra, Australia: Department of Education, Employment and Workplace Relations. Disponible en: <https://apo.org.au/node/15777> [Consultado en junio de 2019].
- Schotten, Michiel; el Aisati, M'hamed; Meester, Wim J.N.; Steinginga, Susanne y Ross, Cameron A. (2017). A Brief History of Scopus: The World's Largest Abstract and Citation Database of Scientific Literature. En: Francisco J. Cantu-Ortiz (eds.). *Research Analytics: Boosting University Productivity and Competitiveness through Scientometrics*. Boca Raton, Florida: CRC Press.
- Tsatsou, Panayiota (2016). Digital technologies in the research process: Lessons from the digital research community in the UK. *Computers in Human Behavior*, 61 (August): 597-608.
- Wulf, William A. (1993). The laboratory opportunity. *Science*, 261 (5123): 854–855.



**ANEXO 1**

**Tabla A1.** Índice de ventajas comparativas reveladas (VCR) por disciplina y país en términos de publicaciones en SCOPUS 2013-2017\*.

Disciplinas	Arg	Chi	Col
Artes y Humanidades	0.227	0.510	0.116
Bioquímica, Genética y Biología Molecular	0.234	-0.171	-0.290
Ciencia de los materiales	-0.394	-0.560	-0.426
Ciencias agrícolas y biológicas	1.901	0.850	0.850
Ciencias ambientales	0.468	0.155	-0.009
Ciencias de la Computación	-0.488	-0.273	-0.007
Ciencias de la Tierra y Planetarias	0.860	1.872	-0.075
Ciencias decisionales	-0.509	0.278	0.725
Ciencias Sociales	-0.034	0.392	0.205
Economía, Econometría y Finanzas	-0.402	-0.025	0.211
Energía	-0.481	-0.450	0.136
Enfermería	-0.574	0.096	-0.162
Farmacología, Toxicología y Farmacia	-0.112	-0.451	-0.152
Física y Astronomía	0.012	0.500	-0.049
Ingeniería	-0.569	-0.444	-0.137
Ingeniería química	0.021	-0.414	-0.035
Inmunología y microbiología	1.045	-0.022	0.326
Matemáticas	-0.217	0.142	0.001
Medicina	-0.087	-0.175	-0.048
Multidisciplinario	-0.269	-0.214	-0.025
Negocios, Gestión y Contabilidad	-0.643	-0.288	0.445
Neurociencias	0.138	-0.090	-0.432
Odontología	-0.631	0.667	-0.028
Profesiones de la salud	-0.503	-0.234	-0.042
Psicología	-0.223	0.215	-0.075
Química	0.043	-0.313	-0.378
Veterinaria	1.283	0.181	0.870

Fuente: Elaboración propia a partir de datos de SCOPUS.

\*Al valor que arroja el índice le restamos una unidad. De esta manera, un valor positivo indica que el país posee VCR y un valor negativo que no posee VCR.



## ANEXO 2

**Tabla A2.** Palabras clave utilizadas para identificar digitalización por disciplina

<b>Disciplina</b>	<b>Palabras clave</b>
Todas las disciplinas	Learning Systems; Machine Learning; Big Data; Artificial Intelligence; Data Mining; Learning Algorithms; Algorith*; Virtual Reality; Digital Storage; Support Vector Machines; Data Handling; Feature Extraction; Social Networking (online); Machine Learning Techniques; Text Mining; Crowdsourcing; Support Vector Machine; Natural Language Processing Systems; Image Processing; Cloud Computing; Deep Learning; Pattern Recognition; Information Retrieval; Distributed Computer Systems; Machine-learning; Computer Simulation; Machine Learning Methods; Clustering Algorithms; Text Processing.
Ciencias biológicas y agrícolas	Machine Learning; Algorith*; Artificial Intelligence; Support Vector Machine; Data Mining; Computational Biology; Big Data; Artificial Neural Network; Software; Remote Sensing; Information Processing; Image Processing; Computer Simulation; Support Vector Machines; Learning Algorithms; Image Analysis; Computer Program; Bioinformatics
Ciencias de la tierra y planetarias	Learning Systems; Machine Learning; Remote Sensing; Artificial Intelligence; Big Data; Learning Algorithms; Algorithm; Neural Networks; Support Vector Machines; Data Mining; Digital Storage; Image Analysis; Data Handling; Artificial Neural Network; Image Processing; Crowdsourcing; Support Vector Machine; Virtual Reality; Digital Elevation Model; Algorith*; Machine Learning Techniques; Optimization; Computer Simulation
Ciencias ambientales	Machine Learning; Big Data; Artificial Intelligence; Algorith*; Artificial Neural Network; Data Mining; Remote Sensing; Learning Algorithms; Support Vector Machine; Digital Storage; Neural Networks; Computer Simulation; Support Vector Machines; Crowdsourcing OR Machine Learning Techniques
Ciencias decisionales	Big Data; Data Mining; Artificial Intelligence; Machine Learning; Data Handling; Learning Algorithms; Social Networking (online); Cloud Computing; Digital Storage; Text Mining; Semantics; Neural Networks; Natural Language Processing; Algorithms; Distributed Computer Systems; Support Vector Machines; Crowdsourcing; Feature Extraction; Machine Learning Techniques; Clustering Algorithms; Text Processing; Deep Learning; Ubiquitous Computing; Data Communication Systems; Knowledge Based Systems

Fuente: elaboración propia.

**ANEXO 3**
**Tabla A3.** Estadística descriptiva de las variables incluidas en el análisis de regresión

	Cs. Biológicas y Agrícolas				Cs. Decisionales				Cs. de la Tierra y Planetarias				Cs. Ambientales			
	Promedio			t.test	Promedio			t.test	Promedio			t.test	Promedio			t.test
	CD=1	CD=0	Total	(p-value)	CD=1	CD=0	Total	(p-value)	CD=1	CD=0	Total	(p-value)	CD=1	CD=0	Total	(p-value)
TCxAño	3.0510	2.2180	2.2480	0.0000***	3.5860	2.0340	2.3070	0.0000***	4.5170	4.0510	4.0840	0.2350	3.8920	3.3220	3.3510	0.0030***
colab_int	1.9110	1.7540	1.7600	0.0000***	1.9460	1.7470	1.7820	0.0000***	2.8310	2.9970	2.9850	0.0640*	2.0080	1.9120	1.9170	0.1050
AU_cantidad	5.4140	4.9030	4.9220	0.0150**	3.3310	3.2470	3.2620	0.5500	11.2700	10.3420	10.4070	0.3330	4.8400	4.9540	4.9480	0.6030
CD	1.0000	0.0000	0.0370	-	1.0000	0.0000	0.1760	-	1.0000	0.0000	0.0700	-	1.0000	0.0000	0.0500	-
T1	0.1490	0.1930	0.1910	0.0000***	0.0810	0.4080	0.3500	0.0000***	0.6940	0.2420	0.2740	0.0000***	0.5630	0.2930	0.3070	0.0000***
T2	0.2170	0.2370	0.2360	0.0090***	0.2160	0.2180	0.2180	0.8830	0.0470	0.2380	0.2240	0.0000***	0.2190	0.2670	0.2650	0.0000***
T3	0.1820	0.3460	0.3400	0.0000***	0.0930	0.2010	0.1820	0.0000***	0.0680	0.1420	0.1370	0.0000***	0.0480	0.2200	0.2110	0.0000***
T4	0.4520	0.2250	0.2330	0.0000***	0.6100	0.1730	0.2500	0.0000***	0.1910	0.3780	0.3640	0.0000***	0.1710	0.2190	0.2170	0.0000***
AR	0.4540	0.5360	0.5330	0.0000***	0.2250	0.2190	0.2200	0.8190	0.3640	0.4160	0.4130	0.0000***	0.4570	0.4970	0.4950	0.0290**
CH	0.3160	0.2760	0.2780	0.0010***	0.5510	0.4110	0.4350	0.0000***	0.4870	0.5240	0.5210	0.0070***	0.3580	0.3330	0.3340	0.1600
CO	0.2560	0.2110	0.2130	0.0000***	0.2470	0.3980	0.3710	0.0000***	0.1720	0.0950	0.1000	0.0000***	0.2090	0.1950	0.1960	0.3600
EDAD	4.7200	5.0070	4.9960	0.0000***	4.2910	3.7910	3.8790	0.0010***	4.5270	4.9550	4.9250	0.0000***	4.9960	4.7230	4.7370	0.0090***

**ANEXO 4**

Con el fin de analizar y/o evaluar el contenido de los tópicos resultantes, buscamos identificar el campo de investigación al que refiere cada uno de ellos en cada disciplina. Para eso nos valimos de las palabras claves que se presentan con mayor frecuencia entre los artículos que tienen una probabilidad mayor a 65% de pertenecer a un tópico<sup>17</sup>. Con estas palabras claves, tras consultas con expertos del campo en cuestión y la revisión de los títulos y *abstracts*, interpretamos que cada uno de ellos refiere a las siguientes subdisciplinas:

**Tabla A4.** Interpretación del campo de investigación relevante asociado a cada tópico en cada disciplina

	Ciencias Biológicas y Agrícolas	Ciencias de la Tierra y Planetarias	Ciencias Ambientales	Ciencias Decisionales
<b>Tópico 1</b>	Ciencia de las plantas	Ciencias de la atmósfera	Energías alternativas	Economía
<b>Tópico 2</b>	Cambio climático en agroecosistemas	Paleontología	Biología de la conservación	Estadística
<b>Tópico 3</b>	Biología Sistémica	Geología	Ecotoxicidad	Modelización
<b>Tópico 4</b>	Ingeniería de alimentos	Astronomía	Cambio climático	Programación

Fuente: elaboración propia.

17 De esta forma, el conjunto de artículos incluidos por disciplina representa el 73% del total de artículos de ciencias biológicas y agrícolas, el 88% de ciencias de la tierra y planetarias, el 74% de ciencias ambientales y el 84% de ciencias decisionales.