

Department of Astronomy and Planetary Science
AST 520 Astroinformatics: Big Data in Astronomy

2021 Spring

3 Credit Units

Prerequisite: Admission to Astronomy and Planetary Science PhD

Updated January 21, 2021

Mode of Instruction: Lecture

- Meets Tuesdays and Thursdays 11:10am - 12:25pm online or in person at HLC 3104. Current NAU rules require remote learning only until January 25.
 - To connect to the virtual classroom: Please use Zoom to connect to the virtual classroom or office hours. Check that you have the latest version by choosing “Check for Updates” under the “zoom.us” menu (on a Mac, probably something similar on a PC and check your app store if on a mobile device). The Zoom link is: <https://nau.zoom.us/j/6032695475> (which requires the passcode, which is 22) or <https://nau.zoom.us/j/6032695475?pwd=VUY3MEFaUmZaNFFyNnJNTVRsaTNqZz09> (which doesn't require the passcode but is harder to type in). Or one tap mobile +16699006833,,6032695475# US (San Jose)
 - About the Virtual Classroom: Please sign into Zoom via your NAU email first otherwise I can't give you attendance or participation points (because without the official NAU login I can't know for sure it's you).
 - About the Virtual Office Hours: I may be available outside my normal office hours, which are Monday and Wednesday 11:30am - 12:30pm. But I can't guarantee it unless you schedule in advance by email or text message to my cell (808-756-4393). But you are welcome to give it a try. To do so, please connect as if you were going to virtual class. I'll get an alert on my phone that someone's in my Zoom room and I'll be on in a few minutes if I'm available.
 - Instructor: Dr. Chad Trujillo Email: chad.trujillo@nau.edu
Office: Physical Sciences (#19) 312 Office phone: 523-6007
Office Hours: Mon/Wed 11:30 – 12:30pm and by appointment Cell phone: 808-756-4393
Availability: Phone response will be within an hour if a message is left or a text is sent, weekday emails will generally be responded to within a day. You can also send class questions or feedback anonymously to Dr. Trujillo at <https://goo.gl/forms/BSvNSPSML7WXMzm22> .
-
-

Course Content: This course provides training in the fundamentals of astroinformatics: applying “big data” techniques to research topics in astronomy. Course material will include case studies of astroinformatics projects that exist presently and that are coming in the future; tutorials in computational approaches; exposure to relevant statistical approaches; and training in creating informatics research topics.

The course will conclude with a term project in which students will apply the skills they have learned to existing data sets.

Goals: The main goal of this course is to prepare you for real-world research in astronomy which is becoming increasingly computationally intensive. Even if you do not do astronomy as a profession in the future, the skills you will learn in this course will help you in almost all STEM-related fields. I believe that the best way to learn is to learn by doing. Therefore, to reach the main goal, we will explore several smaller goals in this course:

In the first 8 weeks of the course, we will review journal publications that are good examples of what astroinformatics is about. We will look at the publications in detail and will identify and summarize the main methodologies and techniques used. The summaries will be done in writing and/or in oral presentations in class. There will be associated technical exercises (primarily programming) related to the readings with the goal of demonstrating skills associated with these publications.

In the next 3 weeks of this course, we will brainstorm research ideas in astroinformatics and create mock proposals for these ideas. These proposals will be peer-evaluated by us in a blind manner, so nobody knows whose proposal is whose. We will rank the proposals and identify the best proposal for the last third of the course.

In the final 4 weeks of the course, we will take the highest ranked research proposal and implement it. The work effort will be distributed among all the people in the class. The goal is that we will, as a team, first implement the proposal on a small amount of data on a desktop/laptop computer then execute the full analysis using the NAU Monsoon High Performance Computing cluster. We will then write up the results of the research project into a single final paper. The last week there will be short presentations by each person on their aspect of the project.

Class Outline: Below is an approximate outline of the topics covered by week. Keep in mind there is the possibility of class cancellation in the event of instructor jury duty, adverse weather such as snow, terrorist threat and/or pandemic. If I deem travel conditions unsafe, I will notify the class by email through NAU's classlists prior to 9am the morning of class. Generally this won't happen until there has been 6 inches of snowfall the previous night or several inches with worsening conditions around class time. Anytime NAU is closed, this class will also be canceled.

- Week 1: (Jan 12 – 14): Introduction, Syllabus, Astroinformatics Overview, RAID arrays, Catalogs: WISE, 2MASS, Vizier. Monsoon Training.
- Week 2: (Jan 19 – 21): Catalog Queries, Cursor Search, Binary Trees, Python and NumPy, Catalogs: Sloan DSS, Kepler, NEOWISE. Homework 1 Due.
- Week 3: (Jan 26 – 28): Algorithm Optimization, Binary Trees, KD Trees, Fast Fourier Transform, Matched Filters, Convolution/Correlation, Experiment: LIGO, Catalog: Gaia. Homework 2 Due.
- Week 4: (Feb 2 – 4): Fast Fourier Transform, Image Processing, Source Extraction, Catalog Creation, Kernels and Convolution, Source Extractor. Homework 3 Due.
- Week 5: (Feb 9 – 11): NumPy, Evil For Loops, Source Extractor, SCAMP. Homework 4 Due.

- Week 6: (Feb 16 – 18): Citizen Science, Zooniverse Projects, Gaussian Statistics, Bayesian Statistics. Homework 5 Due.
 - Week 7: (Feb 23 – 25): Zooniverse Projects, Top Hat, SQL and Relational Databases, Alert Streams and ZTF, Project Proposal Work. Homework 6 Due.
 - Week 8: (Mar 2 – 4): SQL and Relational Databases, Machine Learning. Homework 7 Due. Project Proposal Work.
 - Week 9: (Mar 9 – 11): Draft Project Proposals Due.
 - Week 10: (Mar 16 – 18): Final Project Proposals Due.
 - Week 11: (Mar 23 – 25): Project Proposal Presentations, Rankings, Choice, Delegate Work
 - Week 12: (Mar 30 – Apr 1): Group Project Work
 - Week 13: (Apr 6 – 8): Group Project Work
 - Week 14: (Apr 13 – 15): Group Project Work
 - Week 15: (Apr 20 – 22): Last Week of Instruction: Final Project Presentations, Final Project Due
 - Week 16: (Apr 26 – Apr 29): Finals Week (no exam, no class)
-

Class Times: This course meets Tuesday and Thursday from 11:10am – 12:25pm either virtually (see Zoom connection info on the first page) or in HLC 3104. We all have busy schedules, so please respect the beginning and ending times. If we do meet in person, although this is over lunchtime, I request that you don't eat in class because that means that you won't be wearing a mask.

Office Hours: Looking at everyone's course schedules, it looks like most people are available if I hold office hours Monday and Wednesday from 11:30 am - 12:30 pm. Other hours are available by appointment. You are invited to drop-in virtually to my Zoom room, but please understand that I may not necessarily be available. I check email frequently weekdays, so that is probably the best way to contact me. For more urgent issues, use the phone or text my cell and be sure to leave a message if I cannot answer.

Prerequisites: This course has a prerequisite of being enrolled in the Astronomy and Planetary Sciences PhD program. I am open to considering enrollment for seniors or those with other degrees depending on their background.

Required Materials:

- Computer: You will need a computer to do assignments and experiments which will be computationally intensive.

- Monsoon Training: The latter part of this course will involve using NAU's Monsoon High Performance Computing cluster. You must register for a Monsoon account and take a self-paced training to use Monsoon. Info about this can be found on the HPC website, currently <https://in.nau.edu/hpc/obtaining-an-account>
 - Paper and pencil.
 - Textbook: There are no required textbooks for this course because there are none suitable for this course. There will be reading assigned on a case-by-case basis drawn from journal articles and other sources.
-

Attendance: You are expected to attend every class. This is a small, discussion-style class so having even one person missing can result in a fundamental change in the feel of the class. You may miss 3 classes over the semester without penalty (no excuse needed) if you notify me ahead of time. Institutional excuses do not count towards the 3 classes you can miss. Class sessions will be recorded and made available the day of class via NAU's Google Drive File Stream. As a reminder, all class recordings are to for NAU use only. If you cannot attend class due to schedule conflict or internet connection problems, please view the video within 24 hours after it's been posted.

Homework Policy:

- My goal is to give out homework on Thursdays by 11:59pm and collect homework the Thursday a week later by 11:59pm. The point of this is so that you don't have to do homework on the weekend and I don't have to create it or grade it over the weekend either. Homework will be assigned and turned in through NAU's Google Drive File Stream because of the large file sizes that will often be encountered. Source Code written as part of homework should be turned in as a file and a PDF version of the code should also be turned in, preferably with color syntax highlighting enabled. This is so I can write comments on the PDF to provide feedback.
- For all homework, as in life, you may work together. However, the submitted homework you must create yourself. You must also reference other people that helped you with the homework. The reason for this is because in the real world, not attributing your sources does them a disservice at best and at worst is plagiarism, which is one of the greatest sins in academia.
- Reading assignments with feedback: we will have reading assignments from journal publications and other sources. You will generally have to create some kind of feedback, which may be a written summary of the article or you may have to lead a discussion in class. The purpose of this is because studies show that if you write/discuss an article (rather than just reading it), knowledge of the subject and retention is greatly increased.
- Technical assignments: there will be assignments where computational exercises will include creating code for analysis. In these cases, you may discuss computational techniques with people, but you should write up your code yourself using your own style.
- Calculations: there will be exercises where a calculation is required. In these problems, you must show your work in enough depth that someone else can follow your reasoning. This is because (1) if you make a mistake in computing the final answer, you will still get partial credit for the approach

you use and (2) when doing example calculations in real life (such as in a published paper or thesis), you will have to show your work so that others can follow.

- Writing assignments: being able to communicate effectively in writing is one of the most important skills in the sciences. There will be writing assignments where not only the content but the clarity of your work will be assessed.
 - Late policy: you can turn in 2 assignments in this course 4 days late (nominally Monday at 11:59pm for homeworks normally due Thursday) for full credit if you notify me *prior* to the deadline. No other excuse is needed. This is because stuff happens in life and sometimes you cannot make the deadline. You cannot “bank” this time (turning in 1 assignment 2 weeks late) or “gift” your late assignment quota to others. The four days is because that deadline still lets me talk about the homework on Tuesday after the assignment is normally due. That is because this could result in people falling too far behind in the class progression.
-

Exam Policy:

- In this class you will generally be allowed to use anything at your disposal to take an exam including notes, books or the web. Do *not* assume that this means you shouldn't study the course material.
 - Exams are designed to test individual knowledge. Therefore no group work is allowed on tests.
 - As in the homework, all work must be shown for full credit.
-

Grading: Letter grades will be assigned as follows: A at 90%, B at 80%, C at 70%, D at 60% and F below 60%. The requirements for grades may be relaxed based on class performance.

- Final Exam: This is a project-based research-oriented class, as such there will be no final exam.
- Homework: There will be about 7 homeworks assigned for the first half of the class. These will primarily be reading / response and will count towards 35% of your grade.
- In-Class Presentation: You will also be asked to give a short (5 minute) in-class presentation about a reading assignment. This will count for 5% of your grade (similar to a single homework assignment).
- Final Project: There will be a writing project for the second third of the class where mock research proposals will be written. These are individual works and will count for 20% of your grade. The final third of the class will be a group exercise where we will conduct an actual astroinformatics project. This will also count for 20% of your grade.
- Attendance: Attendance and in-class participation will each count towards 10% of the grade for a total of 20% . For those that cannot attend class contemporaneously participation will be waved (attendance will be assessed once the class video is viewed).
- Extra Credit: There is extra credit for attending the Astronomy and Planetary Science Colloquium. Attend and write two paragraphs with the speaker's name and describe (1) the instrumental setup the speaker used, (2) one of the main conclusions, and (3) what the speaker might be doing in the future along this line of research. One colloquium is equivalent of 12.5% of a single homework, so if you attend 8 colloquia, it is the equivalent of 1 homework assignment.

Plagiarism and Cheating: Plagiarism is using someone else's work or ideas and passing them off as your own. It is considered the most serious breach of scientific integrity. Evidence of this is work suspiciously similar to other's work (exact same wording, or very similar wording) with no attribution. This is considered cheating in this course. All people involved in cheating and/or plagiarism will be given a zero on the assignment / exam, regardless of who cheated off whom. Repeat offenses will be escalated to the Dean following the NAU Academic Integrity Policy.

E-learning Resources Policy: Some course work will take place electronically, including (but not limited to) out-of-class communication, presentation of reading materials, distribution of homework and virtual lectures. It is the responsibility of the student to check email and BbLearn regularly. The student must also allow ample time prior to deadlines to navigate any technological issues that may arise such as computer crashes, internet outages, software version mismatches, etc. Students are also strongly encouraged to back up all data.

Academic Contact Hour Policy: The Arizona Board of Regents Academic Contact Hour Policy (ABOR Handbook, 2-224) states that for a 3 credit course such as this one, a student should expect to spend a **minimum** of 9 hours per week on average. Class time is 2.5 hours per week. Therefore, at least 6.5 hours per week should be spent outside of class on homework, studying and preparation.

Disability Resources: If disability accommodations would benefit you, please talk to me and/or the Disability Resource office at NAU. Many services and accommodations are available at no cost to NAU students.

NAU Policy Statements: The following pages contain the NAU Policy Statements for Course Syllabi. They take precedence over anything earlier in this document.

COVID-19 REQUIREMENTS AND INFORMATION

The following statements in red set forth in this document's first section are specific to NAU's response to the COVID-19 situation. The requirements outlined below are mandatory until further notice. They are based upon current public health conditions and guidance and may change as circumstances warrant or new information becomes available. Additional information about the University's response to COVID-19 is available from the **Jacks are Back!** web page located at <https://nau.edu/jacks-are-back/lumberjack-responsibilities>.

FACE COVERING AND PHYSICAL DISTANCING REQUIREMENTS

Appropriate face masks or other suitable face coverings must be worn by all individuals when present in classrooms, laboratories, studios, and other dedicated educational spaces. To maximize the benefits of physical distancing as an important strategy to help reduce community transmission of the SARS-CoV-2 virus, instructors may implement mandatory student seating arrangements or specific seat assignments. Instructors may remove students who do not cooperate with these requirements from the instructional space in the absence of an approved accommodation arranged through Disability Resources. Failing to comply with these requirements may constitute a violation of the university's *Disruptive Behavior in an Instructional Setting* policy available at <https://nau.edu/university-policy-library/disruptive-behavior>.

USE NAUFLEX TO HELP MAINTAIN PHYSICAL DISTANCING

NAUFlex (available at <https://nau.edu/nauflex/student>) is designed to help all students actively participate in their coursework during the required day and time of a course when they are not physically present in the classroom. This course design model allows students to be fully engaged with faculty and peers and receive the high-quality educational experience for which NAU is known.

CLASS SESSION RECORDINGS FOR STUDENTS AND FACULTY USE ONLY

Certain class sessions may be audio or video recorded to help reinforce live instruction during the COVID-19 pandemic. These recordings are for the sole use of the instructor and students enrolled in the course. Recordings will be stored in approved, accessible repositories. By enrolling, students agree to have their image and classroom statements recorded for this purpose, to respect the privacy of their fellow students, and university-owned intellectual property (including, but not limited to, all course materials) by not sharing recordings from their courses. Questions regarding restrictions on the use of classroom audio or video recordings may be addressed to the appropriate academic unit administrator.

SYLLABUS POLICY STATEMENTS

ACADEMIC INTEGRITY

NAU expects every student to firmly adhere to a strong ethical code of academic integrity in all their scholarly pursuits. The primary attributes of academic integrity are honesty, trustworthiness, fairness, and responsibility. As a student, you are expected to submit original work while giving proper credit to other people's ideas or contributions. Acting with academic integrity means completing your assignments independently while truthfully acknowledging all sources of information, or collaboration with others when appropriate. When you submit your work, you are implicitly declaring that the work is your own. Academic integrity is expected not only during formal coursework, but in all your relationships or interactions that are connected to the educational enterprise. All forms of academic deceit such as plagiarism, cheating, collusion, falsification or fabrication of results or records, permitting your work to be submitted by another, or inappropriately recycling your own work from one class to

another, constitute academic misconduct that may result in serious disciplinary consequences. All students and faculty members are responsible for reporting suspected instances of academic misconduct. All students are encouraged to complete NAU's online academic integrity workshop available in the E-Learning Center and should review the full *Academic Integrity* policy available at <https://policy.nau.edu/policy/policy.aspx?num=100601>.

COURSE TIME COMMITMENT

Pursuant to Arizona Board of Regents guidance (ABOR Policy 2-224, *Academic Credit*), each unit of credit requires a minimum of 45 hours of work by students, including but not limited to, class time, preparation, homework, and studying. For example, for a 3-credit course a student should expect to work at least 8.5 hours each week in a 16-week session and a minimum of 33 hours per week for a 3-credit course in a 4-week session.

DISRUPTIVE BEHAVIOR

Membership in NAU's academic community entails a special obligation to maintain class environments that are conducive to learning, whether instruction is taking place in the classroom, a laboratory or clinical setting, during course-related fieldwork, or online. Students have the obligation to engage in the educational process in a manner that does not interfere with normal class activities or violate the rights of others. Instructors have the authority and responsibility to address disruptive behavior that interferes with student learning, which can include the involuntary withdrawal of a student from a course with a grade of "W". For additional information, see NAU's *Disruptive Behavior in an Instructional Setting* policy at <https://nau.edu/university-policy-library/disruptive-behavior>.

NONDISCRIMINATION AND ANTI-HARASSMENT

NAU prohibits discrimination and harassment based on sex, gender, gender identity, race, color, age, national origin, religion, sexual orientation, disability, or veteran status. Due to potentially unethical consequences, certain consensual amorous or sexual relationships between faculty and students are also prohibited as set forth in the *Consensual Romantic and Sexual Relationships* policy. The Equity and Access Office (EAO) responds to complaints regarding discrimination and harassment that fall under NAU's *Nondiscrimination and Anti-Harassment* policy. EAO also assists with religious accommodations. For additional information about nondiscrimination or anti-harassment or to file a complaint, contact EAO located in Old Main (building 10), Room 113, PO Box 4083, Flagstaff, AZ 86011, or by phone at 928-523-3312 (TTY: 928-523-1006), fax at 928-523-9977, email at equityandaccess@nau.edu, or visit the EAO website at <https://nau.edu/equity-and-access>.

TITLE IX

Title IX is the primary federal law that prohibits discrimination on the basis of sex or gender in educational programs or activities. Sex discrimination for this purpose includes sexual harassment, sexual assault or relationship violence, and stalking (including cyber-stalking). Title IX requires that universities appoint a "Title IX Coordinator" to monitor the institution's compliance with this important civil rights law. NAU's Title IX Coordinator is Elyce C. Morris. The Title IX Coordinator is available to meet with any student to discuss any Title IX issue or concern. You may contact the Title IX Coordinator by phone at 928-523-3515, by fax at 928-523-0640, or by email at elyce.morris@nau.edu. In furtherance of its Title IX obligations, NAU will promptly investigate and equitably resolve all reports of sex or gender-based discrimination, harassment, or sexual misconduct and will eliminate any hostile environment as defined by law. Additional important information about Title IX and related student resources, including how to request immediate help or confidential support following an act of sexual violence, is available at <https://in.nau.edu/title-ix>.

ACCESSIBILITY

Professional disability specialists are available at Disability Resources to facilitate a range of academic support services and accommodations for students with disabilities. If you have a documented disability, you can request assistance by contacting Disability Resources at 928-523-8773 (voice), 928-523-6906 (TTY), 928-523-8747 (fax), or dr@nau.edu (e-mail). Once eligibility has been determined, students register with Disability Resources every semester to activate their approved accommodations. Although a student may request an accommodation at any time, it is best to initiate the application process at least four weeks before a student wishes to receive an accommodation. Students may begin the accommodation process by submitting a self-identification form online at <https://nau.edu/disability-resources/student-eligibility-process> or by contacting Disability Resources. The Director of Disability Resources, Jamie Axelrod, serves as NAU's Americans with Disabilities Act Coordinator and Section 504 Compliance Officer. He can be reached at jamie.axelrod@nau.edu.

RESPONSIBLE CONDUCT OF RESEARCH

Students who engage in research at NAU must receive appropriate Responsible Conduct of Research (RCR) training. This instruction is designed to help ensure proper awareness and application of well-established professional norms and ethical principles related to the performance of all scientific research activities. More information regarding RCR training is available at <https://nau.edu/research/compliance/research-integrity>.

MISCONDUCT IN RESEARCH

As noted, NAU expects every student to firmly adhere to a strong code of academic integrity in all their scholarly pursuits. This includes avoiding fabrication, falsification, or plagiarism when conducting research or reporting research results. Engaging in research misconduct may result in serious disciplinary consequences. Students must also report any suspected or actual instances of research misconduct of which they become aware. Allegations of research misconduct should be reported to your instructor or the University's Research Integrity Officer, Dr. David Faguy, who can be reached at david.faguy@nau.edu or 928-523-6117. More information about misconduct in research is available at <https://nau.edu/university-policy-library/misconduct-in-research>.

SENSITIVE COURSE MATERIALS

University education aims to expand student understanding and awareness. Thus, it necessarily involves engagement with a wide range of information, ideas, and creative representations. In their college studies, students can expect to encounter and to critically appraise materials that may differ from and perhaps challenge familiar understandings, ideas, and beliefs. Students are encouraged to discuss these matters with faculty.

Last revised January 6, 2021