

RICH PAULO, PHD

I'm a data scientist at LWA and spend most of my time in R automating ETL pipelines for sensor networks 📡, building Shiny Apps and dashboards 📊, designing approaches with spatial statistics and hydrologic models, and generally wrangling lots of data.

I have a PhD in Hydrology and my dissertation is titled '*Emerging consequences of regional-scale aquifer depletion: data-driven and numerical models of well failure, basin salinization, and contaminant transport*' (my exit seminar can be viewed here¹). Early in my PhD, I found that I really enjoyed data science and programming, and I used these years to sharpen those skills. My published research includes NLP and network analysis², spatial statistics³, and physical modeling of 3D, subsurface contaminant transport⁴.

I'm an #rstats nerd and automation/reproducibility fanatic. My favorite tools include tidyverse (dplyr, ggplot2, purrr), shiny, flexdashboard, plotly, DT, RMarkdown (for dashboards/reporting), sf, sp, raster, leaflet (for spatial data), and DBI for databases. A few projects I'm proud of include an R package to query water quality data 🗄️⁵, R data science curriculum 📖⁶, a dashboard that makes millions of water quality observations understandable 📈⁷, and a model that predicts the risk of wells going dry 💧⁸ funded by Microsoft's AI for Earth Grant.

🎓 EDUCATION

2020
|
2015



PhD, Hydrogeology

University of California Davis

📍 Davis, CA

- Published 5 scientific papers (3 first-author).
- Won ~\$153,000 in national, competitive grants and awards from NASA, Microsoft AI for Earth, AGU, and others.

2011
|
2006



B.S., Integrative Biology (minor in Conflict Resolution)

University of California Berkeley

📍 Berkeley, CA

- Delivered departmental commencement speech⁹ to ~5,000 people.



PROFESSIONAL & RESEARCH EXPERIENCE

present
|
2020



Data Scientist + Hydrologist

Larry Walker Associates

📍 Berkeley, CA

- Programmed automated ETL pipelines for ~180 real-time sensor networks and dashboards.
- Managed multiple six-figure contracts, scoped work, contributed to strategic marketing, and trained staff.
- Frequent client communication in diverse groups with competing aims.
- Ad hoc geostatistics, hydrologic modeling, remote sensing.



View this CV online with links at richpauloo.com/cv

CONTACT

✉ richpauloo@gmail.com

🐦 [RichPauloo](https://twitter.com/RichPauloo)

🔗 github.com/richpauloo

🌐 richpauloo.com

📞 (415) 275-4981

FOREIGN LANGUAGE SKILLS

Conversational Spanish

LANGUAGE SKILLS

R	
SQL	
Python	
Bash	

Made with [pagedown](https://pagedown.github.io/) in R.

The source code is available at github.com/richpauloo/cv.

Last updated on 2021-08-26.

present 2020	<ul style="list-style-type: none"> ● Data Scientist + Co-Founder Water Data Lab📍 Remote <ul style="list-style-type: none"> • Currently manage \$105k in contracts. • Build ETL pipeline and design strategic approach. • Co-developed r4wrds.com
2020 2018	<ul style="list-style-type: none"> ● Data Engineer UC Water📍 Davis, CA <ul style="list-style-type: none"> • Built a data processing pipeline and web dashboard¹⁰ for real-time groundwater data via a wireless sensor network. View paper¹¹.
2020 2015	<ul style="list-style-type: none"> ● Graduate Student Researcher Fogg Lab📍 UC Davis <ul style="list-style-type: none"> • Process large hydrologic datasets, 3D numerical groundwater flow and contaminant transport models, & network optimization models. • Developed novel models of well failure, groundwater salinization, and contaminant transport in porous media. • Regularly use R, Python, Git, Bash, MODFLOW, RW3D, Paraview, Illustrator, AWS, Linux, ArcGIS, Envi, LaTeX.
2019 2018	<ul style="list-style-type: none"> ● Data Lab Researcher Computational Institute for Geodynamics (CIG)📍 UC Davis <ul style="list-style-type: none"> • NLP, text mining, and network analysis in R on a corpus of ~600 PDFs. • Developed an R Shiny dashboard¹² to understand the corpus. • Results published here¹³.

SELECTED DATA SCIENCE WRITING

2020	<ul style="list-style-type: none"> ● Automating R scripts on Linux with cron¹⁴
2019	<ul style="list-style-type: none"> ● Using Twilio to Text Myself After Long Running Jobs¹⁵
2019	<ul style="list-style-type: none"> ● Race to the Bottom¹⁶ Exploratory data analysis and science journalism California well construction trends.
2019	<ul style="list-style-type: none"> ● Text Analysis of the Mueller Report¹⁷ Text mining and sentiment analysis.
2018	<ul style="list-style-type: none"> ● Tidy Chi Squared stats in infer¹⁸

View all of my [blog posts here](#).

LINKS

- 1: <https://www.richpauloo.com/talk/2020-exit-seminar/>
- 2: <https://www.richpauloo.com/publication/cig/>
- 3: <https://www.richpauloo.com/publication/well-failure/>
- 4: <https://www.richpauloo.com/publication/vhgr/>
- 5: <https://caopenwater.github.io/sdwisard/>
- 6: <https://r4wrds.com/>
- 7: <http://calwaterquality.com/>
- 8: <https://www.gspdrywells.com/>
- 9: <https://www.youtube.com/watch?v=vBnvVL6XQTW&t=2s>
- 10: <https://www.richpauloo.com/project/lcsn/>
- 11: <https://doi.org/10.3390/w12041066>
- 12: https://richpauloo.shinyapps.io/cig_nlp
- 13: <https://ieeexplore.ieee.org/document/8827910>
- 14: <https://www.richpauloo.com/post/crontabs/>
- 15: <https://www.richpauloo.com/post/textme/>
- 16: <https://www.richpauloo.com/post/race-to-the-bottom/>
- 17: <https://www.richpauloo.com/post/mueller/>
- 18: <https://www.richpauloo.com/post/infer/>