

Simplificación Intercambio de Datos Investigación

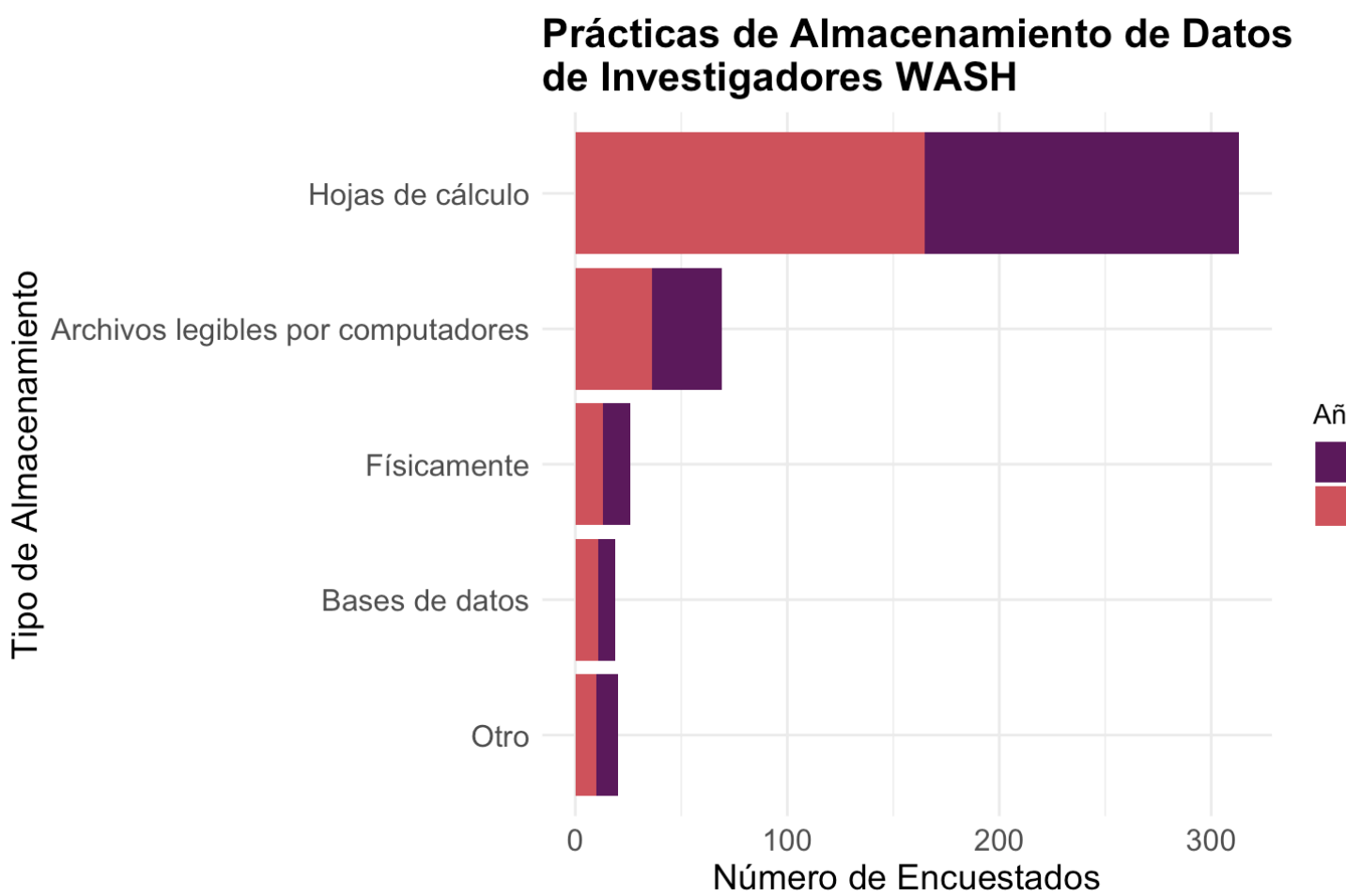
Adriana Clavijo Daza et al
Global Health Engineering (GHE)
ETH Zurich

December 4, 2025

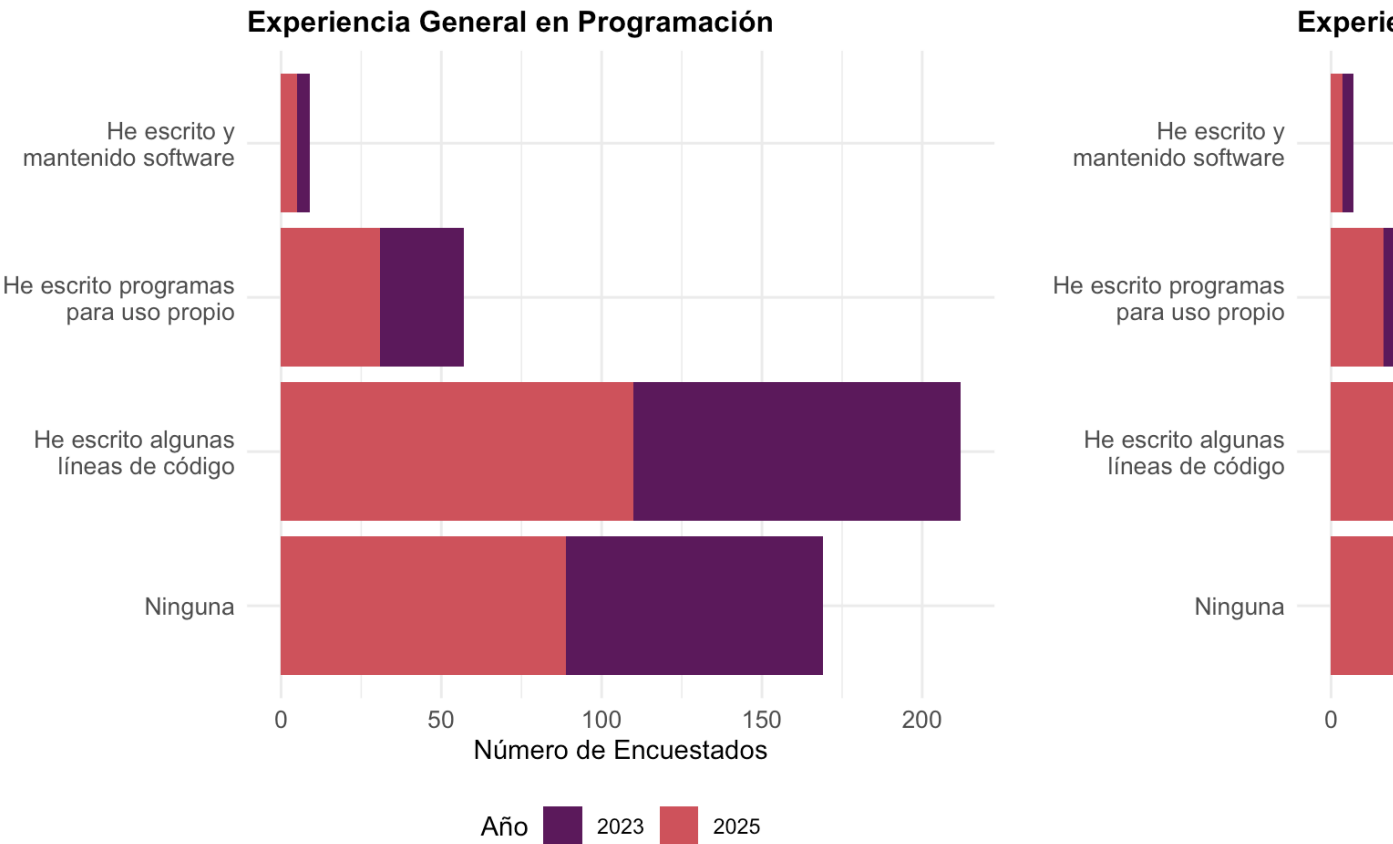
Contexto

- Colaboración desde la academia con **investigadores WASH** (water, sanitation and hygiene) en países con recursos limitados
- Conjuntos de datos valiosos son **desaprovechados** al no ser utilizados
- Familiaridad limitada con los **principios FAIR** de datos (Wilkinson et al., 2016)
 - Localizables (Findable)
 - Accesibles (Accessible)
 - Interoperables (Interoperable)
 - Reutilizables (Reusable)
- El trabajo requerido para generar y recolectar datos no recibe el reconocimiento que merece

Prácticas Actuales de Almacenamiento de Datos



Niveles de Experiencia en Programación



Retos y Soluciones

Retos

- Falta de conocimiento de herramientas accesibles
- Curva de aprendizaje pronunciada
- Flujos de trabajo no estandarizados
- Reconocimiento solo a quienes publican un artículo académico

Soluciones

- Minimiza
- Interfaces usuario
- Guía deta publicaci
- Énfasis en apropiad

Nuestro Recorrido: washr

- Primera iteración: paquete **washr** (Zhong
- Flujo de trabajo simplificado de datos → p
- Basado en utilidades de **devtools**
- Guía completa de publicación (**Walder, S**
2025)

Pero podíamos hacerlo mejor...

Presentando fairenough

github.com/openwashdata/fairenough

Un paquete de R para publicación de datos
principios FAIR

- Mínima entrada del usuario requerida
- Documentación generada por LLMs (opcional)
- Un solo comando para obtener el paquete
datos ordenados (tidy)

Veamos fair enough en

fairenough: Control Granu

1. **setup()** - Inicializa estructura de paquete
2. **process()** - Lectura y validación automática
(`auto_clean = TRUE`)
3. **collect()** - Recolección interactiva de n
4. **generate()** - Diccionarios de datos generados
`ellmer`
5. **build()** - Infraestructura completa de pa
`Roxygen`, archivos de información de cita
web

Principios FAIR

- **Localizables:** Generación integral de metadatos
- **Accesibles:** Paquete R + sitio web + múltiples formatos (CSV, XLS)
- **Interoperables:** Formatos legibles por máquinas
- **Reutilizables:** Licenciamiento y atribución

Además: Facilita la asignación de DOI a través de GitHub con Zenodo

Gracias!

Autores

Adriana Clavijo Daza, Lars Schöbitz, Prof. I

Colin Walder, Nicolo Mass

Global Health Engineering (GHE), I

Yash Dubey, Mian Zhong

ETH Zurich

Paquete

[github.com/openwashdata/
fairenough](https://github.com/openwashdata/fairenough)

g
open

Referencias

- Walder, Colin, Lars Schöbitz, and Yash Dubey. 2025. “ghedat”
doi.org/10.5281/zenodo.1234.
- Wilkinson, Mark D., Michel Dumontier, IJsbrand Jan Aalbers
Myles Axton, Arie Baak, Niklas Blomberg, et al. 2016. “T
for Scientific Data Management and Stewardship.” *Scie*
<https://doi.org/10.1038/sdata.2016.18>.
- Zhong, Mian, Margaux Götschmann, Colin Walder, and Lars
Publication Toolkit for Water, Sanitation and Hygiene (W
doi.org/10.32614/CRAN.package.washr.