

What is CODATA-RDA?

The **Committee on Data for Science and Technology** (**CODATA**) was established in 1966 as an interdisciplinary committee of the [International Council for Science](#). It seeks to improve the compilation, critical evaluation, storage, and retrieval of data of importance to [science](#) and [technology](#).^[2]

CODATA sponsors the CODATA international conference^[3] every two years.

The schools are a working group of CODATA:

<http://www.codata.org/working-groups/research-data-science-summer-schools>

The **Research Data Alliance (RDA)** is a research community organization started in 2013 by the [European Commission](#), the American [National Science Foundation](#) and [National Institute of Standards and Technology](#), and the Australian [Department of Innovation](#). Its mission is to build the social and technical bridges to enable open sharing of data.^[2] The RDA vision is researchers and innovators openly sharing data across technologies, disciplines, and countries to address the grand challenges of society.^[2] The RDA is a major recipient of support in the form of grants from its constituent members' governments.^{[8][9]}

As of August 2019, the RDA has over 8,800 individual members from 137 countries.^[10]

Do you want to get involved?

- Organize workshops at your home institution.
- Use the school's material to teach.
- Attend RDA Plenaries.
- Join our alumni community
- Be a helper / instructor in our schools!
- Have any other ideas on how to get involved? Let us know

RDA Plenary 2020 will be in Costa Rica



16th Research Data Alliance
PLENARY MEETING
NOV.2020 San José, Costa Rica

Knowledge **Ecology**

Organized by CONARE Costa Rica and RDA United States

Logos for CONARE, UNICA, SUCRE, and RDA (Research Data Alliance) are visible in the bottom left corner.

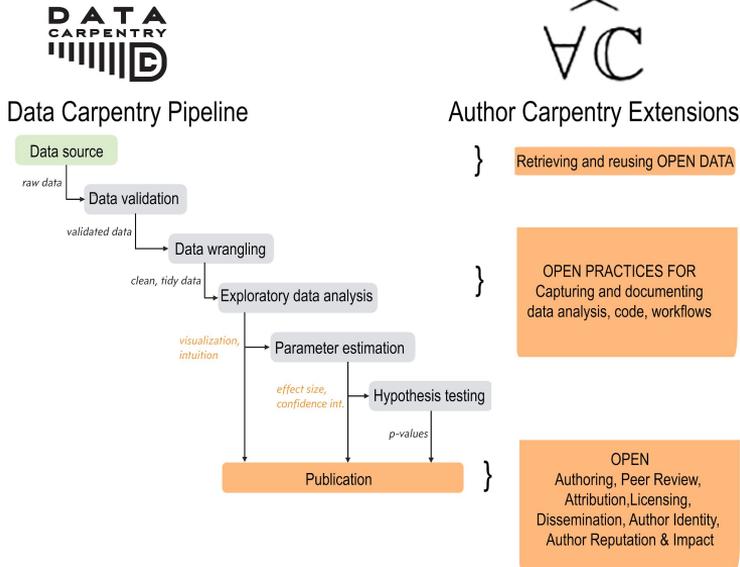
Author Carpentry

CODATA-RDA School of Research Data Science

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



What is Author Carpentry?



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

- Author Carpentry, a researcher-to-researcher training and outreach program in open authoring and publishing, scholarly identity and reputation, and research impact. Inspired by the [Modern Scientific Authoring](#) lesson proposed by Software Carpentry founder Greg Wilson, Author Carpentry builds on this vision to complement and extend the research pipelines covered in Software and Data Carpentry. Author Carpentry picks up with the "last mile" of the research process: writing, reporting, review, dissemination and licensing, impact measurement, and establishing author identity and reputation (as shown in the illustration below).

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

About me

- Assistant Professor at the School of Statistics, Statistics and Computer Science Graduate Programs and CIMPA (Centro de Investigación en Matemática Pura y Aplicada), Universidad de Costa Rica. More info: <https://malfaro2.netlify.io/>
- Founder and organizer of [R-ladies en San José](#) and [ConectaR2019](#), first conference of R users in Central America. Co-chair of [CODATA-RDA Schools of Data Science](#).

[Data Science and Open Science]

- From my perspective as an applied statistician, I see making science as building a big structure, where the documentation and data analysis are the glue/cement.
- In a lot of occasions, after doing the data analysis, we end up (whether we like it or not) being the data managers, collectors, curators, archivers and maintainers.
- Thus, if we are not making the right questions about how to treat, manage, analyze or maintaining our data, then, who is the person responsible of that?



What we are covering this morning

[Top ten suggestions to help increase your reach and impact!](#) de Jon Tennant + [Author Carpentry](#)

1. Place articles on your institutional webpage or repository.
2. Place articles in an appropriate subject repository ([Kérwá](#)).
3. Inform interested users on [Twitter](#). Choose appropriate #hashtags to increase the discoverability of your work.
4. Post links to articles on [Facebook](#), and any relevant groups, as well academic networking sites, including your [LinkedIn](#) profiles.
5. Save it to your reference library and promote in the academic network of reference manager sites such as [CiteuLike](#) or [Zotero](#).
6. Share your research data and code (when possible) and increase your citations.
7. Inform society news outlets and bloggers in your field.
8. Create a Google Scholar profile and track your citations.
9. Add articles to your [ORCID](#) account.
10. And, if you feel inclined to do so, [update your ScienceOpen profile through ORCID](#), and track your article- and author-level metrics for all your research articles.

What we are covering this morning

- Promote your work, articles
 - Public repos, academic social networking
- Establish your academic identity
 - Best practices
 - Identifying good journals
- Reviewing our workflow
 - Open workflow
 - Manage citations
 - Open markup languages for reproducible reports
 - DOIs for GitHub repos

Promote your work

Why do we need to do this?

- Ideally, our research results will be useful, either for other researchers in our same area, or for the general public (through public policies, business decisions and even personal decisions).
- The usefulness of our results is no longer relevant if we do not have a correct disclosure of our work. How will our results be useful if nobody knows about them?
- The goal then is to increase your digital identity and with it, the impact of your research.

How?

- Use of repositories, academic social networks, academic profiles.

Promote your work

Repositories:

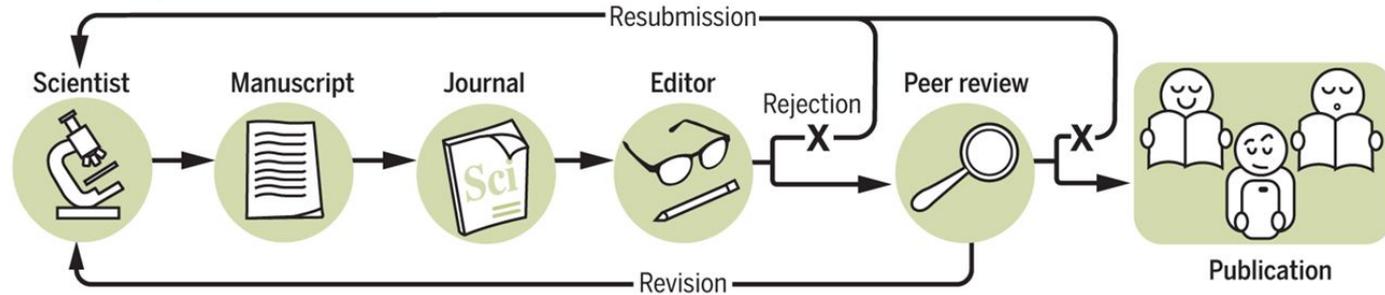
- Place articles on your institutional webpage or repository.
- Place articles in an appropriate subject repository (Kérwá).

Work in pairs:

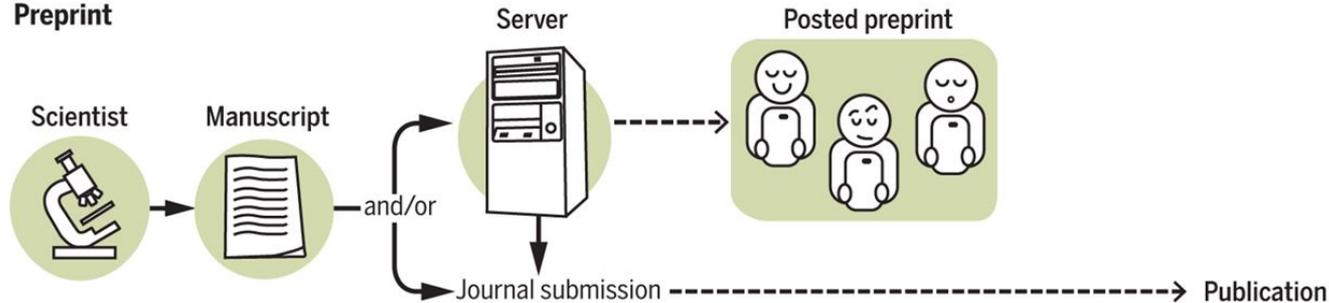
1. Make a list of repositories that you know/use
2. Do you know the difference between them?

But, what exactly are we depositing?

Peer review publication



Preprint



Repositories per area (some of them)



Preprint servers are appearing in many disciplines



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

Institutional Repositories (CONARE - Costa Rica)



¡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)

Otros repositorios institucionales:

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

Preprints are not the only things we can deposit:

- Research Reports
- Teaching materials
- Conference presentations and/or posters
- Images or plots from each of your articles
- Data and code
- A long etc.

Where? Venkat will talk about that this afternoon

Promote your articles

Social Networks can be academic

- Inform interested users on Twitter. Choose appropriate #hashtags to increase the discoverability of your work.
- Post links to articles on Facebook, and any relevant groups, as well academic networking sites, including your LinkedIn profiles.
- Save it to your reference library and promote in the academic network of reference manager sites such as CiteuLike or Zotero.
- Share your research data and code (when possible) and increase your citations.
- Create a Google Scholar profile and track your citations.
- Add articles to your ORCID account.

Promote your articles

Twitter, Facebook, CiteuLike, LinkedIn, Zotero.

Inform society news outlets and bloggers in your field.

What's the difference between ResearchGate, Academia.edu, and the institutional repository?

	Open access repositories	Academia.edu	ResearchGate
Supports export or harvesting	Yes	No	No
Long-term preservation	Yes	No	No
Business model	Nonprofit (usually)	Commercial. Sells job posting services, hopes to sell data	Commercial. Sells ads, job posting services
Sends you lots of emails (by default)	No	Yes	Yes
Wants your address book	No	Yes	Yes
Fulfills requirements of UC's OA policies	Yes	No	No

Twitter

- Academic Twitter is a nice place (not like the rest of Twitter).
- Find researchers from your area.
- You can choose to read twits quietly or to participate and exchange opinions.

[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/waird-and-wonderful-world-academic-twitter>



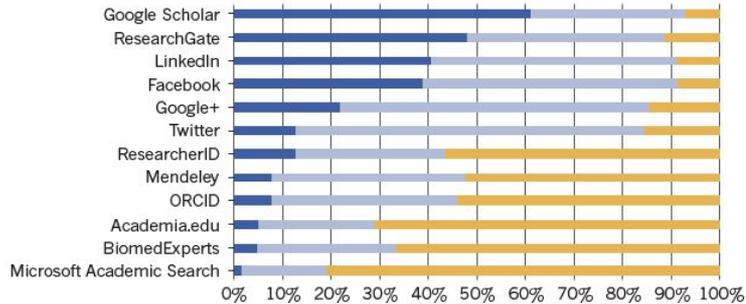
Promote your articles

REMARKABLE REACH

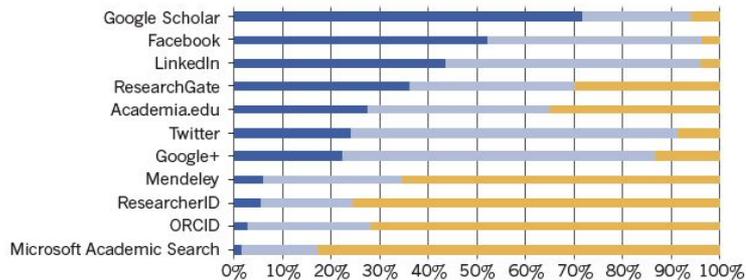
More than 3,000 scientists and engineers told Nature about their awareness of various giant social networks and research-profiling sites. Just under half said that they visit ResearchGate regularly. Another 480 respondents in the humanities, arts and social sciences were less keen on ResearchGate.

- 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

Science and engineering



Social sciences, arts and humanities



Is there one you don't know?
Which ones would you like to use more often?

My favourite combination

ORCID + Webpage + Github + Google Scholar + Twitter

CODATA-RDA School of Research Data Science

		Automatically assigned	Non-Proprietary	Database independent	Can link with Research UNE	Provides metrics
ORCID		✗	✓	✓	✓	✗
Google Scholar Profile		✗	✗	✗ Only Google Scholar	✗	✓ (basic)
Scopus AuthorID		✓	✗	✗ Only Scopus	✓	✓
ResearcherID		✗	✗	✗ Only Web of Science	✓	✓

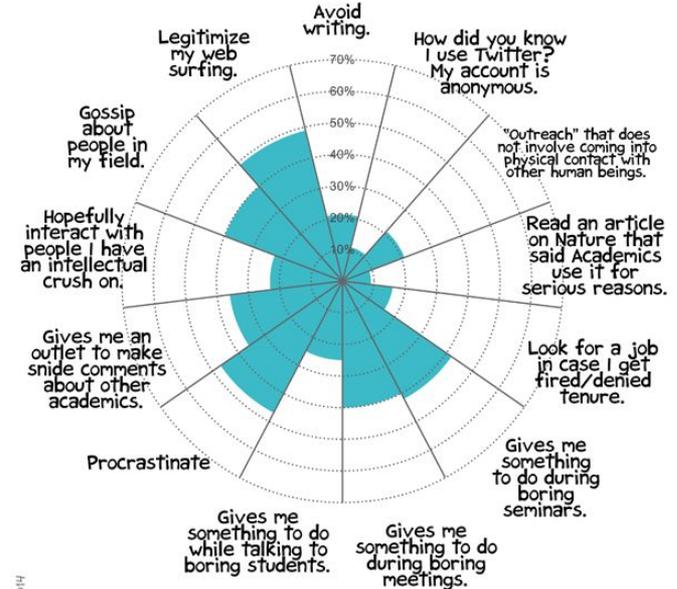
<https://blog.une.edu.au/library/2018/11/26/author-profiles-which-are-right-for-you/>

Promote your articles

Your turn:

In pairs, discuss which combination of academic social networks and academic profile sites is most convenient for your area of research.

Why Academics REALLY Use Twitter



URGE CHAM © 2014

www.phdcomics.com

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

Establish your academic Identity

Why?

- How many people around the world has your same name?
- How can I do it if I don't keep everything in the same place?
- It doesn't only helps you as a researcher, but it also helps your home institution.

How?

- Using [ORCID](#) and from there, update all the other sites.
- A [Google Scholar](#) profile is the most efficient to count citations.

Establish your academic Identity



Connecting Research
and Researchers

<https://vimeo.com/97150912>

You do not have ORCID? Let's do it:

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

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

Already have it? Connect it with science open (if you want):

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

Establish your academic Identity



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

You can only create it if they have publications. However, I highly recommend it because it is the source of one of the web presence rankings of the Universities around the world. In addition, it is a useful tool to control your work citations.

Best practices

- Our research products are not just scientific articles, we have to recognize that it is increasingly common to have many other formats for presenting results.
- In particular, quantitative research needs a good system of attribution, recognition and appointment for non-traditional research products (software, data sets). REPRODUCIBILITY!
- In addition, several studies indicate that articles with available data and code have more citations on average.

How?

- Using repositories and DOIs (Venkat will talk about this)

Identifying good journals

Why?

- *Predatory Journals* are all over Internet, sending us nice offers to publish in their pages.
- These journals are not indexed, probably do not have a reputable editorial board and they will charge us for publication.

How?

- Check this site: <https://predatoryjournals.com/journals/>
- Site built thanks to a community effort, after an information professional was forced to delete his page due to the harassment of these magazines. [Reference](#).

Identifying good journals

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

Journals and publishers engaging in any of the following behaviors will be listed here as possibly predatory:

1. Charging exorbitant rates for publication of articles in conjunction with a lack of peer-review or editorial oversight.
2. Notifying authors of fees only after acceptance.
3. Targeting scholars through mass-email spamming in attempts to get them to publish or serve on editorial boards.
4. Quick acceptance of low-quality papers, including hoax papers.
5. Listing scholars as members of editorial boards without their permission or not allowing them to resign.
6. Listing fake scholars as members of editorial boards or authors.
7. Copying the visual design and language of the marketing materials and websites of legitimate, established journals.
8. Fraudulent or improper use of ISSNs.
9. Giving false information about the location of the publishing operation.
10. Fake, non-existent, or mis-represented impact factors.

Identifying good journals

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

- Now in pairs: check the list.
- Are there any magazines in your area that would have seemed acceptable if you weren't on that list?

 Stop Predatory Journals

Identifying good journals

https://doaj.org/search#.W8_j-BNKjOQ

- Can you find the UCR/CONARE journals?



Reviewing our workflow

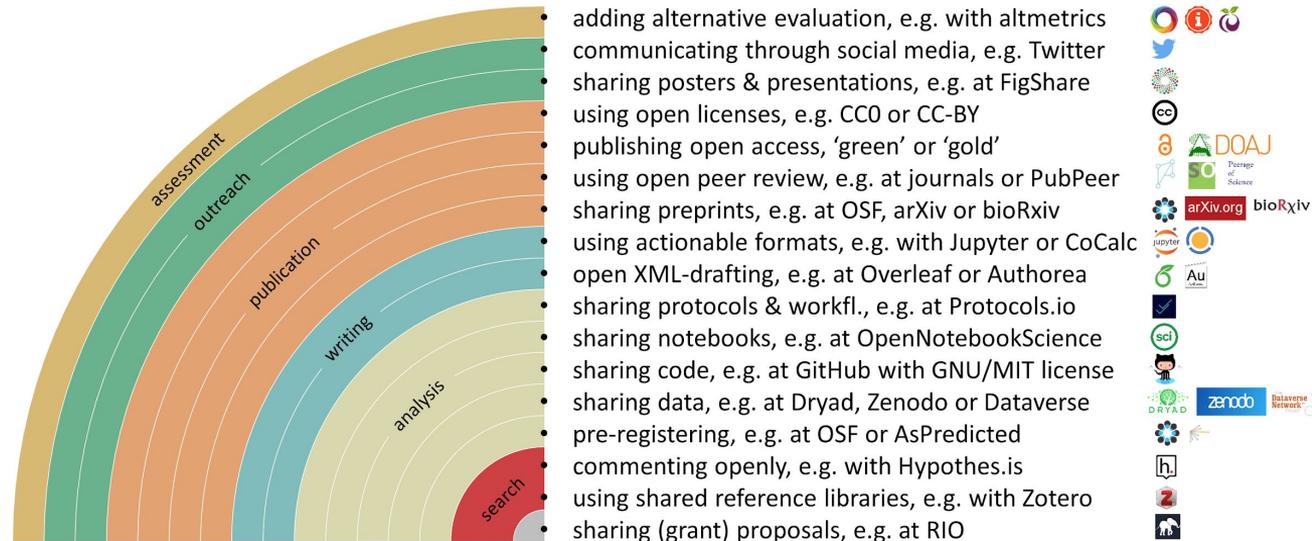
Depósito de Datos y Código:

Some recommendations (Venkat will talk more about this)

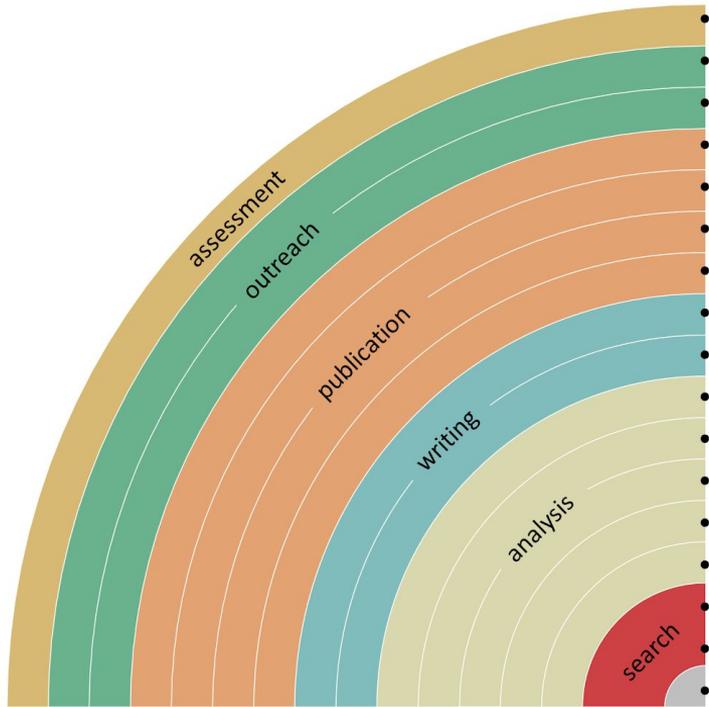
- Código y datos para ciencia reproducible [*A toolkit for data transparency takes shape.*](#) Además, <https://ropensci.org/>, <http://jupyter.org/>

Reviewing our workflow

You can make your workflow more open by ...



You can make your workflow more open by ...

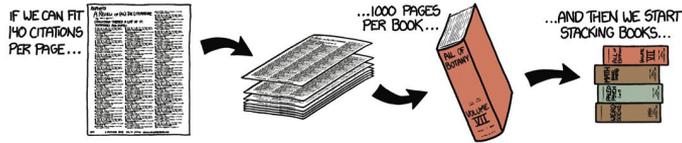


- adding alternative evaluation, e.g. with altmetrics
- communicating through social media, e.g. Twitter
- sharing posters & presentations, e.g. at FigShare
- using open licenses, e.g. CC0 or CC-BY
- publishing open access, 'green' or 'gold'
- using open peer review, e.g. at journals or PubPeer
- sharing preprints, e.g. at OSF, arXiv or bioRxiv
- using actionable formats, e.g. with Jupyter or CoCalc
- open XML-drafting, e.g. at Overleaf or Authorea
- sharing protocols & workfl., e.g. at Protocols.io
- sharing notebooks, e.g. at OpenNotebookScience
- sharing code, e.g. at GitHub with GNU/MIT license
- sharing data, e.g. at Dryad, Zenodo or Dataverse
- pre-registering, e.g. at OSF or AsPredicted
- commenting openly, e.g. with Hypothes.is
- using shared reference libraries, e.g. with Zotero
- sharing (grant) proposals, e.g. at RIO



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?



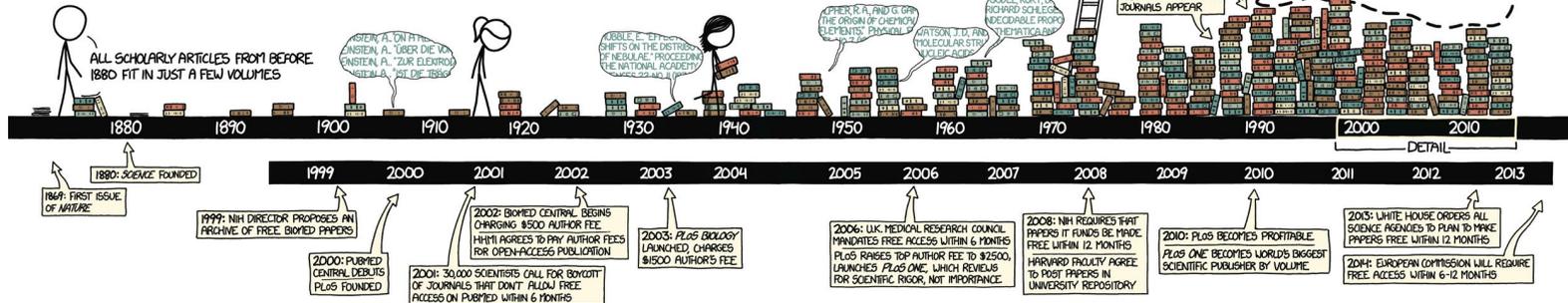
A LIST OF PAPERS PUBLISHED IN 1880 WOULD FILL 100 PAGES.

BY 1920, THE LIST WOULD BE GROWING BY 500 PAGES PER YEAR.

THE 1975 SECTION WOULD FILL FOUR HUGE VOLUMES.

TODAY, WE'RE UP TO 15 VOLUMES PER YEAR—A PAGE EVERY 45 MINUTES.

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



¡Manos a la obra!

OPTION 1

RMarkdown from RStudio:

<https://github.com/bioconnector>
https://www.emilyzabor.com/tutorials/rmarkdown_websites_tutorial.html

Other sources:

- <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>

OPTION 2

BiBTeX + Zotero:

<https://libguides.mit.edu/cite-write/bibtex>

<https://libguides.mit.edu/cite-write/zotero>

OPTION 3

DOIs for Github repos:

<https://guides.github.com/activities/citable-code/>

OR

<https://genr.eu/wp/cite/>

¡Muchas gracias!

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

