

[To view this email as a web page, go here.](#)

THE UNIVERSITY OF ARIZONA



RESEARCH, INNOVATION & IMPACT

Data Science Institute

## The Data Drip

A digest of Data Science news at UArizona

January 2024

*New Year New Opportunities!*

### DATA SCIENCE INSTITUTE (DSI)

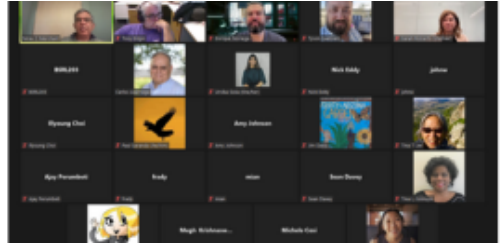
#### UArizona Data Lab Spring 2024 workshops starting NOW!

The [UArizona Data Lab](#) with DSI and ICDI addresses key topics in data science and focuses on skills required when working with modern computing languages, data science tools, and AI.



- [NextGen Geospatial Data Science](#) - Tuesdays 2-3pm
- [Data Science Essentials: From Jupyter to AI Tools](#) - Tuesdays 3-4pm
- [Data Science Tapas: Savor the Tools of Data Mastery](#) - Wednesdays 1-2pm
- [Intro to Classical Machine Learning. Mastering Machine Learning: Your Path to Data-Driven Research](#) - Thursdays 2-3pm
- [Intro to Deep Learning. Deep Dive into Data: A Guide to Deep Learning](#) - Thursdays 3-4pm

Connect with [Carlos Lizarraga](#) for questions and more info about the UArizona Data Lab. Check often for more Spring semester workshops and activities on the [DSI calendar](#) of events.



## Project Chatur: An LLM codeathon

Pronounce it however you want, but Project Chatur – Phase I was an intense but gratifying 3-day codeathon that allowed participants to explore tool, techniques and best practices for building applications using [open Large Language Models \(LLM\)](#) with private data. Team members from [DSI](#), [ICDI](#) and [CS](#) built a proof-of-concept platform for inference, tuning, and augmenting open LLM's with the goal of creating shared infrastructure that can support multiple projects and courses. They created a Retrieval Augmented Generation (RAG) pipeline and an extensible chatbot using the Chatur platform using NSF Jetstream2 GPU's and CyVerse Data Store and cloud native technologies. Phase II will include organizing a larger community codeathon and a white paper describing the capabilities, recommendations and lessons learned. Interested in learning more about Project Chatur or using LLM's for your research and courses, contact DSI Director, [Nirav Merchant](#).

**CyVerse FOSS starting soon. Free to UArizona. Register now!**

[Register](#) to join others who want to discover how FOSS *actually is* for them and their research. Foundational Open Science Skills (FOSS) virtual workshops held **January 25 - March 14, 2024** (Thursdays) at 11 AM – 1 PM PDT/MST/AZ. Use open source tools to manage the full data lifecycle for open science and improved reproducibility; Organize your lab, communications, and research with productivity software; Connect with colleagues across institutional boundaries and research domains; and Engage with the community of Open Science practitioners.





## AEOLUS Soars with Eighteenth Mile TRIF Funding

In a significant lift towards advancing research, the Airborne Environmental Observatory and Laboratory for Uncrewed Systems (AEOLUS) project, led by PI [Dr. Tyson L. Swetnam](#) and Co-PI [Dr. Jeffrey K. Gillan](#), has been awarded Eighteenth Mile Funding Program funds for FY 2024. The \$141,000 grant, part of the UArizona Office for Research, Innovation, and Impact and the Innovative Technologies for the Fourth Industrial Revolution TRIF Fund, will help propel the AEOLUS project to new heights. AEOLUS's goal is to leverage cyberinfrastructure for improved data management for data-intensive research using drones and uncrewed aircraft. [Read more](#) about AEOLUS.



## UArizona Data Lab and Tree Ring Lab Bootcamp

On January 8 and 9, 2024, the UArizona Tree Ring Lab hosted the openDendro Bootcamp. Dendrochronology is the scientific technique using tree rings in order to date past events and environmental changes by using the characteristic patterns of annual growth rings from tree trunks samples. The UArizona Data Lab has been directly involved in modernizing the field and contributing to open Dendrochronology (openDendro) by developing code and hosting the open source openDendro software through CyVerse. For decades, dendrochronology heavily relied on software developed with FORTRAN, a powerful but dated computer language. OpenDendro intends to develop software based on R and Python – contemporary coding languages – which would allow newer dendrochronology cohorts the ability to contribute to the scientific arena. Making the codebase available in R and Python gives researches the possibility to add new functionalities, suggest modifications, and overall expand the dendrochronology field. Learn more about [openDendro](#) and its efforts.

## DSI CALENDAR

Be sure to check regularly the [Data Science Institute's calendar](#) for information about useful data science workshops, trainings, and events offered by the Data Science Institute, UArizona Data Lab, and other departments across campus.

## NEWS AND ACTIVITIES

### CyVerse Webinars Line-up for Spring 2024

Save these Friday 10:00 AM slots for more free webinars to help you do your data-driven research and teaching. Visit [CyVerse webinars](#) for more details.



- February 16: Get Started with CyVerse: How to make the most of your free UArizona account
- February 23: A Conversation with your own data with open LLM's using NSF Jetstream and CyVerse
- March 22: Managing and Analyzing Wearables Data in Your Research
- April 19: The Art and Science of User Experience (UX) in Data Science
- May 3: Geospatial Data Analyses Using CyVerse

### UArizona RSE Group includes research software professionals

The UArizona Research Software Engineers (RSE) Group that [DSI](#), [ICDI](#), and other UArizona departments are forming on campus includes research software professionals such as Software Engineers, User Experience/User Design Specialists, Application Developers, DevOps Engineers, Research Engineers, Research Scientists, and more. The UArizona RSE will help build a campus community to support the UArizona research mission. Join to learn more or be part of the



**UArizona  
Research  
Software  
Engineers  
Group**

UArizona RSE Group by emailing [RII-datascienceinstitute@arizona.edu](mailto:RII-datascienceinstitute@arizona.edu).

### For the Love of UX, Code, and Data

Join [ResBaz AZ](#) for a special co-hosted, in-person Hacky Hour for all folks interested in user experience, technology, coding, programming, software, and more on February 15th, 4:30 PM - 6 PM. Mark your calendars and check out the [event details](#).



### WiDS Tucson Conference planning underway

The Women in Data Science Tucson conference is in its initial planning phase. Contact [Kristina Riemer](#) if you are interested in planning or want more information about the 2024 [WiDS Tucson](#) conference.



### Conduct research in a secure data analysis enclave with Soteria

[Soteria](#) changes the way researchers analyze and manage projects with protected data like PII and PHI. Soteria is a secure data and analysis enclave where teams can collaboratively organize and analyze data while adhering to compliance requirements associated with protected data. Resources and state-of-the-art tools are available through Soteria. Explore a variety of tools that can help propel your research.

[Apply now for Soteria access](#)

### Free UArizona CyVerse accounts





Get your free UArizona CyVerse account today! Thanks to the Technology and Research Initiative Fund (TRIF), 2,500 Academic Pro accounts that offer free data storage, and compute and cloudnative resources are available to UArizona students, faculty, and staff.

[Get your UArizona CyVerse account](#)



### Spotlight you and your research!

*The Data Drip* will highlight UArizona students, faculty, staff, research, data science projects, tools, and/or software. DSI wants to shine the light on you. Let us know what you want to share.

[Email us to contact you](#)

## MUSINGS AND OPINIONS

The views and opinions expressed in *The Data Drip* do not necessarily reflect the official policy or position of the UArizona Data Science Institute or the University of Arizona.

### DSI Spotlights...

#### Coffee and Code

Still strong in 2024, a unique community-driven service in Tucson, AZ, **Coffee and Code**, has been providing support since 2016. Currently hosted by several partners including Dr. Valeria Pfeifer and Julian Pistorius, this weekly gathering serves as a beacon for non-researchers and researchers seeking not only coding assistance but also a supportive community to navigate the challenges of research projects. Coffee and Code is more than just a help desk; it's a community where individuals from diverse backgrounds and cross disciplines come together to troubleshoot code, seek advice on various challenges, and foster connections. The welcoming atmosphere encourages both newcomers and repeat visitors to participate...

[LEARN MORE](#)

Share [The Data Drip](#) with your friends and colleagues! They can [subscribe](#) by completing a brief form. Let us know if you have data science news that you want us to share in The Data Drip or on the [Data Science Institute](#) website.



The UArizona data science community is reading and sharing...

- [Python packaging 101](#)
- [Hearst Health Prize](#)
- [Inside the World Excel Championships \(Yes, You Read That Right\)](#)
- [COGS 500: Computing for the Research Lab](#)
- [For the Love of UX, Code, and Data, Thu, Feb 15, 2024, 4:30 PM | Meetup](#)
- [DateTime](#)
- [Noam Chomsky turns 95: the social justice advocate paved the way for AI. Does it keep him up at night?](#)
- [Generative AI Has a Visual Plagiarism Problem](#)
- ['Merchants of Complexity': Why 37Signals Abandoned the Cloud](#)
- [The oldest-known version of MS-DOS's predecessor has been discovered and uploaded](#)
- [The tech sector is pouring billions of dollars into AI. But it keeps laying off humans | CNN Business](#)
- [DeepMind AI solves geometry problems at star-student level](#)
- [Neuroscience for machine learners](#)

Read more news about the UArizona data science community on the [Data Science Institute news](#) site.

## CALENDAR

Data science workshops are available across campus. Take a look at the DSI [calendar](#) of events for more workshops. To see your data science events listed below and on the DSI Events page, complete the DSI [calendar](#) request form.

JAN-APR  
16-23

[NextGen Geospatial Data Science](#)

Step out of conventional GIS frameworks and discover the latest trends in geospatial data science where open tools, cloud technologies, and the proliferation of sensor data are innovating earth observation and environmental monitoring.

JAN-APR  
16-23

**Data Science Essentials: From Jupyter to AI Tools**

Sessions cover topics such as data wrangling, statistics, visualization, exploratory data analysis, time series analysis, machine learning, natural language processing, deep learning, prompt engineering, and AI tools

JAN-APR  
17-24

**Data Science Tapas: Savor the Tools of Data Mastery**

“Data Science Tapas” offers essential data science tools for Ph.D. dissertation research. Participants will learn command line, version control, project management, data handling, documentation, and workflow automation.

JAN-APR  
18-25

**Mastering Machine Learning: Your Path to Data-Driven Research**

“Mastering Machine Learning: Your Path to Data-Driven Research” covers essential topics in classical machine learning for data analysis and visualization in Ph.D. research.

JAN-APR  
18-25

**Deep Dive into Data: A Guide to Deep Learning**

The “Deep Dive into Data: A Guide to Deep Learning” workshop is a comprehensive resource for Ph.D. students in data modeling and visualization.

JAN-APR  
19-26

**CCT Bayesian Learning Group**

8 biweekly sessions in a peer-to-peer learning collaborative while working through foundational concepts in Bayesian statistics.

JAN  
24

**CCT Demystifying APIs for Researchers**

Learn what an API is, how to access data using APIs, and how to write R code to automate downloading data from APIs using the httr2 package.

JAN-MAR  
25-14

**FOSS Spring 2024**

The workshop is aimed at graduate students and researchers at all



stages of their education and careers who wish to learn computational tools to increase the reproducibility and efficiency of their work.

JAN  
31

### **End-to-end data science workflows with Posit Team**

**Posit** will walk through an end-to-end data science workflow and demo the core functionality of Posit Team while highlighting some of the latest features!

## **Add me to the DSI email list.**

Did a friend send this email to you? Subscribe to let us know to add you to the Data Science Institute email list. Or email the Data Science Institute so we can be of assistance to you and your research.

SUBSCRIBE TO THE DATA DRIP

EMAIL THE DATA SCIENCE INSTITUTE

You are receiving this email from the **University of Arizona's** mailing list.

University of Arizona, 1401 E. University Blvd., Tucson, AZ 85721, US  
© Copyright 2024 The University of Arizona. All rights reserved.

### **Land Acknowledgment**

*We respectfully acknowledge the University of Arizona is on the land and territories of Indigenous peoples. Today, Arizona is home to 22 federally recognized tribes, with Tucson being home to the O'odham and the Yaqui. Committed to diversity and inclusion, the University strives to build sustainable relationships with sovereign Native Nations and Indigenous communities through education offerings, partnerships, and community service.*