

Título del artículo:

**Learning analytics en relación con el acceso abierto a datos de investigación en Perú:
Una comparación interdisciplinaria**

Autores:

Katarzyna Biernacka

Humboldt-Universität zu Berlin

ORCID 0000-0002-6363-0064

katarzyna.biernacka@hu-berlin.de

Investigadora y candidata al doctorado en Educación Informática en el Departamento de Informática de la Humboldt-Universität zu Berlin (HU) en Alemania con enfoque en los retos de la publicación de datos de investigación en relación con la integridad de la investigación y la protección de datos. Copresidenta del Grupo de Interés de Educación y Formación sobre la Gestión de Datos de Investigación del Research Data Alliance (RDA) y miembro del grupo directivo del grupo de trabajo de Datos de Investigación de DINI/nestor en Alemania.

Libio Huaroto

Universidad Peruana de Ciencias Aplicadas

ORCID 0000-0003-4667-0993

llibio.huaroto@upc.pe

Miembro del Board of Directors of Networked Digital Library of Theses and Dissertations (NDLTD). Egresado de la maestría en docencia e investigación en la Universidad Peruana Cayetano Heredia (UPCH). Diplomado en Gestión Internacional de la Tecnología y de la Innovación en Universidad Peruana Cayetano Heredia (UPCH). Actualmente responsable del Repositorio Académico de la Universidad Peruana de Ciencias Aplicadas (UPC), desarrolla e implementa proyectos de repositorios digitales y datos y plataformas de publicaciones científicas.

Learning analytics en relación con el acceso abierto a datos de investigación en Perú: Una comparación interdisciplinaria

Katarzyna Biernacka ¹ and Libio Huaroto ²

¹ *Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany*

² *Universidad Peruana de Ciencias Aplicadas, Prolongación Primavera 2390, Santiago de Sucre, Perú*

Resumen

El objetivo del estudio fue investigar las percepciones acerca del learning analytics en investigadores peruanos sobre las barreras para la publicación de los datos de investigación. Se realizó una revisión de la legislación más importante. Se utilizó las entrevistas como método de investigación, el foco del trabajo se centró en el conflicto entre la publicación de los datos de investigación y la protección de datos personales. Los resultados muestran rangos de factores individuales que influyen en el comportamiento de los investigadores en relación con la publicación de los datos de investigación, enfatizando las barreras relacionadas a la protección de datos en diferentes disciplinas.

Palabras clave

Publicación de datos, compartir datos, datos abiertos de investigación, integridad científica, datos de investigación, ciencia abierta.

Abstract

The aim of this paper is to investigate the perceptions of learning analytics researchers in Peru about the barriers to publication of their research data. A review of the relevant legislation was done. Semi-structured interviews were used as a research method, the focus being on the presumed conflict between the publication of research data and the protection of personal data. The results show a range of individual factors that influence the behaviour of scientists in relation to the publication of research data, emphasizing the barriers related to data protection in different disciplines.

Keywords

Data publication, data sharing, open research data, research integrity, research data, open science

Resumo

O objetivo do estudo foi investigar as percepções da aprendizagem analítica em pesquisadores peruanos sobre as barreiras para a publicação de dados de pesquisa. O foco do trabalho é o conflito entre a publicação dos dados da pesquisa e a proteção dos dados pessoais. Foi realizada uma revisão da legislação mais importante. Entrevistas foram utilizadas como técnica de pesquisa, Os resultados mostram uma gama de fatores individuais que influenciam o comportamento dos pesquisadores em relação à publicação de dados de pesquisa, enfatizando as barreiras relacionadas à proteção de dados em diferentes disciplinas.

Palavras-chave

Publicação de dados, compartilhamento de dados, dados de pesquisa abertos, integridade científica, dados de pesquisa, ciência aberta

1. Introducción

El Perú no ha estado alejado de los movimientos de acceso abierto a nivel global, la consolidación de un trabajo continuo por más de una década se reflejó con la promulgación de la Ley 30035 “Ley que regula el Repositorio Nacional Digital de Ciencia, Tecnología e Innovación de Acceso Abierto” (Congreso de la República del Perú, 2013).

El acceso abierto tiene su comienzo a finales del 1999 en la Universidad Nacional Mayor de San Marcos (UNMSM), con el trabajo de digitalización de las tesis, libros y revistas que se publicaron inicialmente en formato HTML y PDF, para luego aplicar estándares de interoperabilidad. Al inicio, el trabajo se realizaba en un ambiente heterogéneo, con diferentes metodologías y rutas, equipos de cómputo y recursos humanos variados; caracterizado en su forma de objetivos-fronterizo, esto es, instrumentos suficientemente estándares para ordenar conductas, lo mismo que suficientemente flexibles para adaptarse a las necesidades y limitaciones locales.

Entre noviembre del 2003 y junio del 2004 se pone en marcha el Portal de Tesis Digitales denominado “Cybertesis Perú”, que comprende el uso de estándares de procesamiento de información (Dublin Core), protocolos de interoperabilidad (Open Access Initiative-Protocol Metadatos Harvesting-OAI-PMH), ETD-MS, entre otros. En ese mismo año, el CONCYTEC se incorpora como Miembro Coordinador Nacional de Latindex (Sistema Regional de Información en Línea para Revistas Científicas de América Latina, el Caribe, España y Portugal). El 25 de marzo de 2004 se inicia el proyecto SciELO Perú, modelo para la publicación electrónica cooperativa de revistas científicas en Internet, auspiciado por Consejo

Nacional de Ciencia, Tecnología e Innovación Tecnológica (CONCYTEC), Universidad Nacional Mayor de San Marcos (UNMSM) y la Organización Panamericana de la Salud OPS Perú). En el 2010, la Pontificia Universidad Católica del Perú (PUCP) inició su programa de tesis digital mediante la plataforma DSpace. En el ámbito de revistas, la UNMSM publicó en 2010 el Portal de Revistas de Investigación, a través de la plataforma Open Journal System (OJS).

En relación a los repositorios de datos de investigación, a pesar de contar con un marco regulatorio, este tipo de proyecto no ha mostrado el mismo grado de desarrollo como en Argentina, Chile, Colombia y México. En Perú, la Guía ALICIA 2.0.1¹ recomienda el uso del software Dataverse para gestión de datos, en ese sentido, algunas instituciones presentan experiencias que evidencian esfuerzos en favor de los datos abiertos. En esta misma línea, en mayo 2022 se publicó la Directiva N° 003-2022-CONCYTEC-P² que regula el funcionamiento de #PeruCRIS³, proyecto gestionado por el CONCYTEC y que busca operar la Red Nacional de Información en Ciencia, Tecnología e Innovación (CTI) a fin de consolidar la información científica y académica en Perú. En relación al trabajo en learning analytics, el desarrollo de investigaciones y experiencias ha sido muy poca.

2. Situación Problemática en el Perú

Los diversos proyectos e iniciativas de acceso abierto en Perú se incrementaron progresivamente durante la década del 2000, motivados principalmente por visibilizar la producción académica institucional, no existía un marco legal nacional que obligara a las instituciones a desarrollar actividades o proyectos similares. En la actualidad, a pesar de la existencia del marco normativo sobre la gestión de datos abiertos, aspectos como políticas de procesamiento de datos de investigación, directivas sobre gestión de datos y ciclo de vida; así como también plataforma tecnológica siguen sin tener la misma relevancia que los repositorios institucionales en la vida académica y científica del Perú.

2.1. Legislación y Acceso Abierto a Datos de Investigación en el Perú

Latinoamérica y en especial Perú se fortalecieron en el proceso de difusión y promoción de la información científica en acceso abierto como consecuencia de las normativas que fueron promulgadas por algunos gobiernos. El 5 de junio de 2013 se publicó la Ley 30035 “Ley que regula el Repositorio Nacional Digital de Ciencia, Tecnología e Innovación de Acceso Abierto”, su reglamento fue aprobado el 24 de enero de 2015 (Congreso de la

¹ <http://hdl.handle.net/20.500.12390/2231>

² <https://perucris.concytec.gob.pe/documentacion>

³ <https://www.perucris.pe/>

República del Perú, 2013). Esta normativa, la primera ley a nivel global, define aspectos importantes sobre el acceso abierto. Los documentos producidos con fondos del Estado se publican en acceso abierto, compartiendo los derechos patrimoniales con el autor o autores. De igual forma, se crea el Repositorio Nacional Digital en Ciencia, Tecnología, e Innovación de Acceso Abierto, llamado "ALICIA",⁴ gestionado por el CONCYTEC. Así mismo, se crea la Red de Repositorio Peruanos RENARE.

En el 2016 CONCYTEC publica la Directiva N° 004-2016-CONCYTEC-DEGC, que regula la gestión de ALICIA (08 de junio de 2016)⁵. Entre otros aspectos, la norma formaliza el uso de *softwares* para la gestión de los repositorios institucionales (DSpace) y datos de investigación (Dkan) (CONCYTEC, 2016).

En el sector educación, el 6 de septiembre del 2016 se publicó la Resolución del Consejo Directivo N° 033-2016-SUNEDU/CD que aprueba el Reglamento del Registro Nacional de Trabajos de Investigación para optar por grados académicos y títulos profesionales—RENATI.⁶ Esta norma permitió la creación del Portal "RENATI" que gestiona y difunde las tesis y trabajos de investigación (Superintendencia Nacional de Educación Superior Universitaria, 2016).

En relación con los datos de investigación, la Ley 30035 incluye aspectos sobre la gestión de datos abiertos de investigación (datos procesados), define la interoperabilidad, ámbito de aplicación, entre otros aspectos. El marco normativo en Perú guarda relación con otras políticas de países en la región, por ejemplo, Chile lanza su Propuesta de Política de Datos Abiertos en el 2014 (Comisión Nacional de Investigación Científica y Tecnológica, 2014). México publica el decreto que regula la gestión de datos abiertos (Diario Oficial de la Federación, 2015) y en Argentina, el Ministerio de Ciencia Tecnología e Innovación (MINCYT) reglamenta directamente la Ley N° 26.899 de Repositorios Digitales Institucionales de Acceso Abierto (Honorable Congreso de la Nación, 2013). Esta normativa permite la creación del Programa de Datos abiertos y la del Portal de Datos Abiertos de Ciencia y Tecnología. Por su cercanía, Colombia traza un lineamiento para una Política de Ciencia Abierta a través de un plan de trabajo más detallado con la comunidad de interés proyectado al 2030 (COLCIENCIAS, 2018).

En Latinoamérica la dinámica de los datos abiertos y la ciencia abierta toman mayor interés, diversos estudios coinciden en afirmar que los investigadores reconocen la importancia de compartir los datos de investigación para mejorar la investigación científica, pero solo una minoría está dispuesto a iniciar proyectos que permitan compartir estos datos

⁴ <https://alicia.concytec.gob.pe/>

⁵ <https://busquedas.elperuano.pe/download/url/directiva-que-regula-el-repositorio-nacional-digital-de-cien-resolucion-n-087-2016-concytec-p-1390024-1>

⁶ <https://busquedas.elperuano.pe/normaslegales/aprueban-reglamento-del-registro-nacional-de-trabajos-de-inv-resolucion-no-033-2016-suneducd-1425605-1/>

(Ramírez & Samoilovich, 2018). En el Perú, instituciones como la PUCP⁷ y Centro Internacional de la Papa (CIP)⁸ tienen un portal de datos abiertos a través del software Dataverse, mientras que la Universidad San Ignacio de Loyola (USIL)⁹ y la Universidad Peruana de Ciencias Aplicadas (UPC)¹⁰ presentan colecciones de datos abiertos a través del repositorio institucional (DSpace)

Existen pocos estudios sobre las razones por las cuales los repositorios de datos no se han incrementado. El marco regulatorio sobre los accesos a los datos de investigación y la Ley de Protección de Datos Personales 29733 no se contradicen, siempre que estos últimos se utilicen para actividades de valor histórico, estadístico o científico en forma anónima (Vila, 2017).

2.2. Learning Analytics en el Perú

Learning analytics (o analítica del aprendizaje, LA) ofrece un gran potencial para la educación en Perú y América Latina. LA puede ayudar a atender las necesidades individuales de los distintos estudiantes. Gašević (Gašević, 2018) afirma que no solo puede mejorar la calidad de la enseñanza, sino también reducir el número de alumnos que abandonan la escuela y mucho más. Sin embargo, Cobo and Aguerrebere (Cobo & Aguerrebere, 2018) desafortunadamente muestran que América Latina aún no lo ha logrado. Además, los autores señalan una inquietante ignorancia sobre la importancia del desarrollo de una mayor conciencia y mejor comprensión de LA, y temas relacionados a través de todos los diferentes grupos en nuestra sociedad cada vez más basada en datos. Existe un número reducido de investigadores e investigadoras así como iniciativas de LA en la región (Cobo & Aguerrebere, 2018).

En el Perú existe una iniciativa sin fines de lucro—Learning Analytics Perú, dedicada al liderazgo de iniciativas en favor del desarrollo, promoción e investigación en LA. Esta iniciativa tiene el fin de ejecutar el uso responsable, significativo y provechoso de los datos para mejorar los procesos y resultados de aprendizaje en la educación (Learning Analytics Perú). Una parte de sus líneas estratégicas es el proyecto “Comunidad de Investigadores en Learning Analytics”, cuyo objetivo es atraer, identificar y reunir a profesionales en LA.

Actualmente el grupo está integrado por pocos miembros, no existen en el medio peruano científicos y profesionales que trabajen en LA. Términos como "análisis de datos educativos", "retroalimentación basada en datos" o "acciones basadas en datos en el área de enseñanza/aprendizaje" parecen ser más comunes.

⁷ <https://datos.pucp.edu.pe>

⁸ <https://data.cipotato.org/dataverse.xhtml>

⁹ <https://repositorio.usil.edu.pe/>

¹⁰ <https://repositorioacademico.upc.edu.pe/>

3. Problema Principal

Los problemas que existen en el Perú presentados en capítulo 2, combinados con los problemas particulares que surgen con la recopilación de datos de análisis de aprendizaje, ponen al científico en una situación muy difícil. La legislación, en la que las leyes de protección de datos no están claramente definidas para los datos científicos y la publicación de datos, puede tener potencialmente una gran influencia en el comportamiento de publicación de los datos de investigación de los científicos. Esto provoca, por supuesto, la incertidumbre entre los científicos sobre si deben publicar sus datos a disposición del público en absoluto. Esta investigación se centrará en las barreras que pueden resultar de las regulaciones legales.

El objetivo de nuestro estudio de investigación es obtener una mejor comprensión del estado actual de la publicación de datos en el análisis de aprendizaje y las barreras de publicación de datos influidos por las legislaciones. Los resultados del estudio tienen importantes repercusiones en el diseño y desarrollo de una infraestructura de intercambio de datos.

3.1. Preguntas de Investigación

Las preguntas de investigación consideradas en esta fase del proyecto de investigación son las siguientes:

- ¿Qué preocupaciones tienen y que barreras tienen que superar los científicos del análisis de aprendizaje en Perú al publicar los datos de la investigación?
- ¿Influyen las normas legales del Perú en el comportamiento de publicación de datos de los científicos del análisis de aprendizaje en este país?

Las respuestas a estas preguntas ayudarán a desarrollar hipótesis para la segunda fase del proyecto de investigación, a desarrollarse en los próximos meses.

3.2. Trabajos Relacionados

Las diferentes disciplinas exigen un manejo diferente de los datos de investigación. Cada una de ellas también tiene diferentes inconvenientes para su publicación. Se han realizado algunos estudios sobre los obstáculos para la publicación de datos de investigación en general en diferentes países (Lucraft et al., 2019; Schmidt et al., 2016; Tenopir et al., 2011; Van den Eynden et al., 2016), o en áreas del conocimiento específico, por ejemplo, la psicología (Houtkoop et al., 2018) o las ciencias de la ingeniería (Johri et al., 2016). Según estos estudios, las barreras más comunes se relacionan con aspectos como: no es común en la comunidad, el temor al mal uso y la mala interpretación de los datos, el deseo de publicar los resultados antes de publicar los datos o la falta de requisitos de las revistas.

Se realizó una encuesta a los científicos (Eléspuru Briceño, 2017) de la Universidad Nacional Mayor de San Marcos (UNMSM) y Universidad Peruana de Ciencias Aplicadas (UPC), pero no con un enfoque en las barreras de publicación de datos. No obstante, quedó claro que existe un interés básico en el intercambio de datos, sin embargo, parece haber algo en el camino para evitar su proliferación (solo el 24% de los encuestados utilizan repositorios para sus datos).

A pesar de ello, ninguno de los estudios anteriores se ha enfocado específicamente en la publicación de datos de investigación sobre learning analytics, centrados en las barreras específicas que surgen en América Latina o en el Perú en particular.

4. Metodología

El proyecto de investigación consta de tres fases: una cualitativa, una cuantitativa y la aplicación de los resultados. En este artículo se presentan los resultados de la primera fase del proyecto.

La grabación cualitativa consistió en la realización de entrevistas semiestructuradas las que se realizaron en octubre de 2019 con científicos del Perú. Se entrevistó a investigadores de learning analytics, en medicina y de investigación del impacto climático para evaluar su actitud respecto a la publicación de los datos de sus investigaciones. En las entrevistas se incluyó tanto a científicos *junior* (2-4 años de experiencia) como a científicos *senior* (más de 5 años de experiencia).

En las entrevistas, los científicos explicaron cómo manejan los datos de investigación y describieron sus dificultades en el proceso de publicación. Al principio presentaron su trabajo y explicaron con qué datos de investigación están trabajando. La segunda parte de la entrevista aborda su actitud respecto a la publicación de los datos de su investigación y las razones para la no publicación. Dependiendo del curso de la conversación, se hicieron preguntas sobre la propiedad de los datos. La última parte de la entrevista se centró en el proceso de publicación propiamente dicho y en cómo facilitar el intercambio de datos entre los científicos.

Se seleccionaron tres disciplinas diferentes para excluir el comportamiento específico de la comunidad. Learning analytics, la medicina y la investigación sobre el impacto climático muestran grandes diferencias tanto en términos de los tipos de datos más utilizados como en la sensibilidad de los datos. Mientras que el learning analytics se ocupa de los flujos de clic o de los datos de registro, la medicina se ocupa de los datos genómicos o de observación, y la investigación del impacto climático de los datos climáticos o socioeconómicos. El campo de la medicina trabaja con datos muy delicados y, por lo tanto, puede entrar en conflicto con la protección de datos personales. Learning analytics se basa en datos personales para

personalizar el aprendizaje. Mientras tanto, los científicos que investigan el impacto climático no parecen tener ningún obstáculo de protección de datos que superar.

La evaluación cualitativa de las entrevistas se basa en la metodología de la teoría fundamentada (Charmaz, 2006). Las transcripciones de las entrevistas se han codificado con el enfoque de codificación inductiva, centrándose en las opiniones y acciones de los participantes en relación con la publicación de los datos de la investigación. De la codificación se derivó una dirección teórica que ahora se formula como hipótesis.

Estas hipótesis se verificarán y cuantificarán en la segunda fase de la investigación.

5. Los Resultados de las Entrevistas Semiestructuradas

En total se realizaron 7 entrevistas (Biernacka, 2020) con investigadores en el campo del learning analytics, medicina e investigación del impacto climático de 2 instituciones diferentes en Perú. Estos científicos nombraron un total de 15 barreras diferentes para la publicación de los datos. De ellas, dos parecen estar influenciadas por la normativa legal del país (Tabla 1). En las secciones siguientes se examinarán y analizarán de forma más detallada estas barreras en el ámbito del learning analytics, y se compararán con las otras dos disciplinas.

Tabla 1

Códigos surgidos para las barreras a la publicación de datos de investigación en el learning analytics, en comparación con la medicina y la investigación del impacto climático en el Perú. Los códigos resaltados pueden ser vistos como un impacto directo o indirecto de las regulaciones legales en este país.

Learning analytics	Medicina	Investigación del impacto climático
competencia	calidad de los datos	competencia
la falta de infraestructura	competencia	desconfianza
miedo a la malinterpretación	desconfianza	desconocimiento de las normas legales
no es común en la comunidad (científica)	equilibrando la privacidad y la apertura	la falta de infraestructura
no hay motivación u obligación extrínseca	falta la revisión por pares	miedo a la malinterpretación
vulnerabilidad	la cultura del no compartir no es común en la comunidad (científica) no está permitido publicar	

tiempo y/o esfuerzo de
trabajo
valor no visible

5.1. Learning Analytics

Ambos entrevistados son de la misma institución y cada uno tiene más de 5 años de experiencia en la ciencia. Las barreras que mencionaron eran muy diferentes entre sí. Mientras que una persona está muy preocupada por la vulnerabilidad, el segundo entrevistado subraya que la práctica de publicar datos no es común en la comunidad y que falta la infraestructura necesaria. En sus palabras:

Yo pienso que, bueno, en realidad **no hay esa práctica**, pero si hubiera, yo creo que sí, sería pertinente que se pueda colocar, ¿no? Como te digo, o sea, ya hay una práctica previa donde los trabajos de investigación definen la carrera, o sea, las tesis tanto de pregrado como de posgrado van con todo esto completo, ¿no?, y están en el repositorio. Pero no hay un repositorio, donde están los datos de los artículos que se han publicado. Eso no hay. **No hay esa infraestructura.**

(científico *senior*, learning analytics)

Ninguna de las barreras mencionadas parece depender de la protección de datos u otras normas legales.

5.2. Medicina

En el campo de la medicina se realizaron un total de 4 entrevistas. Dos de los entrevistados no entendieron el concepto de publicación de datos y con cada respuesta recurrieron a sus experiencias con la publicación tradicional de artículos en revistas. Un ejemplo de esto:

Entrevistadora: ¿Ya ha publicado sus **datos** de investigación?

Entrevistada: No, justo ahora quiero publicarlos.

Entrevistadora: ¿Y ahora quiere usted publicar un artículo?

Entrevistada: Sí, un **artículo** previo.

Entrevistadora: ¿Con este artículo también se publican **datos**, como en un repositorio?

Entrevistada: ¿Cómo? ¿Un repositorio? **No entiendo.**

(científica *senior*, medicina)

A partir de estas dos entrevistas no han podido averiguar qué es lo que les impide publicar los datos. Lo que se necesita aquí es una explicación de lo que son los datos de investigación,

las posibilidades de publicación de los datos y el valor añadido que podría tener una publicación. Sin embargo, esto no fue posible al realizar las entrevistas.

Para los otros dos entrevistados, en cambio, surgieron 10 barreras diferentes. Una vez más, aunque los entrevistados no se pusieron de acuerdo sobre las barreras, se mencionó aquí por lo menos una barrera que podría indicar la incertidumbre en cuanto a la normativa legal: *equilibrando la privacidad y la apertura*. Un médico también hace mucho énfasis en la *competencia*, como en el caso de learning analytics:

Entonces, **el área nuestra sí es extremadamente competitiva**, y sobre todo porque la industria farmacéutica está encima para encontrar en una publicación un fármaco que pueda ser comercializado y engrosar las cuentas bancarias de sus grandes capitalistas, lamentablemente.

(científico *senior*, medicina)

5.3. Investigación del Impacto Climático

El único científico entrevistado en la investigación sobre el impacto climático en Perú distingue 5 barreras para la publicación de los datos de la investigación. Hizo especial hincapié en los problemas de la desconfianza y la competencia:

Hay mucha **desconfianza**, ¿no es cierto? Si es que uno entrega datos, no se sabe, si esos datos van a ser utilizados de manera adecuada. Entonces, si son datos que uno ha producido, entonces **la utilización sobre esos datos en primer lugar es de los autores**. Luego, esos datos podrían ser utilizados por otros investigadores, siempre y cuando cuenten con autorización nuestra.

(científico *senior*, investigación del impacto climático)

Sin embargo, este científico ve sus debilidades en su conocimiento sobre las regulaciones legales:

Porque además de eso **no tengo muy claro el tema de la protección de la información** en Perú.

(científico *senior*, investigación del impacto climático)

6. Conclusiones

Para aumentar la capacidad científica es importante tener la posibilidad de acceder no solo a los artículos científicos, sino también a los datos de investigación. Mientras que en medicina

la publicación de datos de investigación parece ser una práctica completamente desconocida, en la investigación sobre el impacto climático los científicos se preocupan principalmente por la competencia y el mal uso de los datos. Existe una actitud de cautela fundamental hacia la publicación de datos.

De igual forma, el learning analytics no es una práctica común en el medio. Existe una incertidumbre sobre cuáles son los datos de la investigación en sí mismos, sobre cómo otros gestionan estos datos y sobre el peligro de hacerse vulnerable al publicar los propios datos de la investigación.

Los científicos de LA y de la investigación sobre el impacto climático concuerdan en que se carece de una infraestructura adecuada para la publicación de los datos de las investigaciones en Perú. El gobierno peruano realiza esfuerzos para establecer depósitos institucionales a través de repositorios de documentos (artículos, tesis, conferencias), falta hacer más énfasis sobre la publicación de los datos de investigación a través de repositorios de datos.

Los científicos de todas las disciplinas estudiadas coinciden en que la competencia es un obstáculo importante para la publicación de los datos. Dado que el éxito de la ciencia depende todavía de "Publish or Perish", la competencia está notoriamente muy presente y en especial en la medicina, donde el prestigio siempre ha dependido del número de artículos científicos publicados en revistas indexadas.

Las barreras mencionadas que están potencialmente influenciadas por regulaciones legales, sugieren que este no es necesariamente el verdadero problema de la publicación de datos en Perú. Por el contrario, existe una ley que regula el acceso abierto a la información, incluidos los datos, pero estas secciones están muy ocultas, por lo tanto, no son visibles ni comprensibles para el personal responsable de la implementación de la infraestructura tecnológica ni mucho menos por los científicos.

7. Limitaciones

La principal limitación de este estudio está relacionada con la muestra, la cual fue autoseleccionada y se basó en recomendaciones a través del efecto bola de caracol. Puede ser que los colegas de trabajo tengan una forma similar de trabajar con los datos de la investigación. Por la misma razón, el número de instituciones incluidas en la muestra es limitado.

Debido a la modesta tasa de respuesta a las solicitudes de entrevistas en el período previo, la distribución de la muestra no es pareja. La preocupación es que los investigadores que aceptaron la solicitud pueden haber tenido ya una opinión firme sobre los datos de investigación abiertos o la publicación de datos en general.

Aunque los autores solicitaron la ayuda de la “Comunidad de Investigadores en Learning Analytics” fue difícil encontrar científicos en esta disciplina en el Perú.

8. Referencias

- Biernacka, K. (2020). *Perspectiva de los Investigadores sobre la Publicación de Datos de Investigación: Entrevistas Semiestructuradas de Perú* Humboldt-Universität zu Berlin. <https://doi.org/https://doi.org/10.18452/21394>
- Charmaz, K. (2006). *Constructing grounded theory: a practical guide through qualitative analysis* (Reprinted ed.). Sage.
- Cobo, C., & Aguerrebere, C. (2018). Building Capacity for Learning Analytics in Latin America. In C. Ping Lim & V. L. Tinio (Eds.), *Learning Analytics for the Global South* (pp. 58-67). Foundation for Information Technology Education and Development.
- COLCIENCIAS. (2018). *Lineamientos Para Una Política De Ciencia Abierta En Colombia*.
- Comisión Nacional de Investigación Científica y Tecnológica. (2014). Propuesta de Política de Datos Abiertos. In.
- CONCYTEC. (2016). *Directiva N° 004-2016-CONCYTEC-DEGC: Directiva que regula el Repositorio Nacional Digital de Ciencia, Tecnología e Innovación de Acceso Abierto* (El Peruano, Issue. <http://busquedas.elperuano.com.pe/download/url/directiva-que-regula-el-repositorio-nacional-digital-de-cien-resolucion-n-087-2016-concytec-p-1390024-1>
- Congreso de la República del Perú. (2013). Ley que regula el Repositorio Nacional Digital de Ciencia, Tecnología e Innovación de Acceso Abierto. In. Lima, Perú: Congreso de la República del Perú.
- Diario Oficial de la Federación. (2015). *Decreto por el que establece la regulación en materia de Datos Abiertos*. . http://www.dof.gob.mx/nota_detalle.php?codigo=5382838&fecha=20/02/2015
- Eléspuru Briceño, M. d. P. L. (2017). *Gestión de datos de investigación en universidades en base al ciclo de vida de los datos. Caso de estudio: Área de Ciencias de la Salud* (Publication Number Magister) Universidad Nacional Mayor de San Marcos]. Lima, Perú. <http://cybertesis.unmsm.edu.pe/handle/cybertesis/6459>
- Gašević, D. (2018). Include us all! Directions for adoption of learning analytics in the global south. In C. Ping Lim & V. L. Tinio (Eds.), *Learning Analytics for the Global South* (pp. 2-22). Foundation for Information Technology Education and Development.
- Ley 26899, (2013). <https://www.argentina.gob.ar/normativa/nacional/ley-26899-223459/texto>
- Houtkoop, B. L., Chambers, C., Macleod, M., Bishop, D. V. M., Nichols, T. E., & Wagenmakers, E.-J. (2018). Data Sharing in Psychology: A Survey on Barriers and Preconditions. *aps*, 1(1), 70-85. <https://doi.org/https://doi.org/10.1177/2515245917751886>
- Johri, A., Yang, S., Vorvoreanu, M., & Madhavan, K. (2016). Perceptions and Practices of Data Sharing in Engineering Education. *Advances in Engineering Education*. <https://doi.org/10.1371/journal.pone.0021101.s001>
- Learning Analytics Perú. *Somos*. <https://laperu.org/somos/>
- Lucraft, M., Allin, K., Baynes, G., & Sakellaropoulou, R. (2019). Challenges and Opportunities for Data Sharing in China. <https://doi.org/https://doi.org/10.6084/m9.figshare.7718441.v1>
- Ramírez, P. A., & Samoilovich, D. (2018). *Ciencia abierta. Reporte para tomadores de decisiones 2da edición-Abril 2019*.
- Schmidt, B., Gemeinholzer, B., & Treloar, A. (2016). Open Data in Global Environmental Research: The Belmont Forum's Open Data Survey. *PLoS One*, 11(1), e0146695. <https://doi.org/https://doi.org/10.1371/journal.pone.0146695>
- Superintendencia Nacional de Educación Superior Universitaria. (2016). *Aprueban Reglamento del Registro Nacional de Trabajos de Investigación para optar grados académicos y títulos profesionales - RENATI*.
- Tenopir, C., Allard, S., Douglass, K., Aydinoglu, A. U., Wu, L., Read, E., . . . Frame, M. (2011). Data sharing by scientists: practices and perceptions. *PLoS One*, 6(6), e21101. <https://doi.org/10.1371/journal.pone.0021101>

- Van den Eynden, V., Knight, G., Vlad, A., Radler, B., Tenopir, C., Leon, D., . . . Corti, L. (2016). *Towards Open Research. Practices, experiences, barriers and opportunities*. Welcome Trust. <https://doi.org/https://dx.doi.org/10.6084/m9.figshare.4055448>
- Vila, C. (2017). Datos abiertos y protección de datos personales. In.