Andrew Zaffos

Developer | Data Scientist | Geologist
aazaff@gmail.com | 201.925.6688

EDUCATION

WILLIAM & MARY
B.A. Economics 2008

UNIVERSITY OF GEORGIA
M.S. Geology 2010

UNIVERSITY OF CINCINNATI
Ph.D. Geology 2014

LINKS

Github:// aazaff
Twitter:// @azstrata
ResearchGate:// Andrew Zaffos
Website:// azstrata.org

PROGRAMMING

R Programming Language:
- Eight Years Experience
- Published velociraptr R package
- Taught R-focused college course

Use Regularly:
R | Bash | Web APIs | PostgreSQL and Postico | git and GitHub | PostGIS | QGIS | LaTeX | Markdown | Adobe Illustrator | Microsoft Office

Use Rarely:
C++ | Python | HTML | ArcGIS | CSS

MANAGEMENT

Academic Experience:
- Taught and designed 3 college courses
- 2 Macrostrat map digitizers
- 1 GeoDeepDive research intern
- 5 field assistants

Non-Academic Experience:
- Student Associate (Asst. Manager)

Earl Gregg Swem Library

COMMUNICATION

Academic Communication:
- Published 11 scientific talks
- Published 1 guidebook chapter
- Published 4 scientific papers

Public Communication
- Published online R Tutorial
- Given 5 invited lectures

POSTDOCTORAL RESEARCHER

GEODEEPDIVE | GEODEEPDIVE.ORG

January 2015 – Present | University of Wisconsin-Madison, WI

- Developer and data scientist for the GeoDeepDive initiative.
- GeoDeepDive is a digital library of machine readable documents taken from partnered open and closed-source publishers (e.g., Elsevier, Wiley, Taylor & Francis).
- Regularly participate in design decisions related to the high-throughput computing work flow, metadata schemas, and API design.
- Designed multiple text-mining applications in R with research intern: intelligent matching of references between different databases, extracting the geolocation and age of geologic formations, and identifying aquifers in the US.
- Strong familiarity with the structure and use of derivative products: natural language processing, optical character recognition, and ElasticSearch.
- Currently co-authoring the next grant proposal for continued NSF funding.

PALEOBIOLOGY DATABASE | PALEOBIODB.ORG

January 2015 – Present | University of Wisconsin-Madison, WI

- Developer, data authorizer, and data scientist for the Paleobiology Database.
- PBDB is a synoptic, global dataset of fossil occurrences and associated attributes: geolocation, geologic age, paleoenvironment, and literature metadata.
- Test and develop its web API and user applications that interact with the API, such as its Navigator tool.
- Create educational tools using the database: an upper-level college course, an online tutorial, and an R package.

MACROSTRAT DATABASE | MACROSTRAT.ORG

January 2015 – Present | University of Wisconsin-Madison, WI

- Developer and data scientist for the Macrostrat Database.
- Macrostrat is a compilation of integrated geologic maps and stratigraphic columns. It allows for analyses of how geologic rock formations are distributed in space-time.
- Test and develop its web API and user applications that interact with the API, such as the Rockd social media application for web, iOS, and Android.
- Oversaw two Macrostrat interns who processed and integrated maps into our infrastructure with QGIS, ArcGIS, GitHub, and PostgreSQL.

GRADUATE WORK

August 2008 - December 2014

- Awarded 12 different merit-based grants or fellowships during graduate work.
- Served as a Teaching Assistant for graduate-level Data Analysis course.
- Published papers using multivariate analysis methods, linear and non-linear logistic regression methods, re-sampling techniques, and traditional univariate statistics.
- Hands on data collection in the field: drilling at sea, modern Caribbean beaches, and fossil digs across the continental United States. Comfortable working in difficult environments.