Show me the data*

Rich Pauloo

PhD (2020) | Hydrogeologist | Data Scientist

*clean, standardized, documented, data in an API for sustainable groundwater management

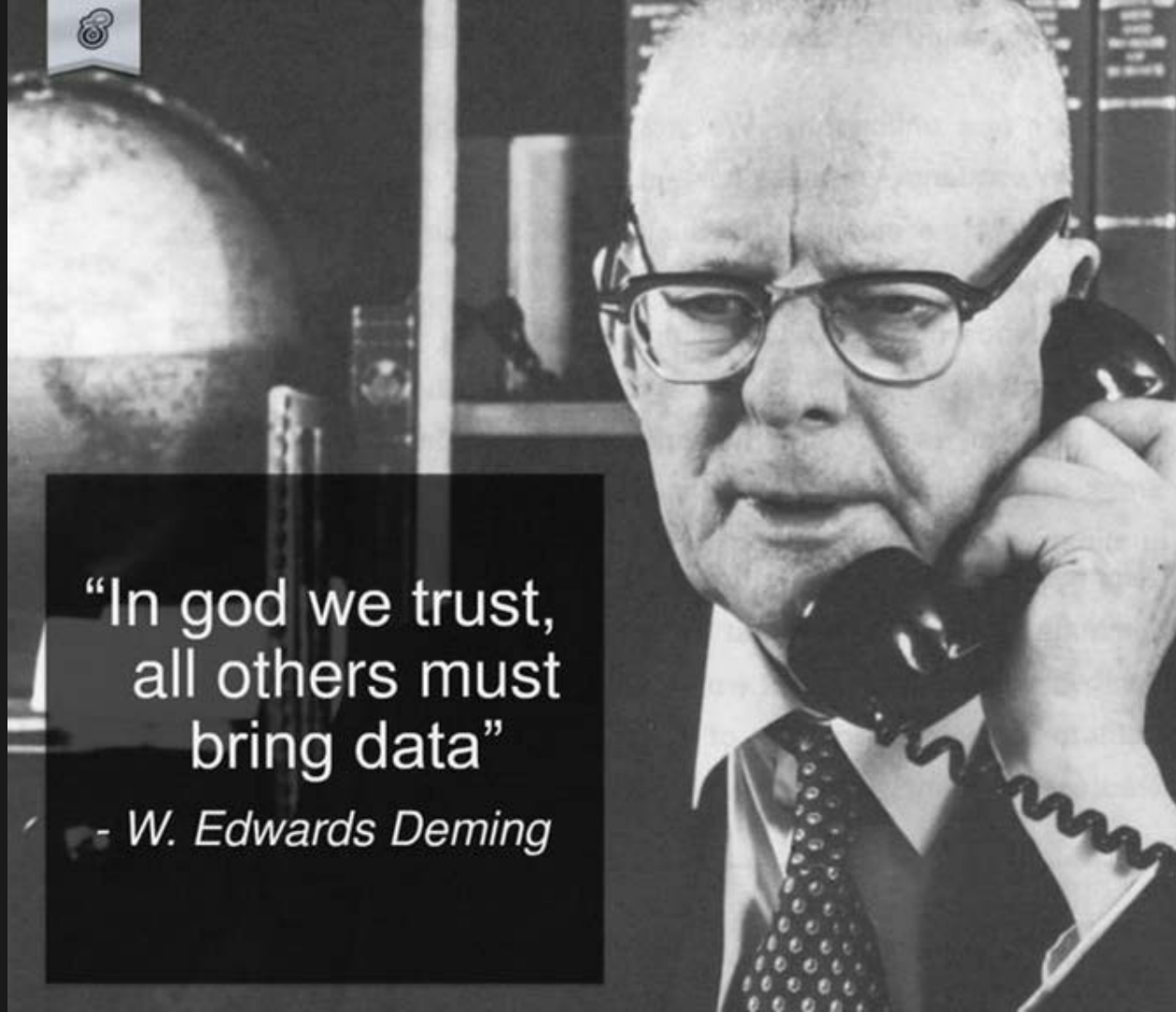
Team

Helen Dahlke
Graham Fogg
Alvar Escriva-Bou
Amanda Fencl
Hervé Guillon

UCDAVIS

NSF

WEST BIG DATA INNOVATION HUB



Agenda

1. The Challenge
2. Open data alone won’t save us
3. Domestic well failure
4. Transformative technologies
5. Take away message

2040

The Challenge: 5,000 days to groundwater sustainability

7,437

The Challenge: 5,000 days to groundwater sustainability

5,298

The Challenge: 5,000 days to groundwater sustainability

Why [open] data?

Data of the right **type** and **amount**

constrains our uncertainty

of the systems we manage in order to

make critical decisions.

338 GSAs

making critical decisions

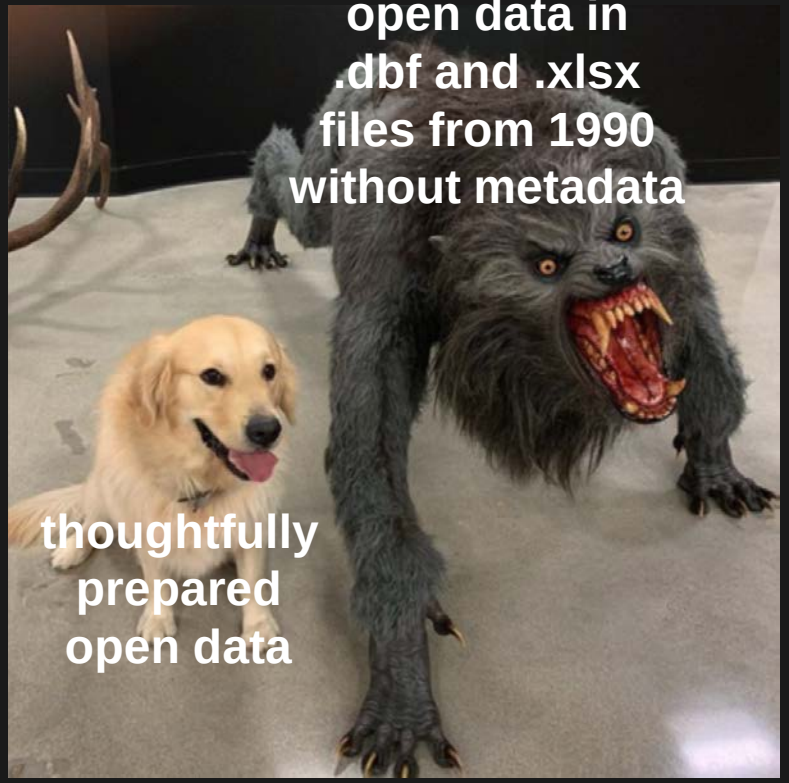
Open data alone won’t save us

Open data isn’t enough

We need to **use** data to constrain uncertainty and make critical decisions.

Thoughtfully prepared data:

- Clean
- Standardized schemas for optimized queries
- Clear documentation
- API



> Open data alone won’t save us

bit.ly/drywells

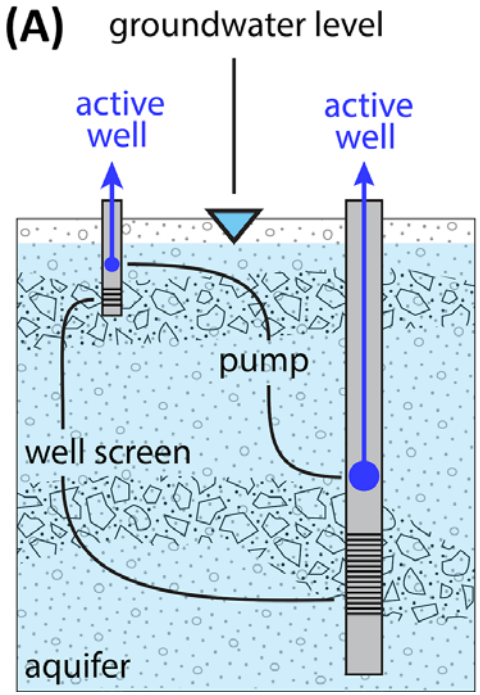
Problem: 2,027 dry wells during 2012-2016 drought, impacting underserved communities.

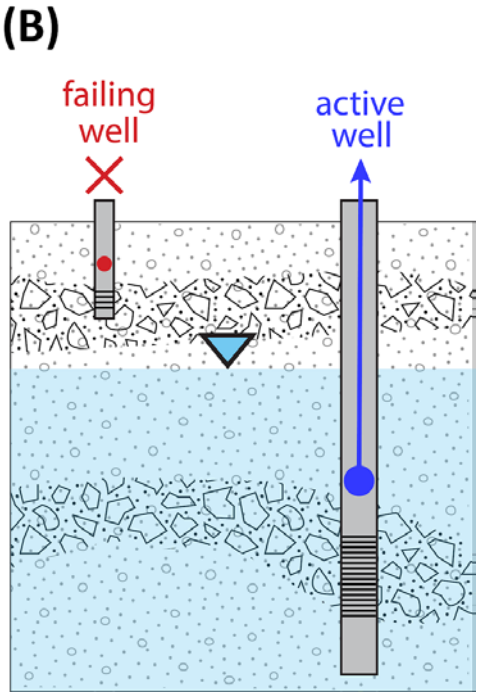
Approach: use open data, AI, and domain knowledge to predict observed well failures, and forecast future well failures.

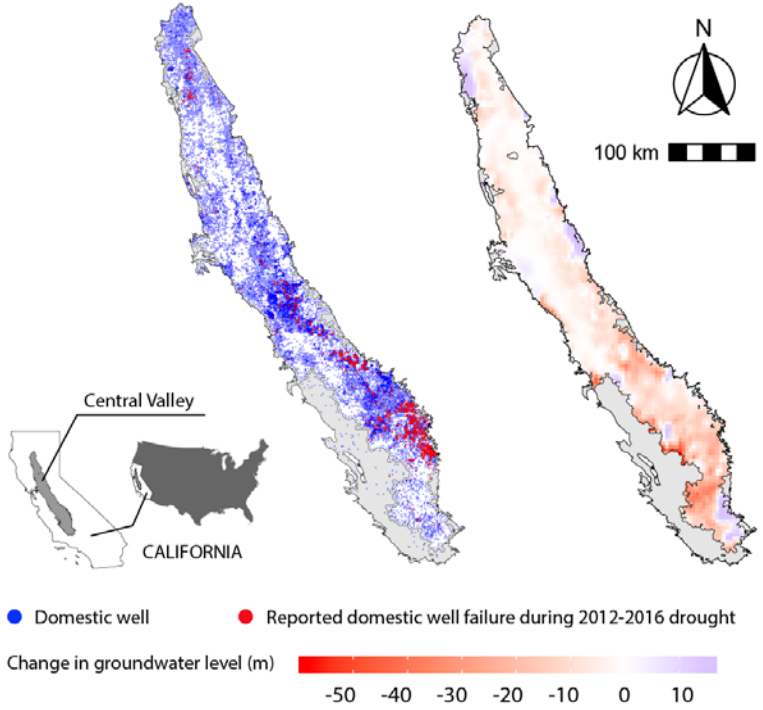
Donna Johnson, 70 (L) lifts pallets of donated bottled water from the back of her truck during her daily delivery run to residents whose wells have run dry, with resident Gabriel Tapia, 31, in Porterville, California October 14, 2014. Picture taken October 14, 2014. Photograph: Reuters/Lucy Nicholson.

Domestic well failure: an example of ***open data for groundwater sustainability***

How does a well go dry?

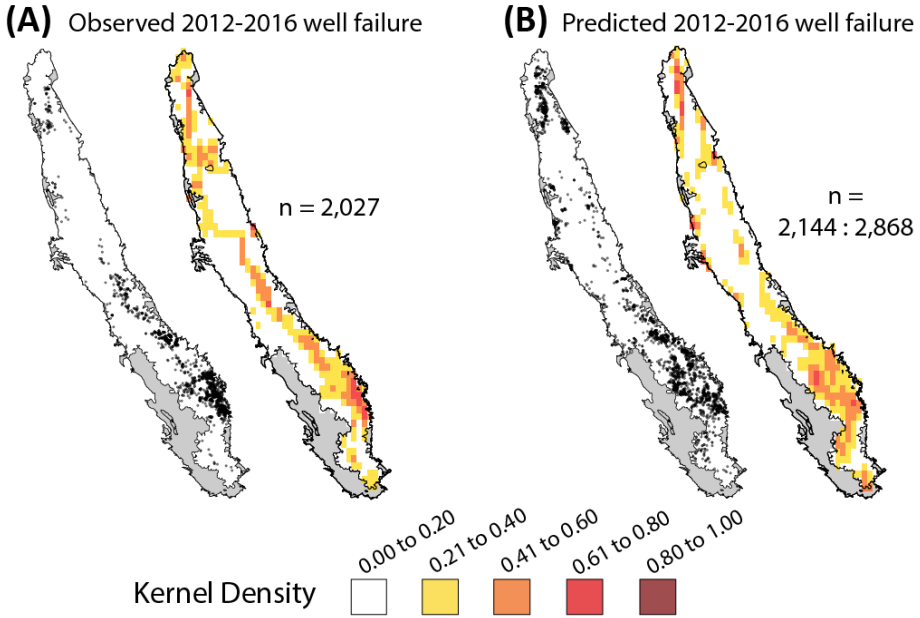






Domestic well failure: an example of ***open data for groundwater sustainability***

Dry well predictions



Domestic well failure: an example of ***open data for groundwater sustainability***

Data Debt

	Open	*Clean*	*Standardized*	*Metadata*	*API*
Well construction	✓	✗	✗	✗	✗
Observed dry wells	✗	✓	✗	✗	✗
Groundwater level	✓	✓	✓	✓	✗

Domestic well failure: an example of ***open data for groundwater sustainability***

Technologies that will transform sustainable groundwater management

- APIs (data infrastructure)
- AI
- Internet of things (IoT)
- Cloud computing
- Remote sensing
- Blockchain (smart contracts)

Transformative technologies

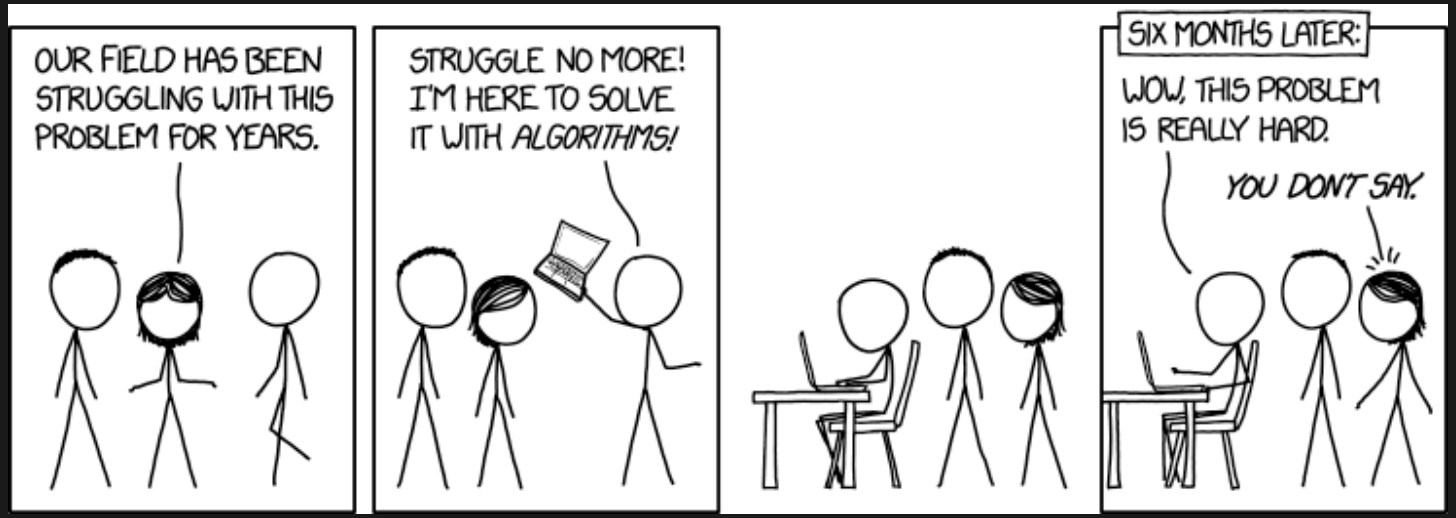
Technologies that will transform sustainable groundwater management

- APIs (data infrastructure)
- ~~AI~~
- ~~Internet of things (IoT)~~
- ~~Cloud computing~~
- ~~Remote sensing~~
- ~~Blockchain (smart contracts)~~

Beware

All the data and AI in the world won’t replace domain knowledge.

Domain knowledge alone won’t save us either. Combining the two will.



Transformative technologies

Summary

- 5,000 days to sustainable groundwater management in California. Make sound data decisions now to improve outcomes tomorrow.
- Open data is a good start, but it’s not enough.
- Avoid data debt by making useful data. Useful data is clean, standardized, documented, and accessible via programming languages (API).
- Neither Data/AI or domain knowledge alone is enough. We need both.
- We must build a culture of data literacy to achieve groundwater sustainability.

Take away message

Thank you.

richpauloo.github.io

@RichPauloo

bit.ly/drywells