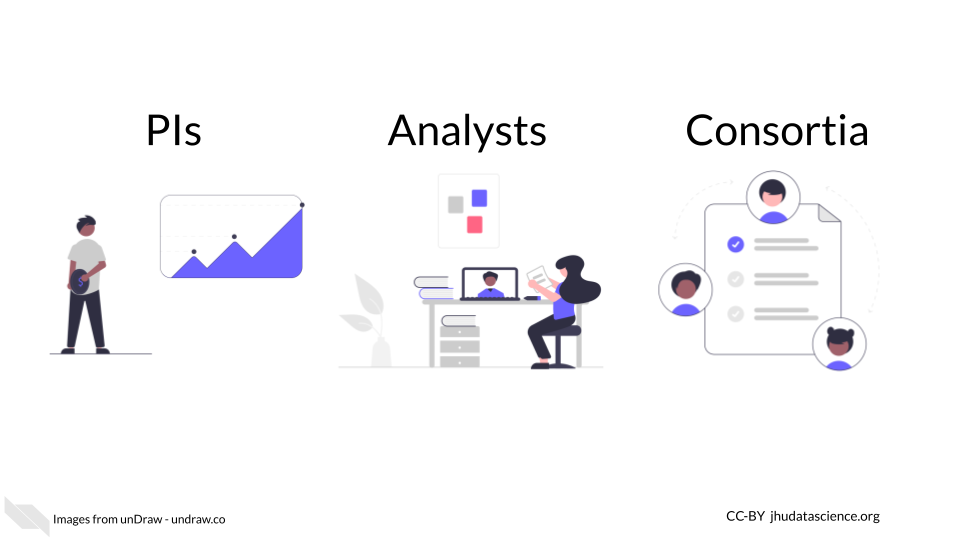
Getting Started on AnVIL

August 29, 2025

Table of Contents

# About this Book

This book is part of a series of books for the Genomic Data Science Analysis, Visualization, and Informatics Lab-space (AnVIL) of the National Human Genome Research Institute (NHGRI). Here, we present opinionated step-by-step guides for setting up accounts focused on three personas: PIs, Analysts, and Consortia.



### Skills Level

Please choose the closest matching persona from the lefthand menu.

*Genetics*

**Novice**: no genetics knowledge needed

*Programming skills*

**Novice**: no programming experience needed

### How to Use This Book

This book is not intended to be read through sequentially, rather, it is a collection of guides that you can reference based on your needs. It is divided into two major sections:

**Account Setup**

Step-by-step instructions for new AnVIL users to set up their accounts and start using the AnVIL platform. We have included recommendations for configuring your accounts based on several common use cases:

* **PIs and Lab Managers**: managing a team of researchers working on AnVIL
* **Data Analysts**: joining a team working on AnVIL
* **Consortia**: using AnVIL as part of a research consortium

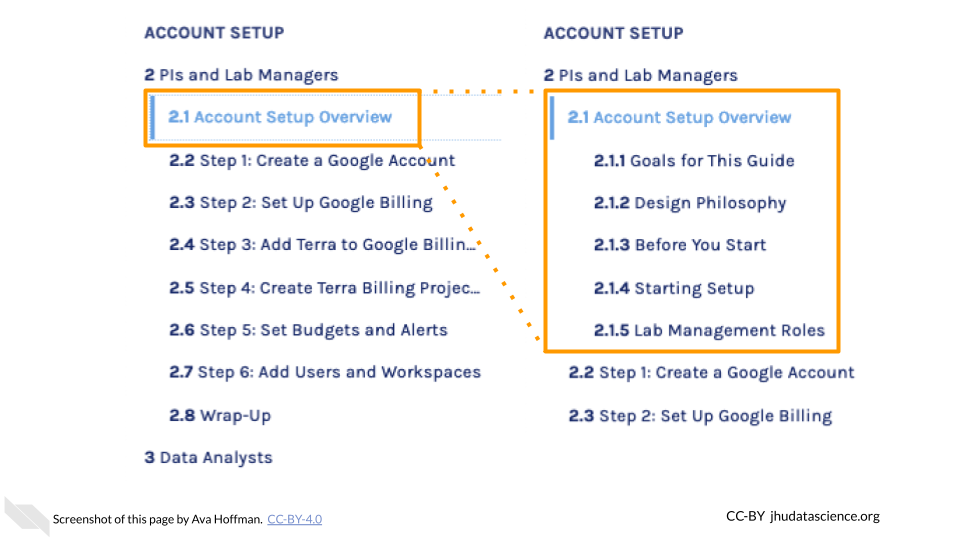
**Working on AnVIL**

Examples and walkthroughs of common tasks on the AnVIL platform:

* **Workspaces**: how to create and clone research spaces on AnVIL
* **Tools**: how to run common tools including Jupyter Notebooks, Galaxy, and RStudio
* **Data**: how to find and access AnVIL datasets, as well as upload and manage your own data
* **Workflows**: how to find and run existing automated data processing pipelines, and how to customize and share your own

### Activate scroll\_highlight Feature {#scroll-highlight}

Note that some sections of this book cover steps in a lot of detail. When navigating the table of contents, you can click subsection (e.g., 2.1, 4.3) headers a second time to expand the table of contents and enable the scroll\_highlight feature. This can help you follow the separate steps within more clearly.



# 1 Student Guide

## 1.1 Activity One

Many of the AnVIL books that are designed for teaching a class or workshop contain information aimed at the instructor.

You might want to create a “student guide” that contains a different subset of Rmd files from your book, or renders to a different output format (e.g. word document). You can specify the output and Rmd files that will be used for the student guide using the \_output.yml and \_bookdown.yml files in the student-guide directory.

By default, repositories based on AnVIL\_Template create a docx file containing this chapter as an example. Once you have adjusted student-guide/bookdown.yml to include the chapters you want in your student guide, you can link to it by replacing {repo-name} in the following URL:

https://github.com/jhudsl/{repo-name}/raw/main/student-guide/docs/Student\_Guide.docx

When someone clicks this link, they will download the docx file. For example, here is the link for the student guide from AnVIL\_Template:

<https://github.com/jhudsl/AnVIL_Template/raw/main/student-guide/docs/Student_Guide.docx>

## 1.2 Activity Two

Steps of the guide could go here.

# About the Authors

These credits are based on our [course contributors table guidelines](https://github.com/jhudsl/DaSL_Course_Template_Bookdown/wiki/How-to-give-credits).

***In*** [***memory***](https://galaxyproject.org/jxtx/) ***of*** [***James Taylor***](https://genomebiology.biomedcentral.com/articles/10.1186/s13059-020-02016-0)***, who was instrumental in initiating this project.***

| Credits | Names |
| --- | --- |
| **Pedagogy** |  |
| Lead Content Instructors | [Katherine Cox](https://publichealth.jhu.edu/faculty/4170/katherine-cox), [Ava Hoffman](https://www.avahoffman.com/) |
| Content Contributors | [Kai Kammers](https://www.hopkinsmedicine.org/profiles/details/kai-kammers), [Frederick Tan](https://carnegiescience.edu/frederick-tan), [Sarah Wheelan](https://publichealth.jhu.edu/faculty/1902/sarah-wheelan) |
| Content Editors/Reviewers | Natalie Kucher, [Jeff Leek](https://jtleek.com/), Valerie Reeves, [Frederick Tan](https://carnegiescience.edu/frederick-tan) |
| Content Directors | [Jeff Leek](https://jtleek.com/), [Frederick Tan](https://carnegiescience.edu/frederick-tan) |
| Content Consultants | [Allie Cliffe](https://www.linkedin.com/in/allie-cliffe/), Candace Patterson |
| **Production** |  |
| Content Publisher | [Ira Gooding](https://publichealth.jhu.edu/faculty/4130/ira-gooding) |
| **Technical** |  |
| Template Publishing Engineers | [Candace Savonen](https://www.cansavvy.com/), [Carrie Wright](https://carriewright11.github.io/) |
| Publishing Maintenance Engineer | [Candace Savonen](https://www.cansavvy.com/) |
| Technical Publishing Stylists | [Carrie Wright](https://carriewright11.github.io/), [Candace Savonen](https://www.cansavvy.com/) |
| Package Developers ([ottrpal](https://github.com/jhudsl/ottrpal)) | [John Muschelli](https://johnmuschelli.com/), [Candace Savonen](https://www.cansavvy.com/), [Carrie Wright](https://carriewright11.github.io/) |
| **Funding** |  |
| Funder | [National Human Genome Research Institute (NHGRI)](https://www.genome.gov/) #5U24HG010263 |
| Funding Staff | Fallon Bachman, Jennifer Vessio, Emily Voeglein |

## ─ Session info ───────────────────────────────────────────────────────────────  
## setting value  
## version R version 4.3.2 (2023-10-31)  
## os Ubuntu 22.04.4 LTS  
## system x86\_64, linux-gnu  
## ui X11  
## language (EN)  
## collate en\_US.UTF-8  
## ctype en\_US.UTF-8  
## tz Etc/UTC  
## date 2025-08-29  
## pandoc 3.1.1 @ /usr/local/bin/ (via rmarkdown)  
##   
## ─ Packages ───────────────────────────────────────────────────────────────────  
## package \* version date (UTC) lib source  
## bookdown 0.43 2025-04-15 [1] CRAN (R 4.3.2)  
## cachem 1.0.8 2023-05-01 [1] RSPM (R 4.3.0)  
## cli 3.6.2 2023-12-11 [1] RSPM (R 4.3.0)  
## devtools 2.4.5 2022-10-11 [1] RSPM (R 4.3.0)  
## digest 0.6.34 2024-01-11 [1] RSPM (R 4.3.0)  
## ellipsis 0.3.2 2021-04-29 [1] RSPM (R 4.3.0)  
## evaluate 1.0.4 2025-06-18 [1] CRAN (R 4.3.2)  
## fastmap 1.1.1 2023-02-24 [1] RSPM (R 4.3.0)  
## fs 1.6.3 2023-07-20 [1] RSPM (R 4.3.0)  
## glue 1.7.0 2024-01-09 [1] RSPM (R 4.3.0)  
## htmltools 0.5.7 2023-11-03 [1] RSPM (R 4.3.0)  
## htmlwidgets 1.6.4 2023-12-06 [1] RSPM (R 4.3.0)  
## httpuv 1.6.14 2024-01-26 [1] RSPM (R 4.3.0)  
## knitr 1.50 2025-03-16 [1] CRAN (R 4.3.2)  
## later 1.3.2 2023-12-06 [1] RSPM (R 4.3.0)  
## lifecycle 1.0.4 2023-11-07 [1] RSPM (R 4.3.0)  
## magrittr 2.0.3 2022-03-30 [1] RSPM (R 4.3.0)  
## memoise 2.0.1 2021-11-26 [1] RSPM (R 4.3.0)  
## mime 0.12 2021-09-28 [1] RSPM (R 4.3.0)  
## miniUI 0.1.1.1 2018-05-18 [1] RSPM (R 4.3.0)  
## pkgbuild 1.4.3 2023-12-10 [1] RSPM (R 4.3.0)  
## pkgload 1.3.4 2024-01-16 [1] RSPM (R 4.3.0)  
## profvis 0.3.8 2023-05-02 [1] RSPM (R 4.3.0)  
## promises 1.2.1 2023-08-10 [1] RSPM (R 4.3.0)  
## purrr 1.0.2 2023-08-10 [1] RSPM (R 4.3.0)  
## R6 2.5.1 2021-08-19 [1] RSPM (R 4.3.0)  
## Rcpp 1.0.12 2024-01-09 [1] RSPM (R 4.3.0)  
## remotes 2.4.2.1 2023-07-18 [1] RSPM (R 4.3.0)  
## rlang 1.1.6 2025-04-11 [1] CRAN (R 4.3.2)  
## rmarkdown 2.25 2023-09-18 [1] RSPM (R 4.3.0)  
## sessioninfo 1.2.2 2021-12-06 [1] RSPM (R 4.3.0)  
## shiny 1.8.0 2023-11-17 [1] RSPM (R 4.3.0)  
## stringi 1.8.3 2023-12-11 [1] RSPM (R 4.3.0)  
## stringr 1.5.1 2023-11-14 [1] RSPM (R 4.3.0)  
## urlchecker 1.0.1 2021-11-30 [1] RSPM (R 4.3.0)  
## usethis 2.2.3 2024-02-19 [1] RSPM (R 4.3.0)  
## vctrs 0.6.5 2023-12-01 [1] RSPM (R 4.3.0)  
## xfun 0.52 2025-04-02 [1] CRAN (R 4.3.2)  
## xtable 1.8-4 2019-04-21 [1] RSPM (R 4.3.0)  
## yaml 2.3.10 2024-07-26 [1] CRAN (R 4.3.2)  
##   
## [1] /usr/local/lib/R/site-library  
## [2] /usr/local/lib/R/library  
##   
## ──────────────────────────────────────────────────────────────────────────────

# 2 References