

Reproducibilidad y ciencia abierta



© Shutterstock

Laura Gómez Navarro

Investigadora postdoctoral de la Universidad de Utrecht (Países Bajos) e IMEDEA (España)

28/02/23

Índice

- ¿ Que és la ciencia abierta?
- ¿Que acciones se pueden llevar a cabo?
 - Ejemplos
 - 1) OceanParcels
 - 2) Cuenta de Twitter y adrift
 - 3) ¡Este Hackatón!
- Consejos prácticos (reproducibilidad)
 - Documentación
 - Organización
 - Automatización
 - Diseminación

¿Qué es la ciencia abierta?

5 tipos de abierto:

- Código abierto
- Hardware libre
- Acceso abierto
- Datos abiertos
- Recursos educativos abiertos



Bosman & Kramer (2017) Figshare

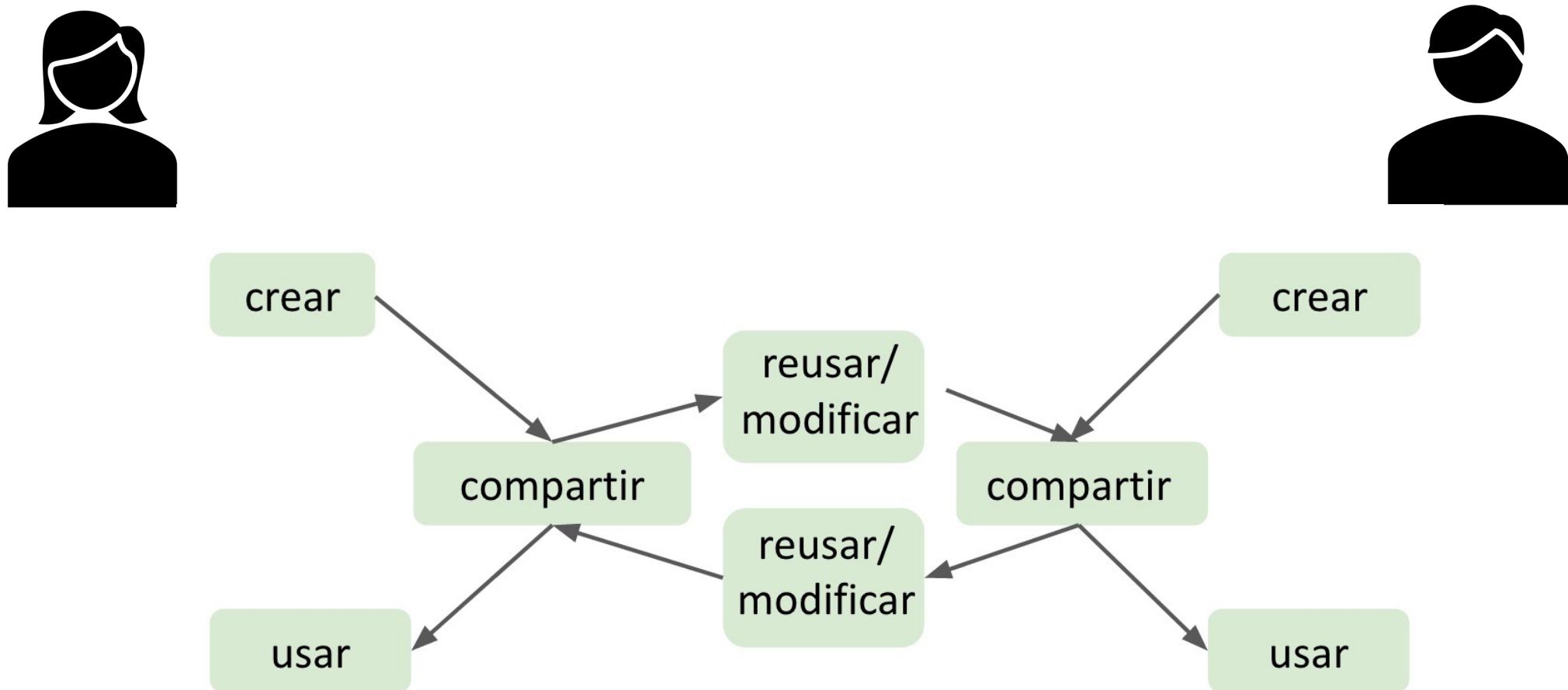
4 sombras se abierto:

- **Acceso** abierto (a datos): ¿puedes conseguirlos?
- **Consumo** abierto: ¿puedes usarlos?
- **Producción** abierta: ¿puedes contribuir?
- **Proceso** abierto: ¿puedes ver de dónde viene?



Hooft & Smeele (2017) Landelijk Coordinatiepunt RDM

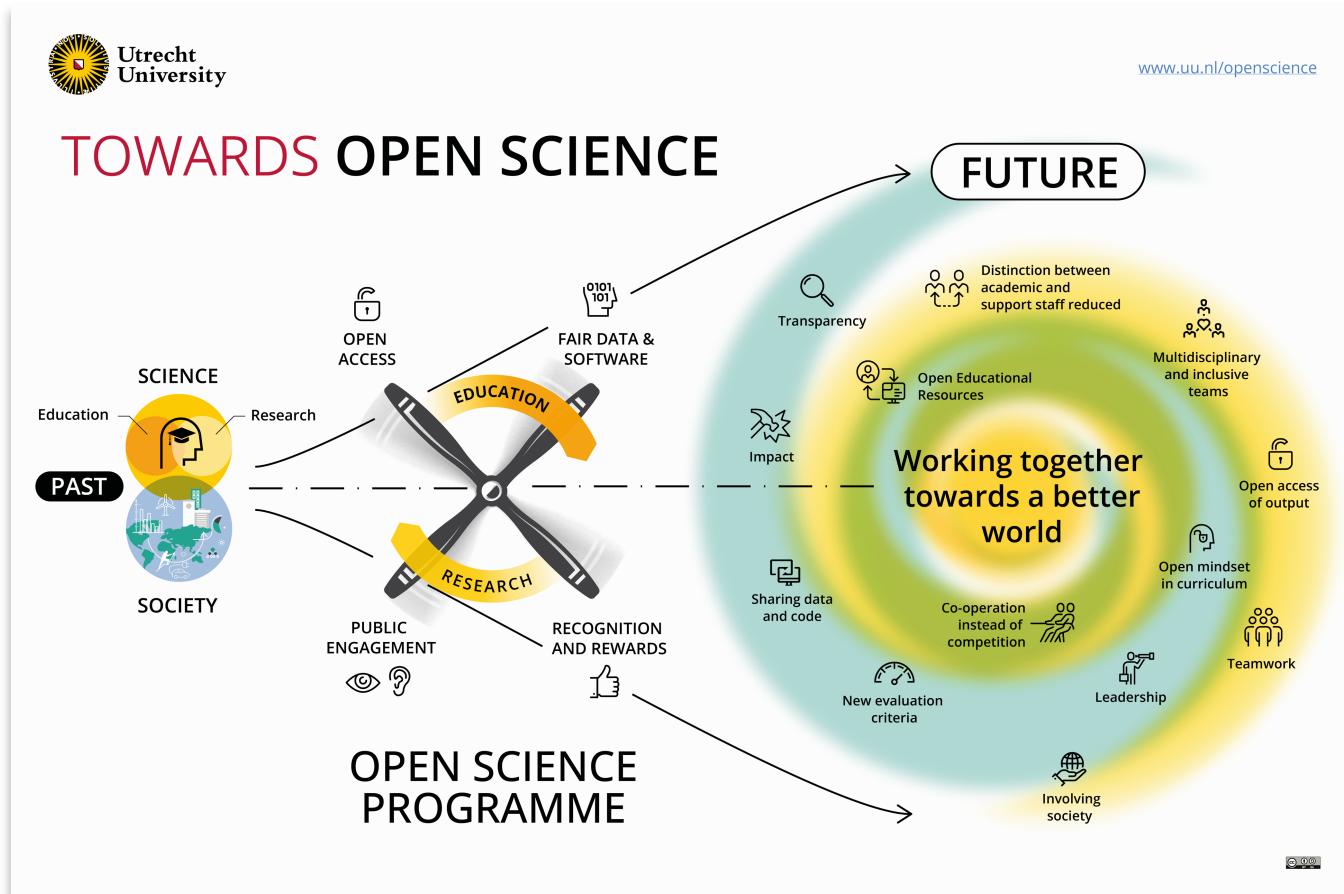
El concepto de Ciencia Abierta



¡Pregunta!

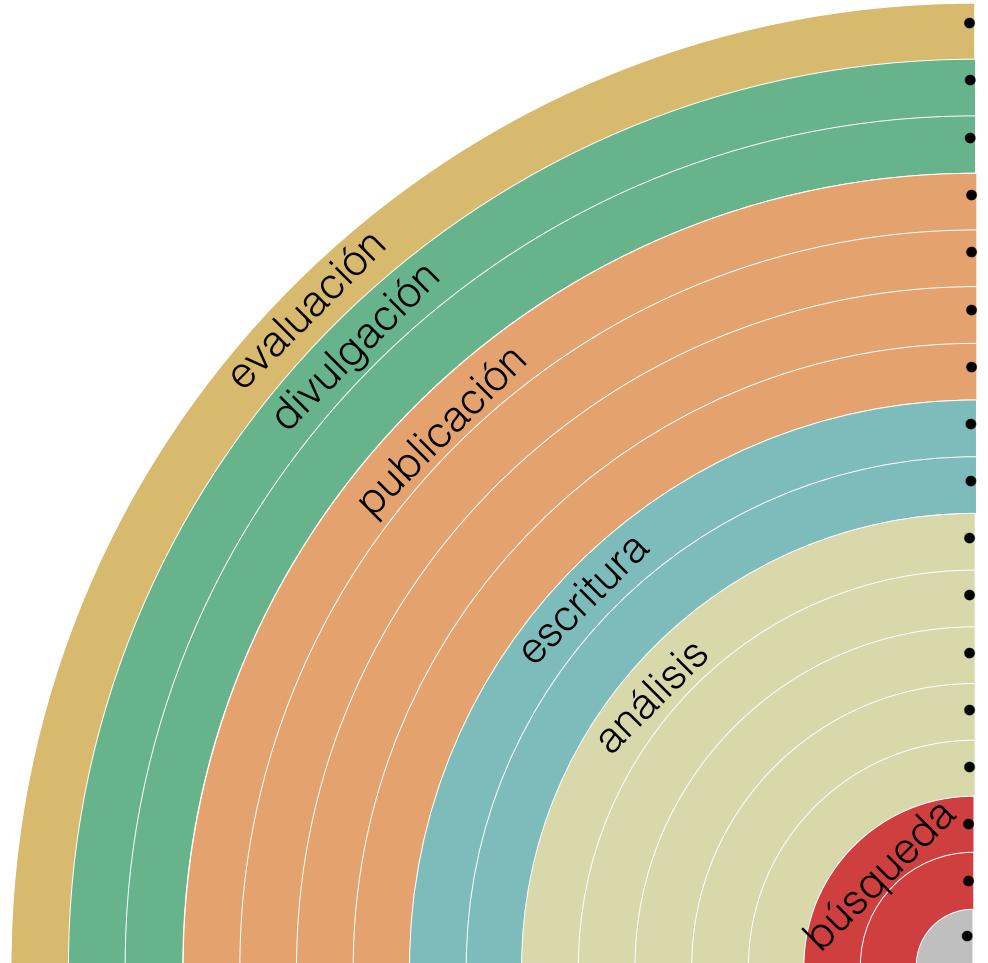
**¿Que actividades realizan en su lugar de trabajo
en relación a la ciencia abierta?**

Caminos hacia la ciencia abierta



- Impacto
- Transparencia
- Reducción de la diferenciación entre el personal académico y de apoyo
- Multidisciplanriedad y equipos inclusivos
- Acceso abierto a los resultados
- Trabajo en equipo
- Involucrar a la sociedad
- Nuevos criterios de evaluación
- Liderazgo
- Mente abierta en el curriculo
- Recursos educativos en abierto
- Compartir datos y código
- Cooperación en vez de competición!

Acciones ciencia abierta



evaluación alternativa
comunicación via redes sociales
compartir posters y presentaciones
uso de licencias abiertas
publicación en acceso abierto
uso de *peer review* en abierto
compartir *preprints*
uso de formatos accionables
redacción XML abierta
compartir protocolos y flujos de trabajo
compartir *notebooks*
compartir código
compartir datos
pre-registro
comentar abiertamente
uso de bibliotecas de referencias compartidas
compartir *grant proposals*



Otras acciones de Ciencia Abierta

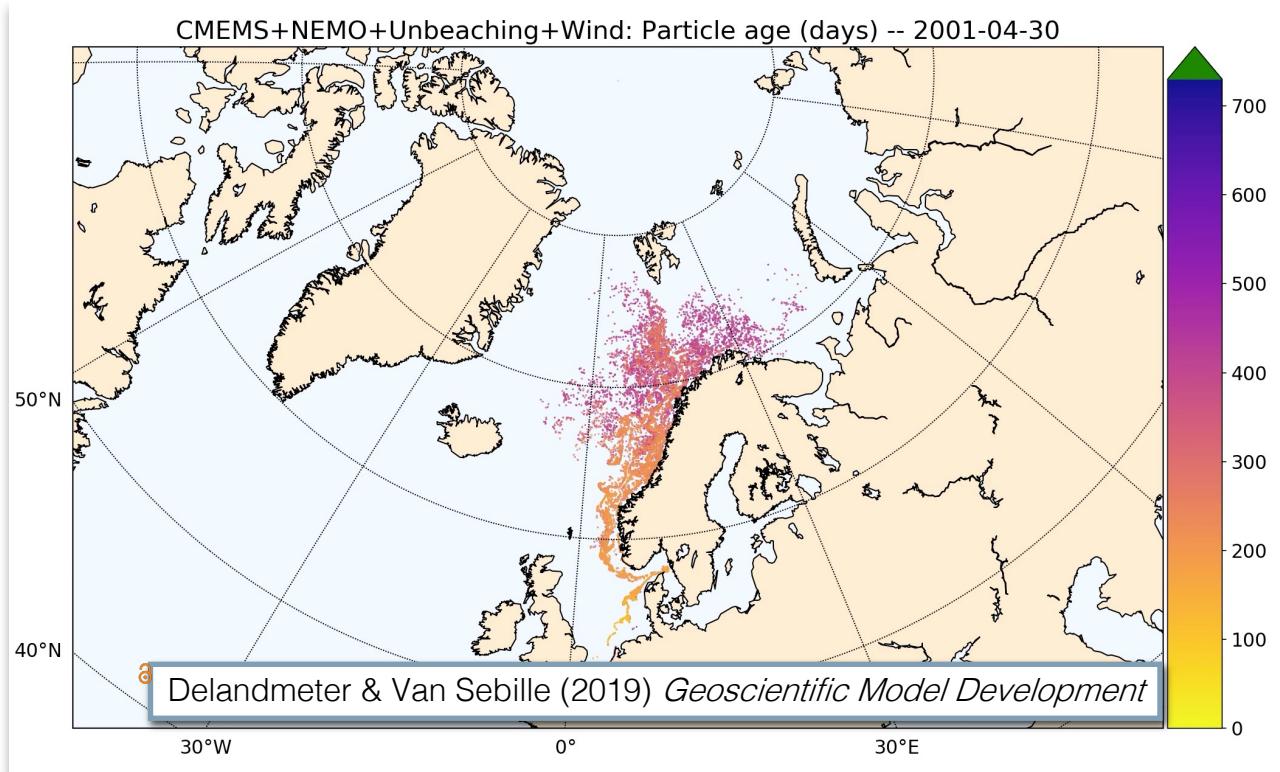


Algunos ejemplos:

- Editar artículos Wikipedia
- Reclamar crédito por la revisión de artículos
- Comunicar resultados analizados con personas expertas, no expertas y al público en general
 - Explicarlos en la TV o radio
 - Compartirlos en la redes sociales (y reaccionar al trabajo de otras personas)
- Presentar a niños y niñas en colegios
- Usar software abierto (y fácilmente usable por todo el mundo)
- Negarse a participar en paneles con solo hombre o personas blancas
- Traducir material científico a otros idiomas

Ejemplo 1: OceanParcels

- Análisis oceánico Lagrangiano:
Seguimiento de partículas virtuales en
simulaciones oceánicas en 3D
- Parcels: “Probably A Really
Computationally Efficient Lagrangian
Simulator” (Probablemente un simulador
Lagrangiano muy eficiente
computacionalmente)
 - Consiste en un conjunto de clases
python y métodos para la construcción de
modelos de partículas Lagrangianas



Ejemplo 1: OceanParcels

Concepto detrás del software

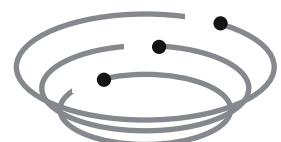
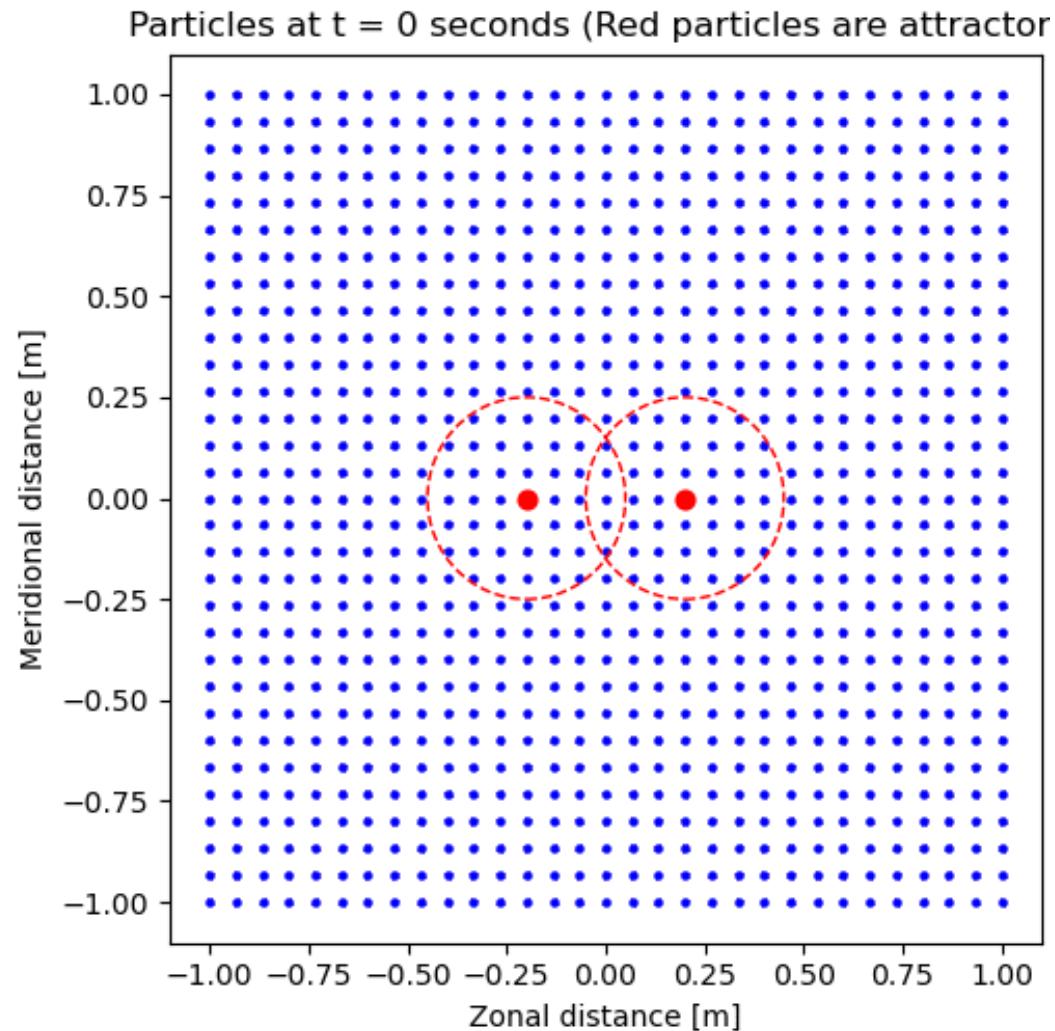


- Escribir buen software (código) es difícil! Las personas científicas a menudo se centran en resolver un problema en particular
 - El software científico se construye gradualmente con un diseño general limitado
 - La optimización del rendimiento suele ser una idea de última hora
- ¿Por qué deberían las personas científicas saberlo todo sobre ingeniería informática?
 - Involucrar tanto personas científicas como analistas numéricas y expertas en programación
- Nadie puede saberlo todo desde ciencia de alto nivel académico hasta la optimización de código!



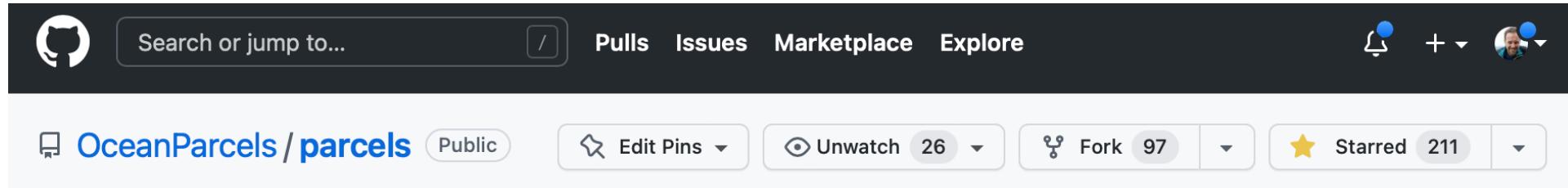
Ejemplo 1: OceanParcels

Concepto detrás del software: Interacción partícula-partícula



Ejemplo 1: OceanParcels

Parcels es usado extensamente



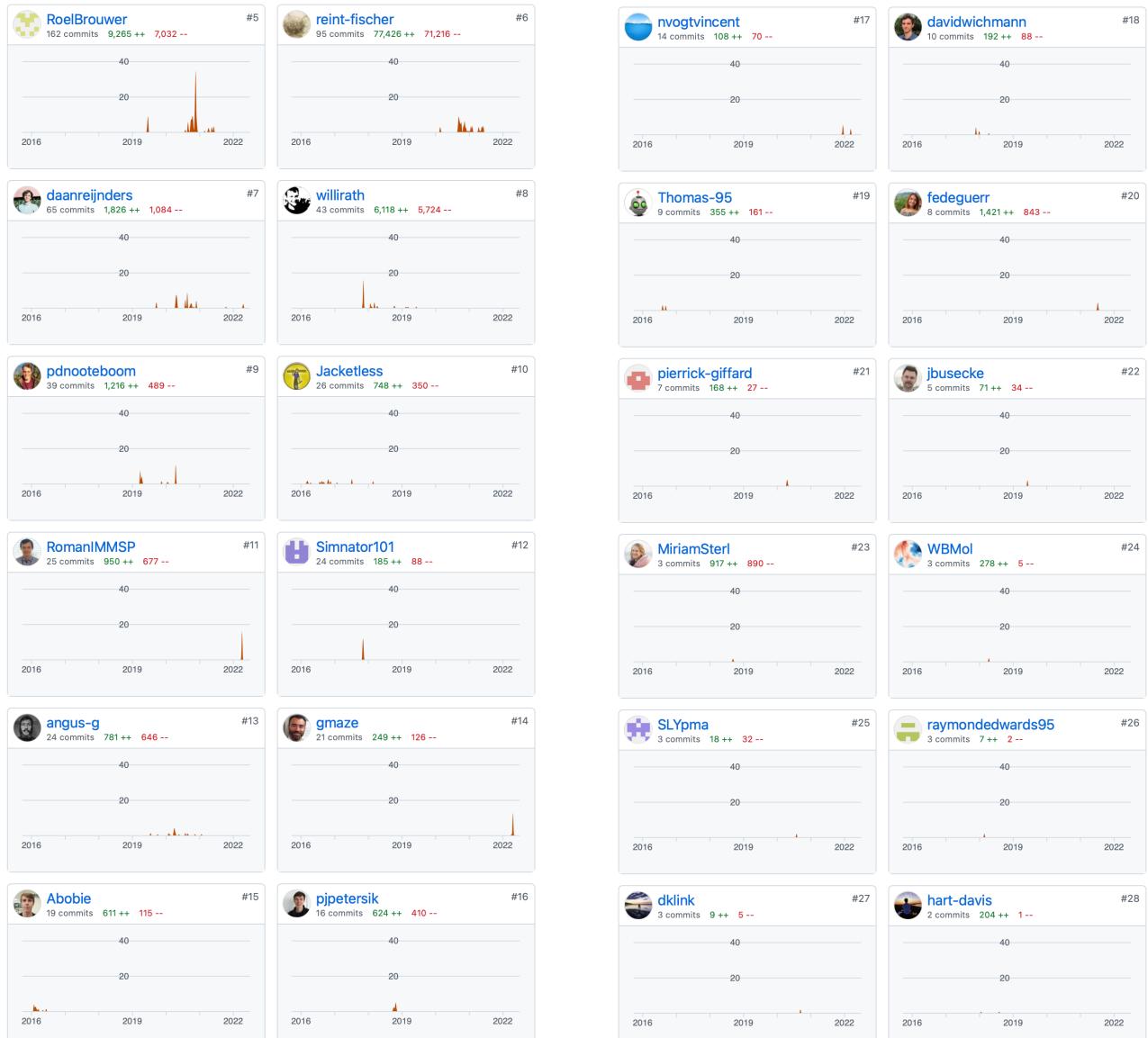
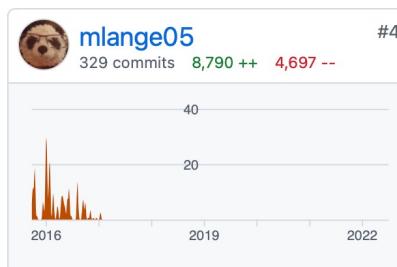
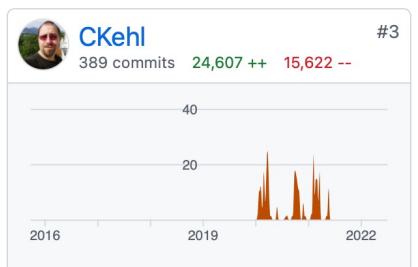
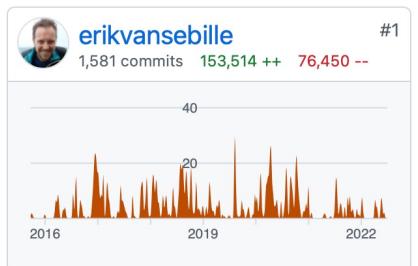
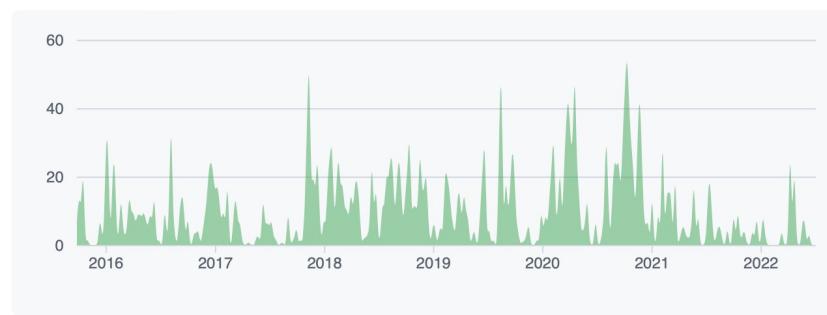
- Más de 80 artículos *peer-reviewed* desde 2017
- Usado al menos por una docena de grupos de investigación alrededor del mundo
 - Desde ecología a biología, matemáticas, arqueología o ingeniería
- Más de 400 *issues* reportados en GitHub
- Licencia MIT (la más permisiva)

Ejemplo 1: OceanParcels

Sep 27, 2015 – Jul 5, 2022

Contributions: Commits ▾

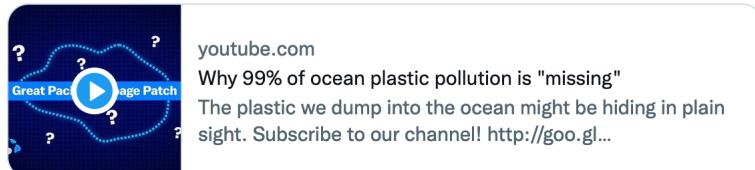
Contributions to master, excluding merge commits and bot accounts



Ejemplo 2: Twitter



Huge thanks to [@voxdotcom!](#) We are thrilled that [@ErikvanSebille](#) and our twitter page featured in this [#vox](#) video about [#ocean #plasticpollution!!!](#)



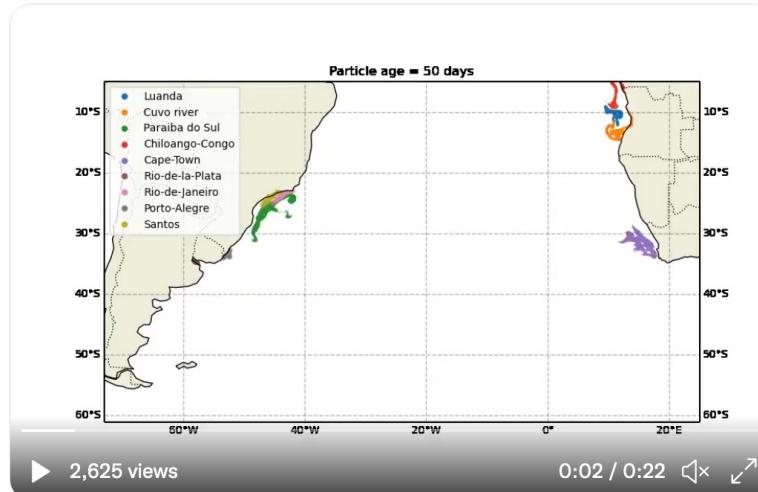
11:55 AM · Apr 28, 2021 · Twitter Web App

53 Retweets 5 Quote Tweets 236 Likes

@UFollowtheOcean



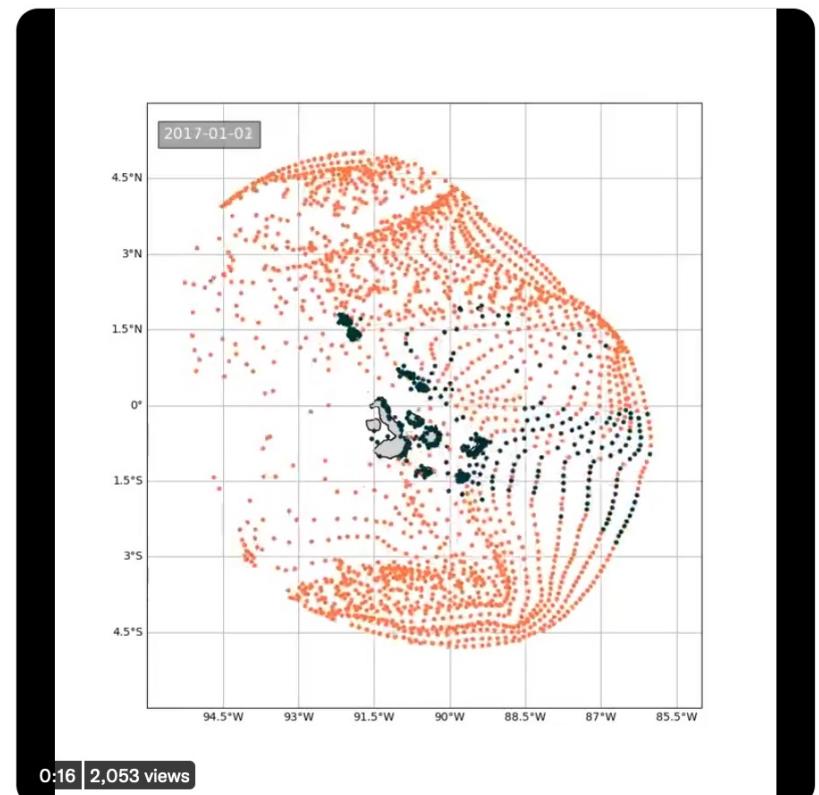
Animation of plastic particles emitted by some of the principal rivers and cities in the South Atlantic. [@cpierard](#) is working on estimating the most likely sources of floating plastic found in this region.



11:35 AM · Mar 19, 2021 · Twitter Web App

60 Retweets 10 Quote Tweets 154 Likes

Lagrangian particle animations are always fun to watch! Check out how these virtual particles get stuck on the [#Galapagos](#) Islands and find out more at the upcoming [#EGU21](#) session ITS2.5/OS4.8 ([meetingorganizer.copernicus.org/EGU21/session/...](http://meetingorganizer.copernicus.org/EGU21/session/))! [#plasticfree #plasticpollution](#)



10:51 AM · Apr 16, 2021 · Twitter Web App

24 Retweets 1 Quote Tweet 84 Likes

¡Re-Pregunta!

**¿Que actividades realizan en su lugar de trabajo
en relación a la ciencia abierta?**



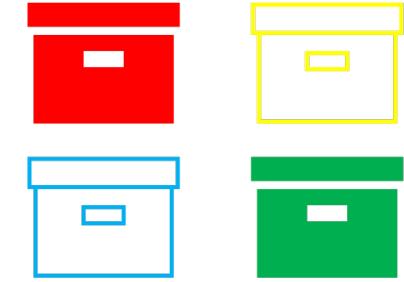
Consejos prácticos
(subjetivo)

Consejos prácticos: Documentación



¡Documentar todo!

1. Notas electrónicas (en el PC)
 - a. Herramientas de notas: por ej. Evernote, OneNote (Windows), Notes (mac)
 - b. Interactivas: por ej. Rmarkdown, Jupyter
 - c. Herramientas de pago: por ej. All-singing-all-dancing
2. Documentos de texto
 - a. Editores de texto: por ej. Notepad
 - b. Procesadores de texto: por ej. Word, Google docs



Consejos prácticos: Organización

Piense en la organización de ficheros **antes** de empezar su proyecto

3 puntos a tener en cuenta:

1. Nombres consistentes, legibles por ordenador, , compresibles humanamente (para ficheros, nombres de bases de datos, etc...)
 - a. Sin espacios (usar mejor _ o -)
 - b. ¿Quiere ordenar los ficheros? Considere enumerarlos, o usar fechas en formato AAAAMMDD
2. Haga que sea fácil encontrar datos, ficheros, documentos, figuras,...
→Tanto para si mismo/a como para un compañero/a
3. Encuentre algo que funcione para usted:
¡no hay respuesta correcta!

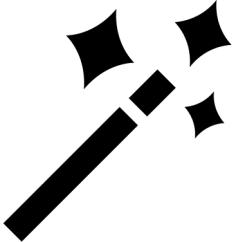
Consejos prácticos: Organización

Expressive Project	Non Expressive Project
/01-scripts	
01-process-landsat-data.py	
02-calculate-ndvi.py	
03-create-ndvi-maps.py	
/02-data	
/raw-data	work.py
/landsat-imagery	plotting.py
/june-2016	plotting-test.py
/july-2016	landsat/
/cold-springs-fire-boundary	data-file.txt
/03-output-graphics	old-stuff/
ndvi-map-june-2016.png	testoutput1.txt
ndvi-map-july-2016.png	testoutput2.csv
/04-final-paper	
veg-impacts-cold-springs-fire.pdf	

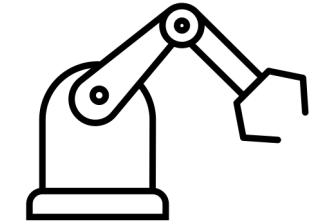
ProjectName/

- README.MD
- Dataset/
 - Raw Data/
 - Processed Data/
 - YYYY-MM-DDVersion
 - YYYY-MM-DDVersion
- Analysis (or Code)/
 - Data cleaning/
 - Data preprocessing/
 - Output/
 - Graphs
 - Tables
- Publications/
 - .tex files
 - .bib file

Consejos prácticos: Automatización



- Pensar en desarrollar métodos en partes pequeñas y reusables
 - Por ej. Para limpiar datos, crear gráficos, hacer análisis, correr un tipo de simulación
- Seleccionar la mejor herramienta para hacer la tarea que quiera
- Si trabaja con hojas de cálculo:
 - Tener un archivo README junto con el fichero para explicarlo
 - Explicar cómo se han creado las figuras
- Si trabaja haciendo código:
 - Intentar usar funciones, especialmente para tareas que repita muchas veces
 - Tener un código individual donde cree un gráfico (intentar añadir todos los detalles ahí sin depender de post-ediciones por ej. mediante Photoshop)



Consejos prácticos: Diseminación (divulgación y difusión)



- Algunas formas en las que compartimos nuestras investigaciones: artículos científicos, presentaciones orales, presentaciones poster
 - Usar bibliotecas de referencias compartidas para mantener un seguimiento de las citas (por ej. **Zotero**, EndNote, Mendeley)
 - Hacer que sea sencillo rehacer / modificar figuras. Por ej. Puede referenciar el código donde creó la figura en los documentos donde la use (por ej. Powerpoint, overleaf, ...)
- Otras cosas en las que pensar:
 - A menudo también tenemos que compartir nuestros datos en repositorios en línea (también se pueden publicar bases de datos como artículos)
 - Para que nuestro estudio sea realmente reproducible, nuestros métodos también tienen que ser compartidos, por ej. Compartiendo el código (Github) o el flujo de trabajo desarrollado

¡Re-Pregunta!

**¿Alguna otra buena práctica de
Ciencia Abierta a compartir?**

Más información?

Online resources

<https://www.earthdatascience.org/courses/intro-to-earth-data-science/open-reproducible-science/get-started-open-reproducible-science/best-practices-for-organizing-open-reproducible-science/>

<https://guides.lib.berkeley.edu/c.php?g=652220&p=4575532>

https://kbroman.org/Tools4RR/assets/lectures/06_org_eda_withn_otes.pdf

<https://openrefine.org/>

23:37 part on joseph Gum

<https://www.youtube.com/watch?v=YB3-zuolzCc>



¿Más información?

- **Github** : <https://github.com/Intercoonecta/Aula-invertida/tree/main/Intro-a-github>
- <https://es.unesco.org/fieldoffice/montevideo/DerechoALaCiencia/CienciaAbierta>
- En inglés:
 - <https://www.earthdatascience.org/courses/intro-to-earth-data-science/open-reproducible-science/get-started-open-reproducible-science/best-practices-for-organizing-open-reproducible-science/>
 - <https://openrefine.org/>
 - https://kbroman.org/Tools4RR/assets/lectures/06_org_eda_withnotes.pdf
 - <https://guides.lib.berkeley.edu/c.php?g=652220&p=4575532>

Agradecimientos y referencias

- Erik van Sebille
- Felix Weijdema & Jeroen Bosman
- Emilio Mayorga
- Nick Record :
 - <https://www.youtube.com/watch?v=YB3-zuolzCc>
 - <https://github.com/SeascapeScience/reproducibility>