

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

THE UNIVERSITY OF ARIZONA

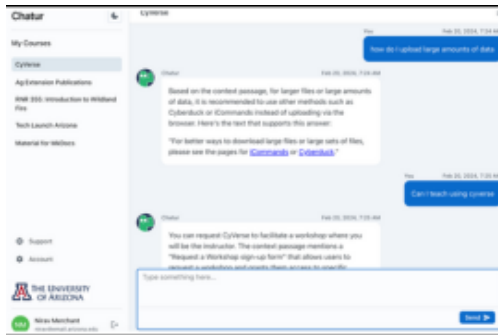


RESEARCH, INNOVATION & IMPACT
Data Science Institute

The Data Drip



DATA SCIENCE INSTITUTE (DSI)



UArizona, chat and ask questions with Chatur, the chatbot!

In January, team members from UArizona [DSI](#), [ICDI](#), [CyVerse](#) and [Computer Science](#) built a proof-of-concept application using open Large Language Models (LLM) with private data as part of a three-day codefest. The goal of the codefest was to explore how LLM capabilities can be made available to UArizona researchers and educators with data that has privacy restrictions. On February 15, 2024, the experimental version of the chatbot, [Chatur](#), created as part of the codefest was launched for testing and evaluation.

This completely in-house chatbot has many capabilities that are on par with any commercially available solution, with all data ingested by Chatur remaining at UArizona. Having the ability to reply with human-like answers, Chatur is currently trained and equipped to answer user questions only from content from five data sources and use cases. A student or researcher can chat or ask questions related to: 1) course content for RNR 355 Introduction to Wildland Fire from instructor provided PPT and PDF documents; 2) selected publications from UArizona Cooperative Extension PDF's from <https://extension.arizona.edu/pubs>; 3) CyVerse learning and training material from <https://learning.cyverse.org>; 4) the Tech Launch Arizona website; and 5) MKDocs, a popular utility for creating documentation.

Chatur runs on state-of-the-art hardware with advanced graphical processing units and is currently hosted on CyVerse and at NSF JetStream2 infrastructure. [Email](#) or contact DSI Director, [Nirav Merchant](#), if you want to participate in the next codefest or are interested in learning more about Chatur or using LLM's for your research and courses.

CHAT WITH CHATUR

NEW - webinars and workshops - now on YouTube

CyVerse Webinars Level up your computational game!

February 23 - [A Conversation with Your Own Data with LLMs Using NSF JetStream2 and CyVerse](#) Presented by Nirav Merchant, Mithun Paul, and Tyson Swetnam.

Introducing **NEW** UArizona Data Lab **Workshops**

- **NEW** - [Navigating the World of Data Engineering](#) - Mondays 2-3pm
- [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
- **NEW** - [Cracking the Coding Interview](#) - Wednesdays 2-3pm
- [Mastering Machine Learning: Your Path to Data-Driven Research](#) - Thursdays 2-3pm
- [Introduction to Deep Learning](#) - Thursdays 3-4pm
- **NEW** - [Cracking the Coding Interview-Virtual Only](#), Fridays 4-5PM



UArizona celebrated Love Data Week/Day 2024

Feb. 12-16 was Love Data Week. And who doesn't love data? UArizona celebrated with a special "Love Data Day" – February 14th – the day of love! [ICDI](#) and [DSI](#) set up in front of Main Library with banners, stickers, and "data" games. Can't be sure, but the baskets of candies might have attracted several students who learned how data impacts their every day lives. Many of the students were most interested in learning more about the UArizona Data

Lab's recently added new workshops and CyVerse webinars. They also learned how the Data Lab helps to improve research, provide career opportunities, facilitate industry partnerships, and encourage innovation and entrepreneurship – gotta love it. [Learn more.](#)

LOVE DATA DAY



-43 degrees for uncrewed systems research

Drs. [Jeffrey Gillan](#) (Data Science Institute) and [Tyson Swetnam](#) (Institute for Computation and Data-enabled Insight) travelled to the University of Alaska Fairbanks in January to meet with members of the [Alaska Center for UAS Integration \(ACUASI\)](#) to discuss research collaborations and partnerships with [CyVerse.org](#) and the University of Arizona. ACUASI is a global leader in uncrewed systems research, leveraging the open skies of Alaska to test new UAS technology for academic and civil applications. Swetnam hopes to build on collaborative research with ACUASI and other federal land management agencies around UAS data management and platform integration to CyVerse and expand the use of UAS and small UAS (sUAS) data in Arizona for research. A polar vortex welcomed Drs. Gillan and Swetnam to Fairbanks with -40 degree weather, last seen over two winters ago. Swetnam and Gillan will return to Alaska to present at ACUASI's [Global Autonomous Systems Conference](#) in August as invited speakers.



Data Science Fellows meet up

On February 6, 2024, the Data Science Fellows (DSF) program held the first of four meetups that will be held during the Spring 2024 semester. In addition to a robust agenda, the meetup included





introductions of the current Fellows cohort, Data Science Institute staff, and several UArizona data science community members. The Distinguished Guest was Nirav Merchant, Director of the Data Science Institute, who presented the attendees with an online interactive discussion about "the future of data science." Learn more about the [DSF program.](#)



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.](#)

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

UArizona Getting Started with CyVerse webinars

"Getting Started with CyVerse and Support Hour" starting in



March for UArizona CyVerse accounts holders. Check the [DSI Calendar](#) for more info. Create your free [UArizona CyVerse account](#). Also, save these Friday morning slots for free [CyVerse webinars](#).



- 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

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 cloud native resources are available to UArizona students, faculty, and staff.

CyVerse: Cyberinfrastructure for open science

CyVerse, the largest publicly-funded open-source research cyberinfrastructure for life sciences, has played a crucial role in advancing data-driven research since the 2010s. As the technology landscape evolved with the emergence of cloud computing platforms... [Read more](#) about this paper.

Read the ICDI newsletter, AI Update

Have you read the latest AI Update? [Subscribe](#) to receive the newsletter, AI Update, and stay connected with the latest breakthroughs in [Artificial Intelligence](#) directly from the University of Arizona.



CCT Data Science Team Workshop Wednesdays: Customizable Resumes with R Tools

February 28, 11 AM - 1 PM, learn how to create and customize your CV / resume using R Markdown templates. From this workshop, you will become familiar with the options in R for making resumes and come away with one with your own information. While previous experience with R is useful, it is not a prerequisite to participate in this workshop. [Register](#).



WiDS Tucson Conference - Register Now

Registration for the [2024 Tucson Women in Data Science conference](#) is now open! The WiDS Tucson conference will be held Friday, March 29, 2024, 8:30 AM – 5:30 PM MST. The conference is a great opportunity for women to celebrate, connect, share, and learn about data science related research approaches and tools in various domains. Everyone is invited to attend the free annual occurrence!



**Women in
Data Science
Worldwide**

Tucson

WiDS Tucson is an independent event organized by the University of Arizona in association with the WiDS Worldwide nonprofit organization, which hosts a central conference and supports an estimated 200+ locations worldwide annually. [Register](#) soon for the WiDS Tucson conference which is available virtually, in addition to the morning keynote speaker and first concurrent speakers who will be in-person on the University of Arizona campus.

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. Learn more or be part of the UArizona RSE



Group by emailing RII-datascienceinstitute@arizona.edu.



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

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...

- [AI breakthrough enables scientists to read Roman scrolls once buried by Mount Vesuvius](#)
- [NIST Establishes AI Safety Consortium](#)
[More than 10,000 research papers were retracted in 2023 — a new record](#)
- [FOSDEM 2024 - AI and Machine Learning devroom](#)
- [AMD Quietly Funded A Drop-In CUDA Implementation Built On ROCm: It's Now Open-Source](#)
- [ResearchHub | Open Science Community](#)
- [Microsoft BitLocker encryption cracked in just 43 seconds with a \\$4 Raspberry Pi Pico](#)
- [Dear Colleague Letter: Request for Information on Researcher and Educator Use Cases for the National Artificial Intelligence Research Resource \(NAIRR\)](#)

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

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.

FEB-APR
Mondays

[Navigating the World of Data Engineering](#)

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.

FEB-APR
Tuesdays

[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.

FEB-APR
Tuesdays

[FOSS Spring 2024](#)

An 8-week long workshop of 2 hour weekly sessions on a variety of foundational principles, skills, and technologies that are essential and valuable for doing Open Science.

FEB-APR
Tuesdays

[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

FEB-APR
Wednesdays

[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.

FEB-APR
Wednesdays

Cracking the Coding Interview

NEW

“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.

FEB-APR
Thursdays

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.

FEB-APR
Thursdays

Intro to Deep Learning

In this series you will learn everything about deep learning. We will start with how deep learning works fundamentally and by the end of the workshop series, you will be building your own ChatGPT.

FEB-APR
Fridays

CCT Bayesian Learning Group

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

FEB
23

A Conversation with Your Own Data with LLMs Using NSF JetStream2 and CyVerse

NEW

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.

FEB-APR
Fridays

Cracking the Coding Interview-Virtual Only

A virtual only workshop to get techniques and tips that can help better prepare you for your next computer science job interview.

NEW

FEB-APR
FRIDAYS

Deep Learning for Astrophysics

This is a comprehensive introduction to deep learning with practical applications in astrophysics.

FEB
28

Customizable Resumes with R Tools

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.

MAR
29

Tucson Women in Data Science Conference 2024

The Tucson one-day technical conference provides an opportunity to hear about the latest data science related research and applications in a number of domains, and connect with others in the field.

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 **Data Science Institute** 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.