Earth Lab Undergraduate Student Onboarding Manual

Tyler L. McIntosh
Last update: 5/30/24

Overview

Many Earth Lab scientists hire undergraduate students to help with fieldwork or other projects throughout the year. Working with Earth Lab is also an incredible opportunity for these students to gain exposure to scientific research! This document contains information that may be helpful for onboarding your students.

Some portions of this document are meant to be copied and pasted into an email or other document for your students so that you can pick and choose which sections or bullet points are most appropriate for your student. Copy-paste intended portions are in regular font. Sections in italics are contextual information that are not meant for copy-paste.

Student Persons of Interest (POIs)

If you are working with a student who is not affiliated with CU Boulder, you will need to request an IdentiKey and email for them through the CU 'person of interest' (POI) process (if they will need access to these). The easiest way to do this is by requesting an IdentiKey from CIRES HR (cireshr@colorado.edu or the HR representative associated with the student) through the POI process. An IdentiKey and CU email are necessary for students to get access to Research Computing, as well as the RCR CITI course.

Responsible Conduct of Research (RCR) CITI Training

CITI Responsible Conduct of Research is an online course aimed at educating on appropriate research practices. CU Boulder requires that all professional researchers complete RCR. The CITI requirement applies to all research regardless of review categories. Additionally, any faculty, other senior personnel, postdoctoral fellow ***or student*** *who is paid from a National Science Foundation (NSF) or USDA NIFA grant award is required to complete RCR. CU Boulder recommends that all student researchers complete RCR, regardless of their funding source. The training is actually a quite good introduction to guiding scientific principles on topics like publication practices and authorship, peer review, collaborative science, etc.*

*Note that **if students or postdocs are required** to complete RCR and complete an online RCR course other than through CU Boulder, **they must submit an RCR Completion Form**. They must sign the original and submit it to the RCR Coordinator (rcr@colorado.edu). They must also provide a copy to their mentor/PI for the NIH- or NSF-funded award.*

In order to complete RCR through CU Boulder, students will need a colorado.edu email address! See above on student POIs.

- ***IF REQUIRED:*** In the next few days or next week, please take an hour or so to complete the online CITI course on responsible conduct of research (RCR) (unless you've already completed it in the past 3 years). This is a course that is **required** for students funded through the grant that you will be working on. It provides a great overview of some basic scientific topics like publication practices and authorship, peer review, collaborative science, etc. Here's the CITI RCR instructions for CU Boulder students. Once you complete it, please fill out this form and then email it to both your supervisor and rcr@colorado.edu.
- ***IF RECOMMENDED:*** In the next few days or next week, it would be great if you could take an hour or so to complete the online CITI course on responsible conduct of research (RCR) (unless you've already completed it). This short course provides a great overview of some basic scientific topics like publication practices and authorship, peer review, collaborative science, etc. Here's the CITI RCR instructions for CU Boulder students. Once you complete it, please fill out this form and then email it to both your supervisor and rcr@colorado.edu.

Timesheet Submission

For your timesheets, you'll first log into MyCUInfo : http://mycuinfo.colorado.edu/ with your identikey. Next, you'll find the "**My Leave**" tile. If the tile isn't already on your homepage, you can try to search for My Leave in the search bar at the top of the page.

There is a two-step process in MyLeave for submitting your timesheet: you first add your hours for each individual day worked on the Calendar. Next, you navigate to the "timesheet" tab where it summarizes all of your hours for the selected pay period – you'll scroll to the bottom to "Submit" the overall timesheet. **This document is a step by step guide that has additional instructions.**

Fieldwork

Some Earth Lab projects involve fieldwork in various environments. If you are receiving this, you're involved in one of these projects! Please review the below information and fill out all surveys and/or forms.

Field Gear

A number of risks are involved with our fieldwork, primarily as a result of the remoteness and high elevation of the locations involved. Each day you will need to be prepared for sun, wind, rain, dehydration, difficult terrain, etc. **Please review the field gear list that we have prepared**, and let us know if there are any required items that you do not have and which acquiring would pose a financial burden or hardship.

Field Survey

We have a number of questions for you that are important for field safety, logistics, and coordination. **Please fill out this survey at your earliest convenience**.

Field Waiver

We need you to fill out a CU Boulder risk management waiver prior to going out in the field. **Please download, fill out, and email the PDF here to your supervisor or mentor.**

Software and system access

Different Earth Lab projects require different access to different software and systems. Below are a few that you may need students to access. Note that there are student computers available in the office if students do not have their own personal laptop.

QGIS

QGIS is an open-source software alternative to ESRI's ArcGIS suite. A benefit of learning QGIS in addition to ArcGIS is that you will be able to use it as a tool even if, in the future, you work for an organization without access to ArcGIS, which may be prohibitively expensive for some organizations. If you've used ArcGIS before, you'll find that QGIS feels relatively familiar! You can download QGIS here.

R & RStudio

R is an open-source coding language originally developed for statistical analysis. It has evolved, along with Python, into one of the most commonly used coding languages for scientific

computing. RStudio is an application that has been developed to make R more user friendly. You can download R and RStudio here.

There are TONS of resources for learning R and its many packages. It can be a bit overwhelming! Below are some options for learning R.

If you want to dive right into an interactive tutorial on how to code in R, check out Swirl.

One cool resource is "The Big Book of R," which is an open source compilation of OTHER R books and resources (most of which are in turn open source). This is a good bookmark and has everything from pretty general content to fairly niche content.

This Introduction to R course seems like a useful one, and is aimed at people who have never used R before—or any other coding language. One thing to note is that it focuses on base R rather than the Tidyverse, which is a very common set of packages often used for working with data in R. However, it is always good to know base R, and the basic information in it about how R works (e.g. objects, functions, etc) remains the same no matter what way you decide to work in R. The course also has videos if you prefer to learn that way. Here's the open source book that goes along with the course.

As you continue to develop your skills you will begin to install and use more packages that have been created for performing different tasks. These packages contain specialized functions. Whenever you are using a new library, I recommend looking for 'cheatsheets' that the R community often makes for different libraries. There are thousands of libraries out there—and so many functions in each library—that these cheatsheets serve as a way to remember what common functions are in a library, what they do, and what their basic syntax is. For example, 'ggplot' is a very commonly used package for visualizing data, and by googling "R ggplot cheatsheet" you can find something like this: https://statsandr.com/blog/files/ggplot2-cheatsheet.pdf.

When it comes down to it, the best way to continue learning any coding language is through trial and error. Try not to be frustrated when your code throws errors, as even the best coders will have errors! Instead, becoming a good coder is about learning to recognize different errors and figure out what might have caused them.

Access to CU Research Computing (CURC)

In order for students to access data on CURC or Earth Lab's Petalibrary account, or for them to process data on CU's High Performance Computing (HPC) system, they will need a CURC account.

If the student has a CU Identikey, they can just sign up for an account as normal at https://rcamp.rc.colorado.edu/accounts/account-request/create/verify/ucb. Once they have an

account, email rc-help@colorado.edu to request access for them to any user groups/accounts on the system they would need to access Earthlab software/files.

If the student does NOT have an Identikey, you will need to request one through the POI process (see above). Once that process is finished, they can sign up for a CURC account per the above.

- Please visit *https://rcamp.rc.colorado.edu/accounts/account-request/create/verify/ucb* and sign up for a research computing account. Once you have made an account, please email your supervisor to let them know and they will ensure that CURC gives your account access to Earth Lab software and files.

Metashape

To get a student access to Metashape on CURC, you will need to request that someone at CURC move the Metashape initialization file into the student's home directory so that they can start Metashape with the following commands from the command line:

- *source ~/.metashape*
- *metashape*

Background Science Information

At the beginning of an undergraduate student spending a summer or semester working with Earth Lab it is helpful to provide them with some foundational information about topics which are essential to their work, but which they may be unfamiliar with (or need a refresher on).

For many Earth Lab projects, students will need familiarity with spatial data formats and processing in addition to topic-specific scientific background. Below are examples of some resources that may be appropriate to provide students with, depending on your project. If you have additional topics that students should be provided information about, please find good introductory resources for those topics and add them below under a new heading for others to use.

Project descriptions

Make sure to provide your students with the full description of your project that was used for acquiring funding for it. Even if a student is only working on a small section of a project, it is important for them to understand the broader scientific context and importance of their work.

Remote Sensing & Geographic Information Systems (GIS)

- Spatial data is a bit different from non-spatial data, as it needs to take into account how the data is related to positions on the Earth's surface. When people work with spatial data, they often work in some type of geographic information system (GIS). Learn more about GIS here:
    - What is GIS?
- When working with spatial data, you will often have two different types of data: Vector and Raster. For our projects you will work with both. You can familiarize yourself with the differences between them here:
    - Vector vs Raster data
- Our projects heavily rely on remotely sensed data from drones, airplanes, and satellites. Familiarize yourself with some basics of remote sensing and related concepts here:
    - Fundamentals of Remote Sensing
- Drone technology allows us to capture images of ecosystems at a much higher resolution than satellites. Essentially, we use drones to take photographs of the land they fly over, and then weave those photos together into a single image, called an 'orthomosaic'. Learn more about orthomosaics below.
    - What is an orthomosaic?
    - To learn more about mapping with drones, including basics of flight planning and data processing, check out this short chapter on "How to make maps with drones"
- Spatial data can be stored in many different formats. One way that you are probably familiar with through using Google Maps or similar applications is a 'tile package'. In our work we sometimes create and then use tile packages. Learn about them below:
    - About tile packages
    - Why should I create tile packages?