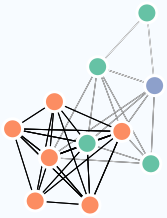
ANDERSON BANIHIRWE

I contribute to and maintain several libraries within the open source scientific Python stack, particularly around improving scalability of Python tools in order to handle petabyte-scale datasets on HPC and cloud platforms.



EDUCATION

2018 | 2014

B.S., Computer Systems Engineering

University of Arkansas at Little Rock — Little Rock, AR

PROFESSIONAL EXPERIENCE

current | 2018-10

Software Engineer

National Center for Atmospheric Research — Boulder, CO

- Project lead & core maintainer of intake-esm[1], a Python package for loading cataloged Earth System Model data on an HPC system or in the cloud.
- Contributing to existing open source software libraries namely Dask[2], xarray[3].
- Designing and building climate data analysis tools leveraging existing scientific computing tools for Python.

2018-09 | 2018-05

Software Developer Intern

Quansight — Austin, TX

- Developed xndframes[4], a Pandas ExtensionDtype/Array backed by xnd[5], a container type that maps most Python values relevant for scientific computing directly to typed memory.
- Worked on integrating cuDF[6] - GPU dataframe library with Apache Arrow[7] library.
- Worked closely with a customer to port existing Postgres code base to Dask[8] based workflow.

2018-04 | 2017-11

Data Science Intern

First Orion — Little Rock, AR

- Built scoring, predictive models with scikit-learn, dask, and apache spark using First Orion's proprietary telecommunication data.

2017-08 | 2017-05

Research Intern

National Center for Atmospheric Research — Boulder, CO

- Developed spark-xarray[9], a Python package that integrates PySpark and xarray for climate data analysis.

View this CV online with links at *cv.andersonbanihirwe.dev*

CONTACT

axbanihirwe@gmail.com

github.com/andersy005

blog.andersonbanihirwe.dev

linkedin.com/in/andersy005

Last updated on 2020-06-28.

SELECTED PUBLICATIONS, POSTERS, AND TALKS

2020

The Pangeo Ecosystem: Interactive Computing Tools for the Geosciences: Benchmarking on HPC[10]

2019 Supercomputing Conference Workshop on Interactive High-Performance Computing

- Authored with Tina Erica Odaka, Guillaume Eynard-Bontemps, Aurelien Ponte, Guillaume Maze, Kevin Paul, Jared Baker, Ryan Abernathey.

2020

Intake / Pangeo Catalog: Making It Easier To Consume Earth's Climate and Weather Data[11]

2020 EarthCube Annual Meeting

- Contributed Jupyter notebook about Pangeo[12]'s data cataloging efforts.

2019

Perceptual Judgments to Detect Computer Generated Forged Faces in Social Media[13]

IAPR Workshop on Multimodal Pattern Recognition of Social Signals in Human-Computer Interaction

- Authored with Suzan Anwar, Mariofanna Milanova, Mardin Anwer.

2019

Interactive Supercomputing with Dask and Jupyter[14]

2019 Scientific Computing with Python conference — Austin, TX

- Contributed talk about Dask and Jupyter.
- Recorded talk[15]
- Slides[16]

2018

Beyond Matplotlib - Tutorial: Building Interactive Climate Data Visualizations with Bokeh and Friends[17]

2018 UCAR Software Engineering Assembly conference — Boulder, CO

- Contributed tutorial about interactive visualization with Python.
- Tutorial materials[18]

2018

PySpark for "Big" Atmospheric Data Analysis[19]

Eighth Symposium on Advances in Modeling and Analysis Using Python — Austin, TX

- Contributed talk about spark-xarray[19].
- Recorded talk[20]
- Slides[21]

LINKS

1: https://github.com/NCAR/intake-esm
2: https://dask.org/
3: https://xarray.pydata.org/en/stable/
4: https://github.com/xnd-project/xndframes
5: https://github.com/xnd-project

6. https://github.com/rapidsai/cudf
7. https://arrow.apache.org/
8. https://dask.org/
9. https://ncar.github.io/PySpark4Climate/
10. https://doi.org/10.1007/978-3-030-44728-1_1
11. https://github.com/earthcube2020/ec20_banihirwe_etal
12. https://pangeo.io/
13. https://www.researchgate.net/profile/Mariofanna_Milanova/publication/333414231_Perceptual_Judgments_to_Detect_Computer_Ge
Judgments-to-Detect-Computer-Generated-Forged-Faces-in-Social-Media.pdf
14. https://youtu.be/vhawO8fgD64
15. https://youtu.be/vhawO8fgD64
16. https://andersonbanihirwe.dev/talks/dask-jupyter-scipy-2019.html
17. https://sea.ucar.edu/event/beyond-matplotlib-building-interactive-climate-data-visualizations-bokeh-and-friends
18. https://github.com/andersy005/beyond-matplotlib-tutorial-sea-2018
19. https://ncar.github.io/PySpark4Climate/sparkxarray/overview/
20. https://ams.confex.com/ams/98Annual/webprogram/Paper334546.html
21. https://opensky.ucar.edu/islandora/object/conference%3A3443