



Author Carpentry

Marcela Alfaro Córdoba, Ph.D.
Profesora Invitada, Escuela de Estadística
Universidad de Costa Rica



TEC

UNA
UNIVERSIDAD
NACIONAL
COSTA RICA

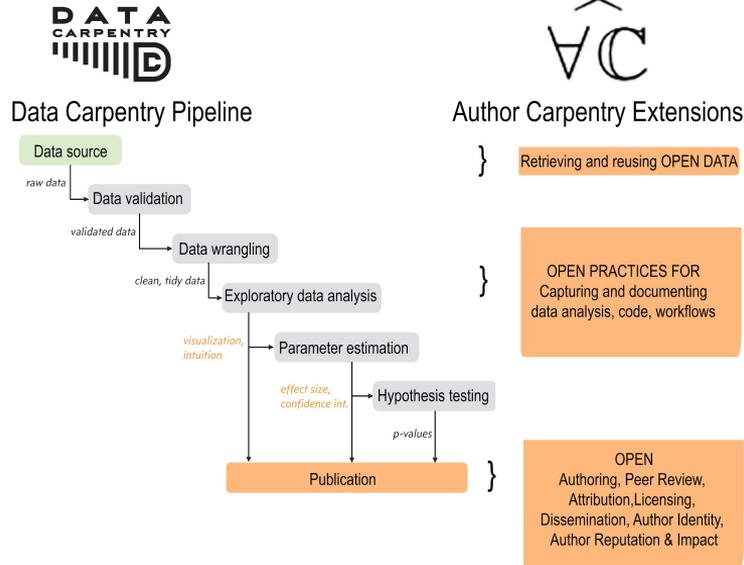


UTN
Universidad
Técnica Nacional

IIP Instituto de
Investigaciones
Psicológicas

EES Escuela de
Estadística

¿Qué es Carpintería para Autores?



Adapted from Bois, Justin, "Reproducibility through computing Caltech Data Carpentry", Caltech Library, April 26, 2017.

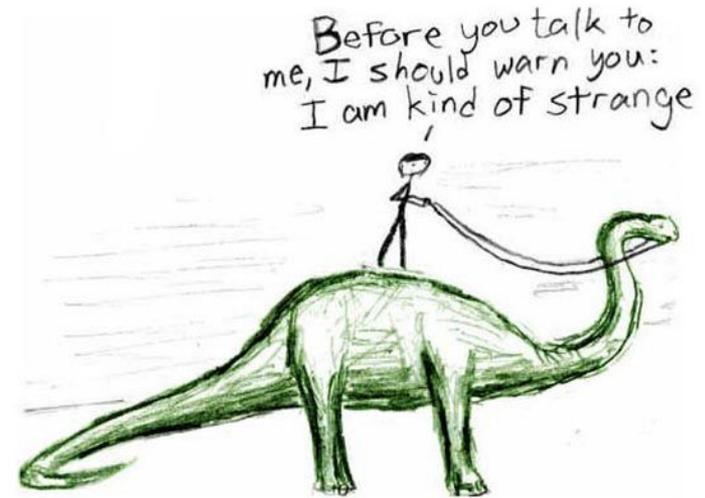
Fuente: <https://authorcarpentry.github.io/>

- Recursos preparados por una comunidad de profesionales de la información que trabajan en apoyo a la investigación, con algunas mejoras hechas por personas investigadoras.
- El programa está inspirado en la filosofía de enseñanza propuesta por Greg Wilson, fundador de [Software Carpentry](#) y se basa en proveer herramientas para la "última milla" del proceso de investigación: redacción, informe, revisión, difusión y licencia, medición de impacto y establecimiento de la identidad y reputación del autor.

Cuéntenme de Uds

En parejas van a hacerse las siguientes preguntas. Tienen 10 minutos para presentarse entre Uds y luego vamos a presentar a la otra persona:

- Afiliación institucional.
- Área de investigación.
- ¿Ya ha publicado en revistas científicas?
- ¿Ha escuchado de alguno de los temas de *Author Carpentry*?
- ¿Cuáles de los temas que mencioné le interesa más conocer?



Fuente: <https://xkcd.com/15/>

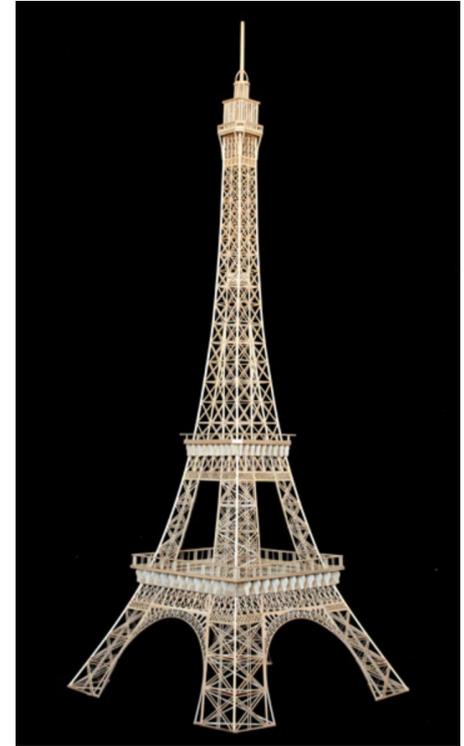
¿Quién soy yo? ¿Por qué enseño esto?

- Profesora e investigadora de la Escuela de Estadística, del Posgrado en Estadística y del Centro de Investigación en Matemática Pura y Aplicada de la Universidad de Costa Rica. Más información: <https://malfaro2.github.io/>
- Organizadora del capítulo [R-ladies en San José](#) y de [ConectaR2019](#), la primera Conferencia de usuarios de R en América Central. Co-directora de [CODATA-RDA Schools of Data Science](#), y en este mes fui instructora del curso [Universidad a través de Datos Abiertos](#) del PROLEDI.

CODATA: Committee on Data of the International Council for Science y *RDA: Research Data Alliance*, son organizaciones internacionales preocupadas por el estado y el futuro de los datos, las habilidades para manejarlos y los esfuerzos de estandarización que aseguren un uso eficiente.

[Estadística y Ciencia Abierta]

- Estadística Aplicada: a mi me gusta verla como la disciplina/ciencia que juega el papel del pegamento, si nos imaginamos el proceso de creación científica como una estructura.
- En muchas ocasiones nos toca ser los administradores, recolectores, curadores, analistas y los que archivan y mantienen los datos. Esto no es exclusivo de nuestra profesión, pero es central en nuestro quehacer.
- Entonces, si nosotros no nos estamos haciendo las preguntas acerca de cómo tratar, administrar, analizar y mantener los datos de investigación, ¿a quién le corresponde esa responsabilidad?



¿Qué vamos a hacer en estos dos días?

- Author Carpentry es un programa de capacitación e información de investigador(a) a investigador(a) sobre autoría y publicaciones abiertas, identidad y reputación académica e impacto en la investigación.
- Mi objetivo es compartirles lo que he aprendido hasta ahora como investigadora que quiere abrir sus productos de investigación, pero también quiero aprender de todos(as) Uds.
- Producto de mi entorno científico, es posible que mis recomendaciones, herramientas, y experiencias estén sesgadas a mi rango de aplicaciones. Sin embargo, me encantaría aprender cómo éstas pueden ser útiles (o inservibles!) en sus áreas de acción.
- También, me encantaría que tomen este material y lo compartan con otras personas que se dedican a la investigación, por lo que esta presentación tiene una licencia de:



Agenda del taller

Miércoles 24 de octubre:

- Promocionar sus artículos para aumentar su identidad digital (repositorios, perfiles, uso de redes sociales y académicas) y el impacto de su investigación.
- Establecer su identidad académica con ORCID y gestionar la reputación con plataformas de intercambio académico habilitadas por ORCID, como ScienceOpen.
- Identificar mejores prácticas en atribución, reconocimiento y cita, en particular para productos de investigación no tradicionales (software, conjuntos de datos).
- Identificar las publicaciones de acceso abierto de buena reputación.
- Revisar las opciones de Intercambio abierto de investigación utilizando licencias Creative Commons, exenciones y marcas de dominio público.

Miércoles 25 de octubre:

- Manejar herramientas para la escritura científica en lenguajes de marcado útiles, potentes y abiertos: RMarkdown.
- Formular y administrar datos de citas, listas de publicaciones y bibliografías en formatos abiertos como BiBTeX, y aprender a usar herramientas de administración de referencias de código abierto como Zotero.

Requisitos para los dos días: traer su propia computadora, con [R y R Studio instalados](#).

¿De dónde salió la agenda?

Top ten suggestions to help increase your reach and impact! de Jon Tennant + *Author Carpentry*

1. Coloque los artículos en un repositorio temático apropiado.
2. Coloque los artículos en su página web institucional y en el repositorio ([Kérvá](#)).
3. Informe a los usuarios interesados en [Twitter](#). Elija #hashtags apropiados para aumentar la capacidad de descubrimiento de su trabajo.
4. Publique enlaces a artículos en [Facebook](#), en cualquier grupo relevante, o en sitios de redes académicas, incluidos sus perfiles de [LinkedIn](#) u otro de su elección.
5. Guárdelo en su biblioteca de referencia y hágales promoción en la red académica de gestores de referencia tales como [CiteuLike](#) o [Zotero](#).
6. Comparta los datos y el código de su investigación y aumente sus citas.
7. Informe a los medios de comunicación y bloggers en su campo.
8. Cree un perfil de [Google Scholar](#) y haga un seguimiento de sus citas.
9. Agregue sus artículos a su cuenta de [ORCID](#).
10. Y, si se siente inclinado a hacerlo, actualice su perfil de [ScienceOpen a través de ORCID](#), y realice un seguimiento de sus métricas a nivel de artículo y autor para todos sus artículos de investigación.

D1: Promocionar sus artículos

¿Para qué queremos eso?

- Lo ideal es que nuestros resultados de investigación sean útiles, ya sea para otras personas investigadoras de nuestra misma área, o para el público en general (por medio de políticas públicas, decisiones de negocios y hasta decisiones personales).
- La utilidad de nuestros resultados deja de ser relevante si no tenemos una divulgación correcta de nuestro trabajo. ¿Cómo van a ser útiles nuestros resultados si nadie sabe acerca de ellos?
- El objetivo entonces es aumentar su identidad digital y con ello, el impacto de su investigación.

¿Cómo?

- Repositorios, perfiles, uso de redes sociales.

D1: Promocionar sus artículos

Repositorios:

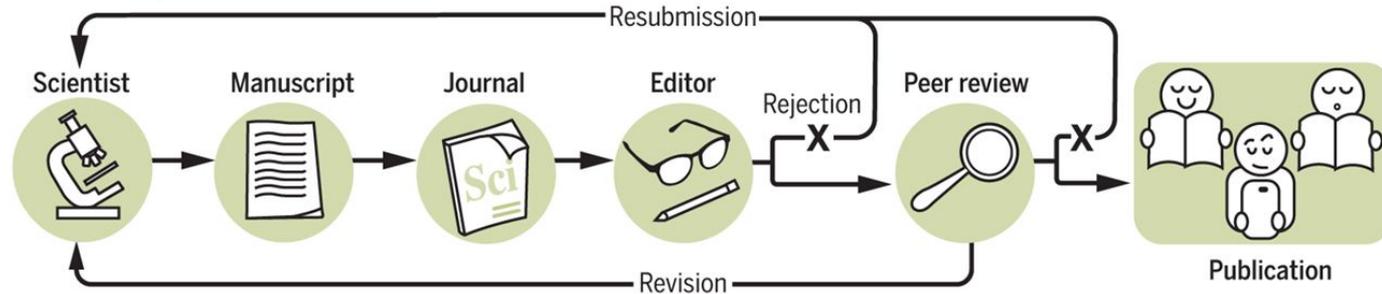
- Coloque los artículos en un repositorio temático apropiado.
- Coloque los artículos en su página web institucional y en un repositorio.

Para trabajar con su compañero(a) de clase:

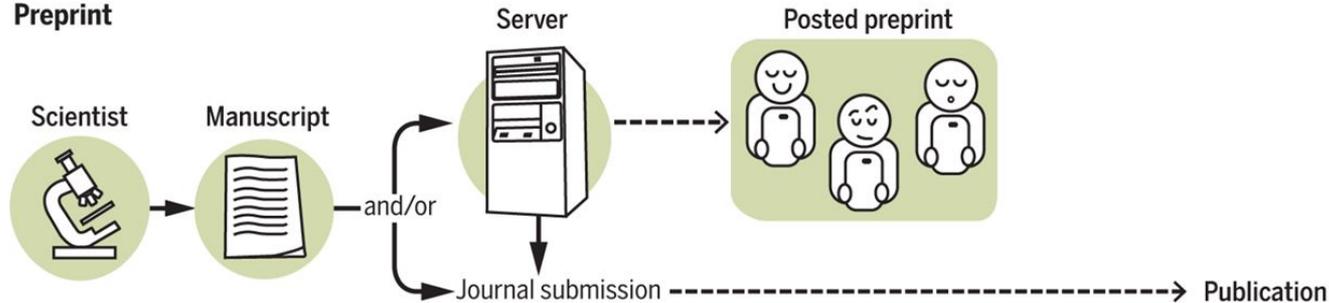
1. Haga una lista de los repositorios apropiados para su área de conocimiento.
2. Haga una lista de repositorios apropiados para su afiliación institucional - ¿tiene varias opciones?

¿Qué vamos a depositar?

Peer review publication



Preprint



Repositorios por área (internacionales)



Preprint servers are appearing in many disciplines



Fuente:
<https://www.slideshare.net/BaltimoreNI/inglis-preprints-in-biology-and-medicine>

Repositorios institucionales (CONARE)



Otros repositorios institucionales:

<http://eprints.rclis.org/18151/1/Repositorios%20Costa%20Rica.pdf>

¡Kérwá es el único de auto depósito! Y no todos los repositorios institucionales son interoperables: lo que no está en Kimuk no se cosecha a nivel mundial (La Referencia, y OpenAIRE)

Además de los preprints de artículos, podemos depositar:

- Informes de investigación.
- Material didáctico.
- Presentaciones de conferencias.
- Pósters presentados en conferencias.
- Imágenes incluidas en el artículo.
- Datos y código
- Un largo etc.

¿Dónde? <https://www.nature.com/sdata/policies/repositories>

¿Cuántas personas han depositado documentos?

¿Dónde? ¿Cómo? ¿Cuándo?

¿Tienen dudas acerca de cómo depositar en Kérwá?

Vamos a dar 20 mins para el depósito de un documento o la búsqueda de los que ya depositaron.

Busquen su nombre: <https://www.base-search.net/> o <http://www.redalyc.org/autorHome.oa>

D1: Promocionar sus artículos

¡Las redes sociales pueden ser redes académicas!

- Informe a los usuarios interesados en Twitter. Elija #hashtags apropiados para aumentar la capacidad de descubrimiento de su trabajo.
- Publique enlaces a artículos en Facebook, en cualquier grupo relevante, o en sitios de redes académicas, incluidos sus perfiles de LinkedIn u otro de su elección.
- Guárdelo en su biblioteca de referencia y hágales promoción en la red académica de gestores de referencia tales como CiteuLike o Zotero.
- Informe a los medios de comunicación y bloggers en su campo.

D1: Promocionar sus artículos

- Twitter, Facebook, CiteuLike, LinkedIn, Zotero.
- Informe a los medios de comunicación y bloggers en su campo.

Para comentar en clase:

¿Cuáles de las redes sociales y/o opciones anteriores ya está utilizando? ¿cómo las utiliza? ¿Conocen bloggers de su campo de investigación? ¿han seguido discusiones académicas en redes sociales?

[Social media is more than simply a marketing tool for academic research \(2012\)](#)

[Personal reputation in a social media world](#)

<https://www.timeshighereducation.com/blog/weird-and-wonderful-world-academic-twitter>

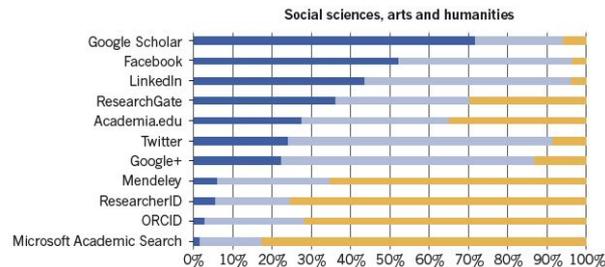
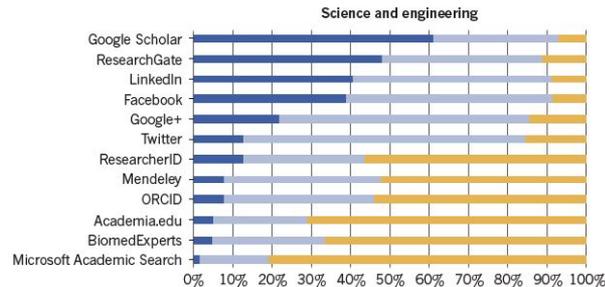


D1: Promocionar sus artículos

REMARKABLE REACH

More than 3,000 scientists and engineers told Nature about their awareness of various giant social networks and research-profiling sites. Just under half said that they visit ResearchGate regularly. Another 480 respondents in the humanities, arts and social sciences were less keen on ResearchGate.

- I am aware of this site and visit regularly
- I am aware of this site but do not visit regularly
- I am not aware of this site



¿Cuáles usan Uds?
 ¿Cuáles no conocen?
 ¿Cuáles les gustaría conocer?

Mi combinación favorita:

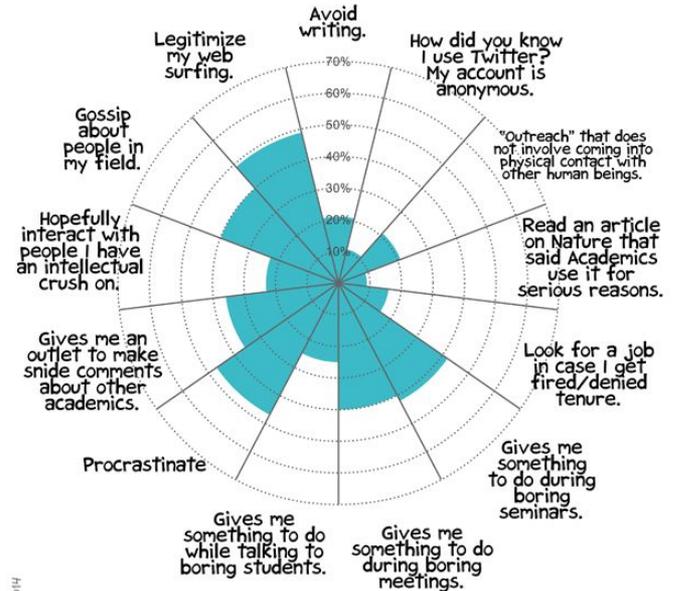
ORCID + Webpage + Github + Google Scholar + Twitter

D1: Promocionar sus artículos

Su turno:

En parejas, discutan cuál combinación de redes sociales académicas y sitios de perfiles académicos es más conveniente para su área de investigación.

Why Academics REALLY Use Twitter



JORGE CHAM © 2014

www.phdcomics.com

Original graph from: nature.com/news/online-collaboration-scientists-and-the-social-network-1.15711

D1: Establecer su identidad académica

¿Para qué?

- ¿Cuántas personas hay en el mundo con su mismo nombre o similar?
- ¿Cómo voy a establecer mi identidad académica si no tengo todo en un solo lugar?
- No solo le ayuda a su imagen de investigador(a), sino que también le ayuda a su Universidad.

¿Cómo?

- Con [ORCID](#) y desde ahí, gestionar la reputación con plataformas de intercambio académico habilitadas por ORCID, como ScienceOpen, Redalyc, etc.
- Un perfil de [Google Scholar](#) es el más eficiente para llevar un recuento de sus citas (aunque mucha gente prefiera ResearchGate).
- Y, si se siente inclinado a hacerlo, actualice su perfil de [ScienceOpen a través de ORCID](#), y realice un seguimiento de sus métricas a nivel de artículo y autor para todos sus artículos de investigación.

D1: Establecer su identidad académica



Connecting Research
and Researchers

<https://vimeo.com/97150912>

¿No tienen ORCID? Hagámoslo de una vez:

<https://authorcarpentry.github.io/orcid-profile/00-orcid-profile.html>

<https://authorcarpentry.github.io/orcid-profile/01-adding-works.html>

¿Ya lo tiene? Conéctelo con ScienceOpen:

<http://blog.scienceopen.com/2016/06/orcid-integration-at-scienceopen/>

D1: Establecer su identidad académica



<https://www.youtube.com/watch?v=t1ZwgDeX2eQ>

Solo pueden crearlo si tienen publicaciones. Sin embargo, yo lo recomiendo mucho porque es la fuente de los rankings web de las Universidades del mundo. Además, es una herramienta útil para tener control de las citas de nuestros trabajos.

D1: Identificar mejores prácticas

¿Para qué?

- Nuestros productos de investigación no son solo artículos científicos, tenemos que reconocer que es cada vez más común tener muchos otros formatos de presentación de resultados.
- En particular, las investigaciones cuantitativas necesitan un buen sistema de atribución, reconocimiento y cita para productos de investigación no tradicionales (software, conjuntos de datos). ¡REPRODUCIBILIDAD!
- Además, varios estudios (la Dra. Mckiernan ha hablado de esto) señalan que los artículos con datos y código disponibles tienen más citas en promedio.

¿Cómo?

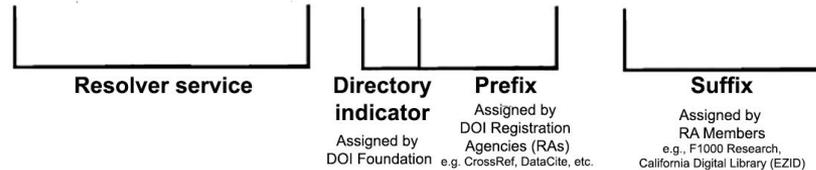
- Uso de repositorios y de DOI (identificadores únicos)

D1: Identificar mejores prácticas

Uso de repositorios y de DOI (identificadores únicos)

Anatomy of a Digital Object Identifier (DOI)

`https://doi.org/ 10.7935 / K5H41PBP`



[Adapted from Australian National Data Service (ANDS)
Digital Object Identifier System and DOI Names (DOIs) Guide,
<http://www.ands.org.au/guides/doi>]

D1: Identificar mejores prácticas

Uso de repositorios y de DOI (identificadores únicos)



¿Cómo se hace? <https://authorcarpentry.github.io/doi-citation-data/01-register-doi.html>

¿Quién más asigna DOIs? ¿Qué cuidados debemos tener?

D1: Identificar mejores prácticas

Repository Name	Information on fees/costs	Size limits	Integrated with <i>Scientific Data's</i> manuscript submission system	Re3data / FAIRsharing entry
Dryad Digital Repository	\$120 USD for first 20 GB, and \$50 USD for each additional 10 GB	None stated	Yes ✓	view FAIRsharing entry
figshare	100 GB free per <i>Scientific Data</i> manuscript. Additional fees apply for larger datasets	1 TB per dataset	Yes ✓ - To qualify for the 100 GB of free storage, data must be uploaded to figshare via our submission system. Download instructions.	view FAIRsharing entry
Harvard Dataverse	Contact repository for datasets over 1 TB	2.5 GB per file, 10 GB per dataset	No	view re3data entry
Open Science Framework	Free of charge	5 GB per file, multiple files can be uploaded	No	view FAIRsharing entry
Zenodo	Donations towards sustainability encouraged	50 GB per dataset	No	view re3data entry
Mendeley Data	Contact repository for datasets over 10 GB	10 GB per dataset	No	view FAIRsharing entry

Depósito de Datos y código

¿Por qué no hacerlo o no en Kérwá? ¿qué necesita la UCR para tener uno de estos?

Fuente: <https://www.nature.com/sdata/policies/repositories>

D1: Identificar mejores prácticas

Depósito de Datos y Código:

¿Qué cuidados hay que tener?

- Código y datos para ciencia reproducible [*A toolkit for data transparency takes shape.*](#) Además, <https://ropensci.org/>, <http://jupyter.org/>
- Datos sin conflictos de privacidad ([o anonimizados para evitarlo](#)) ¿Tenemos conflictos legales?
- ¿A quién pertenecen los datos recolectados con FONDOS PÚBLICOS?
- FAIR DATA: <https://www.nature.com/articles/sdata201618>
- Sitio recomendado: <https://www.nature.com/collections/prbfkwmwvz>

D1: Identificar publicaciones de buena reputación

¿Para qué?

- Cada día hay más Revistas depredadoras (*Predatory Journals*) interesadas en lucrar de nuestra necesidad de publicación.
- Estas revistas no están indexadas, probablemente no tienen un comité editorial aceptable y nos cobrarán por la publicación.

¿Cómo?

- Revisar el sitio: <https://predatoryjournals.com/journals/>
- Sitio construido gracias a un esfuerzo comunitario, después de que un profesional de la información se viera obligado a borrar su página debido al acoso de estas revistas.

[Referencia.](#)

D1: Identificar publicaciones de buena reputación

<https://predatoryjournals.com/journals/>

- Una revista se considera como depredadora si cumple con estos criterios:
 - Cobra montos exorbitantes por la publicación de artículos, y además con una falta de revisión por pares o de supervisión editorial.
 - Avisa a los autores de las tarifas solo después de la aceptación,
 - Se dirige a los académicos a través del correo masivo o de correo no deseado para intentar que publiquen o sirvan en consejos editoriales.
 - Rápida aceptación de artículos de baja calidad, incluidos artículos falsos.
 - Utilizan el nombre de académicos como miembros de juntas editoriales sin su permiso o no permitiéndoles renunciar. Además, muchas veces hacen un listado de eruditos falsos como miembros de consejos editoriales o autores.
 - Copia el diseño visual y el lenguaje de los materiales de marketing y los sitios web de revistas legítimas y establecidas. Uso fraudulento o indebido de ISSNs.
 - Da información falsa sobre la ubicación de la operación de publicación.
 - Factores de impacto falsos, inexistentes o mal representados.

D1: Identificar publicaciones de buena reputación

<https://predatoryjournals.com/journals/>

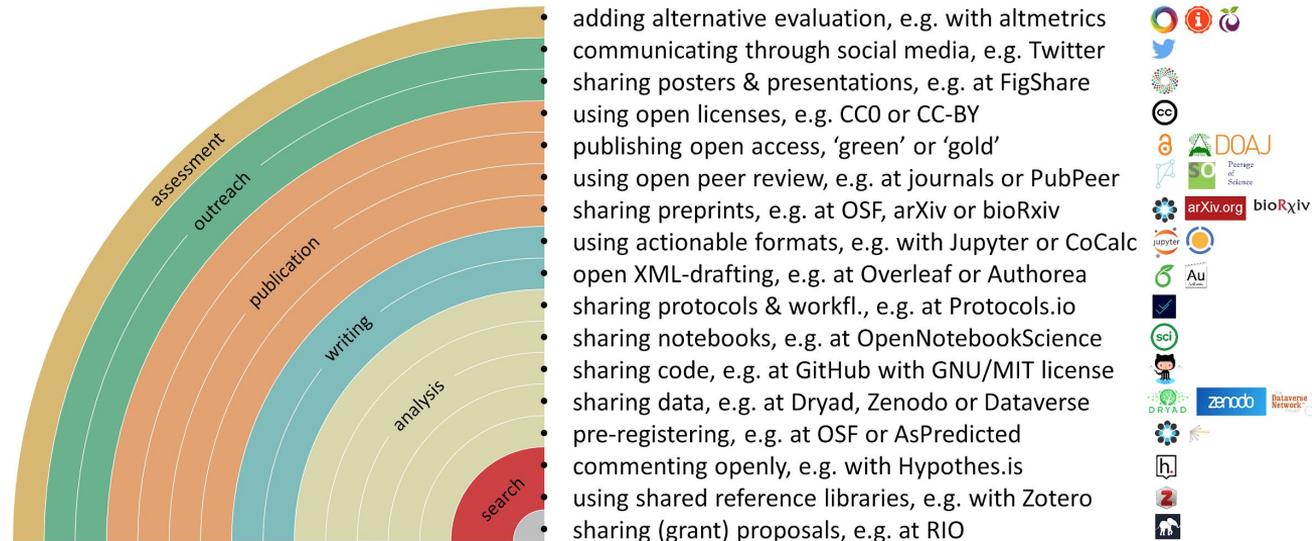
- Ahora en parejas: revisen la lista.
- ¿Hay alguna revista de su área que les hubiera parecido aceptable si no estuviera en esa lista?



Stop Predatory Journals

D1: Revisar las opciones de Intercambio abierto

You can make your workflow more open by ...



D1: Revisar las opciones de Intercambio abierto

¿Para qué?

- No podemos compartir todo sin antes detenernos a pensar en el futuro uso que se le va a dar a nuestras obras.

¿Cómo?

- Licencias de Creative Commons



D1: Revisar las opciones de Intercambio abierto

- En lugar de tener “Derechos Reservados”, CC hace un balance entre las cosas que queremos (y/o podemos) compartir y las cosas que no queremos que se modifiquen.
- ¿Cómo escoger la licencia? <https://creativecommons.org/choose>
 - ¿Quiere permitir adaptaciones a su trabajo cuando se comparte?
 - Si, No, o Si pero tienen que compartir mi trabajo como está.
 - ¿Quiere permitir usos comerciales de su trabajo?
 - Si o No.

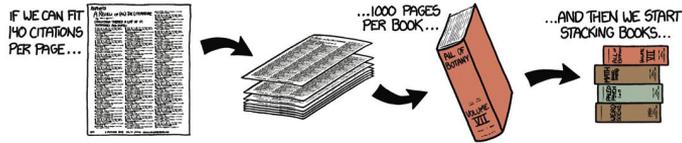
Explicación detallada de cada licencia en español:

https://creativecommons.org/licenses/?lang=es_ES

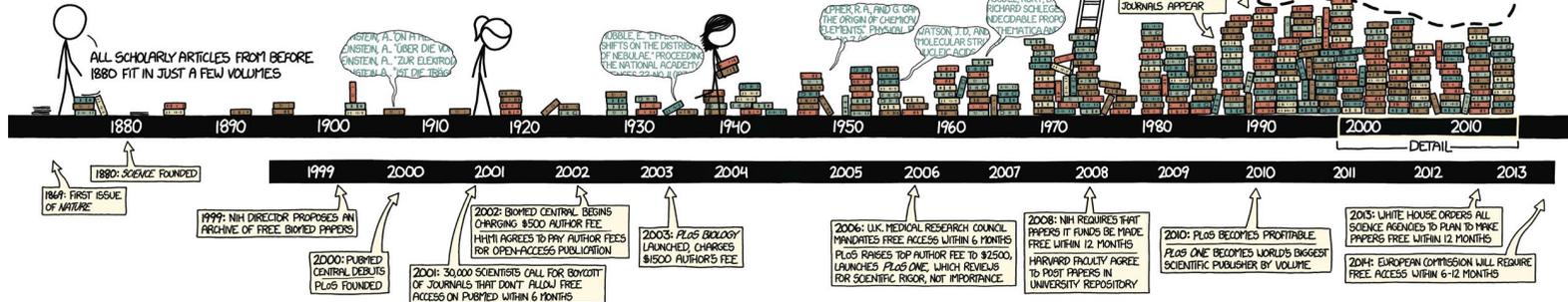


HOW MUCH SCIENCE IS THERE?

SCIENTIFIC PUBLISHING HAS BEEN ACCELERATING—A NEW PAPER IS NOW PUBLISHED ROUGHLY EVERY 20 SECONDS. LET'S IMAGINE A BIBLIOGRAPHY LISTING EVERY SCHOLARLY PAPER EVER WRITTEN. HOW LONG WOULD IT BE?



...THIS IS WHAT THE FULL LIST WOULD LOOK LIKE:



HOW OPEN IS IT?

SINCE THE ADVENT OF THE WEB, MUCH OF SCIENTIFIC PUBLISHING HAS BEEN MOVING TO OPEN ACCESS. ACCORDING TO SCIENCE-METRIX, OPEN ACCESS REACHED A "TIPPING POINT" AROUND 2011: MORE THAN 50% OF NEW RESEARCH IS NOW MADE AVAILABLE FREE ONLINE.

OPEN-ACCESS PAPERS

AS JOURNALS MOVE TO OPEN ACCESS AND DIGITIZE THEIR ARCHIVES, OLD PAPERS FROM EVERY PERIOD MOVE UP HERE...

...IN ADDITION TO THE FLOOD OF NEW PAPERS BEING PUBLISHED HERE DIRECTLY.

25% OF OPEN-ACCESS PAPERS ARE FREELY AVAILABLE ON PUBLICATION. THE REST BECOME FREE WITHIN 12 MONTHS ON JOURNAL WEBSITES OR OTHER REPOSITORIES.

TRADITIONAL PUBLICATION

MOVED TO OPEN ACCESS

¡Nos vemos mañana!

<https://malfaro2.github.io/>
marcela.alfarocordoba@ucr.ac.cr



D2: Manejar herramientas para la escritura científica

¿Para qué?

El uso lenguajes de marcado útiles, potentes y abiertos (como RMarkdown) facilita el flujo de trabajo para que sea reproducible.

¿Cómo?

- RMarkdown desde RStudio y usando Github como Repositorio (si hay tiempo).
- <https://bookdown.org/yihui/rmarkdown/>, <https://swcarpentry.github.io/r-novice-gapminder-es/15-knitr-markdown/index.html>
- <https://rmarkdown.rstudio.com/lesson-6.html>



D2: Formular y administrar datos de citas

¿Para qué?

Listas de publicaciones y bibliografías en formatos abiertos como BiBTeX, así como el uso de herramientas de administración de referencias de código abierto como Zotero nos ayudan a ser más eficientes con nuestras referencias.

¿Cómo?

- BiBTeX + Zotero: <https://libguides.mit.edu/cite-write/bibtex>
- <https://libguides.mit.edu/cite-write/zotero>

BIBTEX + zotero

101 INNOVATIONS IN SCHOLARLY COMMUNICATION



Jeroen Bosman @jeroenbosman
Utrecht University Library

THE CHANGING RESEARCH WORKFLOW

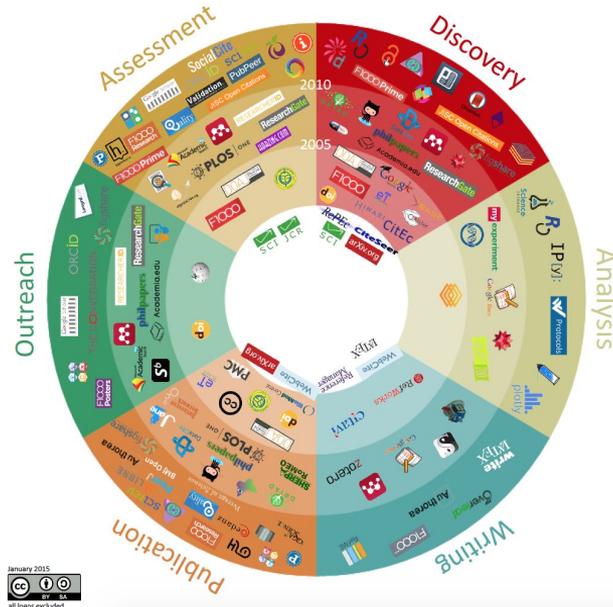


Bianca Kramer @MsPhelps
Utrecht University Library

Science is in transition. This poster gives an impression of the exploratory phase of a project aiming to chart innovation in scholarly information and communication flows from evolutionary and network perspectives.

We intend to address the questions of what drives innovation and how these innovations change research workflows and may contribute to more open, efficient and good science.

101 Innovative tools and sites in 6 research workflow phases (< 2000 - 2015)



Most important developments in 6 research workflow phases

	Discovery	Analysis	Writing	Publication	Outreach	Assessment
Trends	social discovery tools	data-driven & crowdsourced science	collaborative online writing	Open Access & data publication	scholarly social media	article level altmetrics
Expectations	growing importance of data discovery	more online analysis tools	more integration with publication & assessment tools	more use of "publish first, judge later"	use of altmetrics for monitoring outreach	more open and post-publication peer review
Uncertainties	support for full-text search and text mining	willingness to share in analysis phase	effect of collaborative online writing	journal/publisher status	requirements of funders & institutions	who pays for costly qualitative assessment?
Opportunities	discovery based on aggregated OA full text	open labnotes	semantic tagging while writing/citing	reader-side paper formatting	using author-, institutional- and affiliation-IDs	using author-, publication- and affiliation-IDs
Challenges	real semantic search (concepts & relations)	reproducibility	safety/privacy of online writing	globalization of publishing/access standards	making outreach a two-way discussion	quality of measuring tools
Most important long-term development	multidisciplinary + citation-enhanced databases	collaboration + data-driven	online writing platforms	Open Access	more & better connected researcher profiles	Importance of societal relevance + non-publication contributions
Potentially most disruptive development	semantic/concept search + contextual/social recommendations	open science	collaborative writing + integration with publishing	circumventing traditional publishers	public access to research findings, also for agenda setting	moving away from simple quantitative indicators

Typical workflow examples



¡Muchas gracias!

<https://malfaro2.github.io/>
marcela.alfarocordoba@ucr.ac.cr

