Collin J. Schwantes

https://orcid.org/0000-0002-9882-941X

Summary Technically savvy data librarian with a record of fluidly working on multiple large projects.

Research	Communication	Programming
Biological Systems Modeling	Live Demonstrations	R
Geospatial Analytics	Technical Guides	Python
Graph Analytics	Data Visualization	SQL
Knowledge management	Spanish - *fluent*	Airtable, AWS

Professional Experience **Data Librarian** - *Yale University - Remote - Washington D.C.*
2024-Present

- Support and expand Verena's open data infrastructure.
- Develop software and a culture that facilitate open science.
- Provide analytical support for disease ecology projects.

Data Librarian - *EcoHealth Alliance - Remote - Washington D.C.*
2021-2024

- Build and support the culture, infrastructure, skills, and strategies necessary to have collaborative, reproducible, and open science be the norm at EcoHealth Alliance.
- Lead the transformation of data management by developing strategic plans, emphasizing documentation and flexible planning at the project level, and creating integrated and collaborative systems that unify laboratory, field, and qualitative data streams across multiple international institutions in different resource settings.
- Write software in R (90%), Javascript (8%), and Python (2%) to facilitate open research, including writing, testing, and maintaining software packages in R.
- Develop and improve core databases for multiple one health projects, including investigations involving comparative immunology of wild bats, using relational databases and Airtable.
- Design and implement automated quality assurance and quality control checks for complex data sets.
- Streamline, integrate, administer, and maintain organizational databases that capture and disseminate institutional knowledge.
- Create, curate, and teach a collection of materials on research reproducibility and FAIR/CARE practices.
- Provide on-demand technical support across multiple projects for software testing, data modeling, automated data curation, data integration, data imputation, geospatial analyses, and graph analytics.

Data Scientist - *Accenture Federal Services - Washington D.C.*
2019-2021

- Rapidly develop and deploy data science applications to solve analytics problems for the United States government.

- Data Science Craft lead. Help data scientists within the Discovery Lab thrive by connecting them with resources, building best practices around code and data with an emphasis on documentation, and creating a supportive development community.
- Lead COVID-19 modeling efforts. Develop and deploy near real time forecasting models, phase classification algorithms based on CDC guidance, and a network based COVID-19 transmission simulation. These products provided essential information via cloud based solutions to leadership at three major government agencies and two fortune 500 companies.
- Data visualization expert who uses human centered design techniques to generate insightful and ergonomic tools for exploring data.
- Developed an R-python hybrid application to improve search results and usability of contracting documents with completeness and quality issues. The application uses a semantic search based on the BERT algorithm and a graph database to overcome the inconsistencies in the underlying documents.
- Lead Graph Analytics Community of Interest. Build and nurture expertise in graph analytics and semantic data with colleagues across Accenture Federal Services.

Biosurveillance Scientist - *Accenture Federal Services* - *Washington D.C.*
2016-2021

- Integrated open source health surveillance data from multiple jurisdictions to provide situational awareness for emerging infectious diseases in animals, plants, and people.
- Championed reproducible research best practices, including efforts to improve and standardize documentation for data and code.
- Lead geospatial analysis efforts, train team members to use GIS systems, and provide subject matter expertise on geospatial data standards to ensure interoperability of systems developed by the client. Develop and use geospatial databases like PostGIS to store data.
- Wrote flexible programs in R to parse unstructured or semi-structured text and image based data for use in analyses. This improves reproducibility of analyses, facilitates data sharing, and reduces person hours dedicated to data entry.
- Produced data visualizations and event summaries used in in-depth reports, twice daily communications materials, and congressional/leadership briefings.
- Tested prototype applications in multiple spaces including epidemiological forecasting and natural language processing of medical text.
- Wrote SOPs for complex data integration and maintenance tasks.
- Moderate inter-agency analyst forum and maintain key relationships across DHS to ensure access to mission critical data. **Application Developer** - *Free Lance* - *Washington D.C.*
2015-2018
- Developed web-based interactive graphical user interface for vector control optimization models with clients at the University of Maryland using R Shiny.
- Developed an interactive Hymenoptera taxonomy learning tool that leverages the the Hymenoptera Anatomy Ontology.
- Developed python applications for automating image processing tasks for the Smithsonian National Museum of Natural History Entomology Department. These significantly reduced transcription error rates and the processing time per-specimen.

GIS Specialist and Spatial Analyst - *University of Colorado Museum of Natural History - Boulder, Colorado*
2015

- Developed and began implementation of an open research plan for a multi-year citizen science project using open databases, R, and GitHub to transparently study the relationship between urban land-use and bee diversity in the Front Range of Colorado.
- Created a web application that allows 500+ citizen scientists and interested members of the public to explore bee diversity data.

Research Assistant - *University of Colorado Boulder - Boulder, Colorado*
2013-2015
- Used multivariate statistical analysis to better understand how land-use policy impacts native bee communities. - Collected, integrated, and analyzed biological data, remotely sensed agricultural data, and infrastructure data in R and ArcGIS while managing and mentoring 10 undergraduate research technicians. - Managed data documentation and transmission for multiple studies. - Developed and documented standard operating procedures for data collection, editing, verification and management.

Education Master of Arts Ecology and Evolutionary Biology
University of Colorado Boulder - 2013-2015

Bachelor of Science Evolutionary Biology, Spanish Linguistics, Zoology
University of Wisconsin Madison - 2006-2011

Professional Certifications ICAgile Certified Professional
Accenture Agile Institute - 2016

Publications

Baker, Heather, Asher Grady, Collin Schwantes, Emily Iarocci, Rachel Campbell, Gus Calapristi, Scott Dowson, Michelle Hart, Lauren E. Charles, and Teresa Quitugua. 2018. "NBIC Biofeeds: Deploying a New, Digital Tool for Open Source Biosurveillance Across Federal Agencies." *Online Journal of Public Health Informatics* 10 (1). https://doi.org/10.5210/ojphi.v10i1.8947.

Carper, Adrian L., Virginia L. Scott, Collin J. Schwantes, Stacy B. Endriss, Andrew P. Norton, Deane Bowers, and Mary A. Jamieson. 2014. "Conservation Grasslands Support Diverse Native Bee Assemblages in Agroecosystems." In. Portland, Oregon. https://esa.confex.com/esa/2014/webprogram/Paper86096.html.

Dolan, Levi, Lucy Carr Jones, Barrie Hayes, Hao Ye, Helenmary Sheridan, Marla Hertz, Nicole Contaxis, et al. 2022. "Working Group on NIH DMSP Guidance." https://doi.org/10.17605/OSF.IO/UADXR.

Iarocci, Emily, Collin Schwantes, Anne Folley, Chandra Lesniak, Tiana Garrett-Cherry, and Teresa Quitugua. 2018. "NBIC Collaboration at Multiple Jurisdictional Levels During the Zika Epidemic." *Online Journal of Public Health Informatics* 10 (1). https://doi.org/10.5210/ojphi.v10i1.8585.

Schwantes, Collin J. 2015. "Digitized Data and the Bees of Colorado." In. Minneapolis, Minnesota. https://ecnweb.net/wp-content/uploads/2020/05/ECN_Program2015.pdf.

Schwantes, Collin J., Adrian L. Carper, and M. Deane Bowers. 2018. "Solitary Floral Specialists Do Not Respond to Cryptic Flower-Occupying Predators." *Journal of Insect Behavior* 31 (6): 642–55. https://doi.org/10.1007/s10905-018-9706-9.

Schwantes, Collin J., Benno Lee, Benjamin Ortiz, Marjorie Willner, and Viveca Pavon-Harr. 2021. ""Integrating COVID Models at Different Scales for Infection Risk Estimation and Control Optimization"." In. https://apps.dtic.mil/sti/trecms/pdf/AD1187600.pdf.

Schwantes, Collin, Jeff Demers, Sharon Bewick, and William Fagan. 2023. "Mosquito Control Optimization App." Zenodo. https://doi.org/10.5281/zenodo.8410638.

Schwantes, Collin, Alexandra Rose, Virginia Scott, Adrian Carper, and Deane Bowers. 2023. "Bees Needs Data Exploration App." Zenodo. https://doi.org/10.5281/ZENODO.8408622.

Schwantes, Collin, Noam Ross, Alix Augusto Armero Villanueva, Nathan Layman, Kevin J. Olival, Ernest Guevarra, Evan Eskew, Maëlle Salmon, and Shirley X. Chen. 2023. "EHA Modeling & Analytics Handbook." Zenodo. https://doi.org/10.5281/ZENODO.8408552.

Talamas, Elijah J., Joseph Thompson, Amy Cutler, Samantha Fitzsimmons Schoenberger, Anthony Cuminale, Trenton Jung, Norman F. Johnson, et al. 2017. "An Online Photographic Catalog of Primary Types of Platygastroidea (Hymenoptera) in the National Museum of Natural History, Smithsonian Institution." Plazi.org taxonomic treatments database. https://doi.org/10.15468/RR87G2.