Planning and tracking work with GitHub





Objectives



- Describe some of the general benefits of an issue tracking system
- Review the features of GitHub issues, including issue templates
- Discuss what makes a "good" issue
- Illustrate how issue scope can have broader implications
- Briefly mention how we currently use GitHub issues to plan our work at the Data Lab

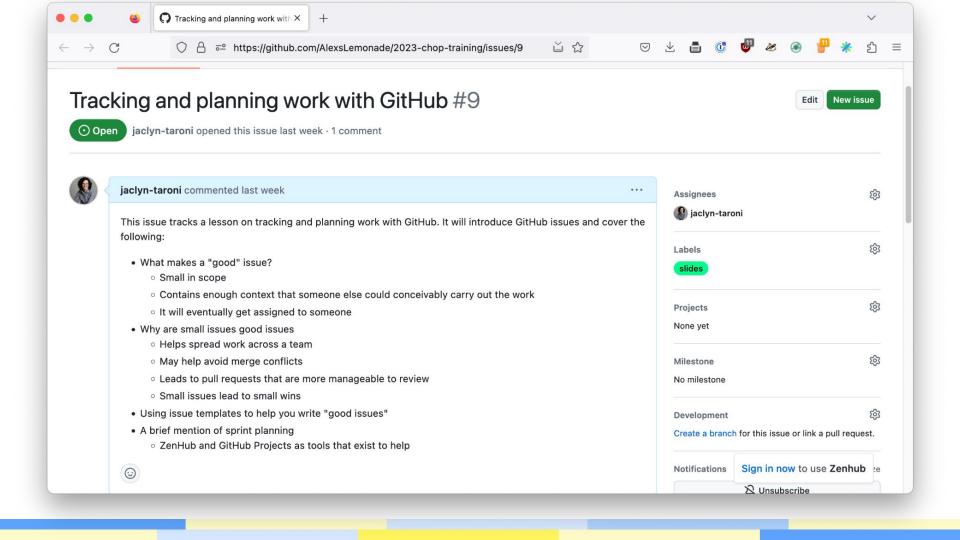


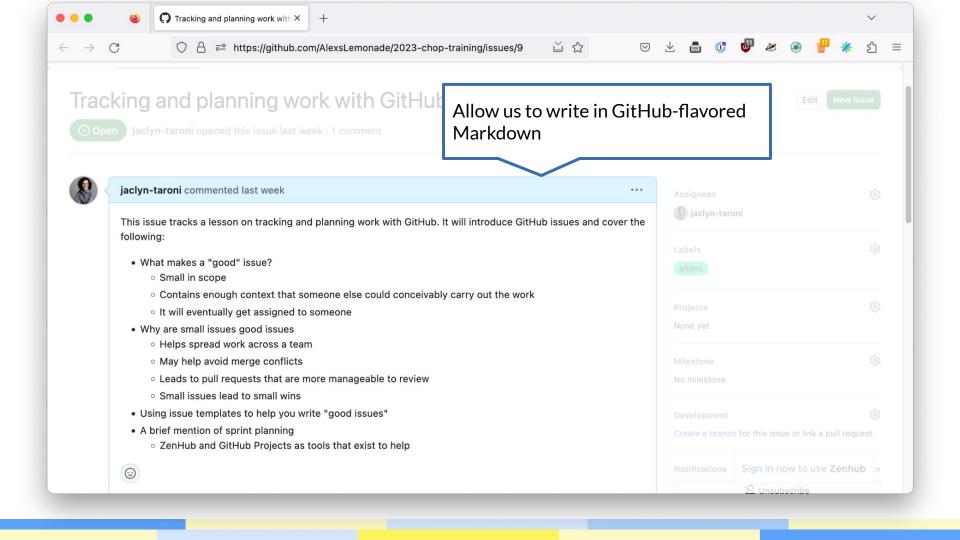


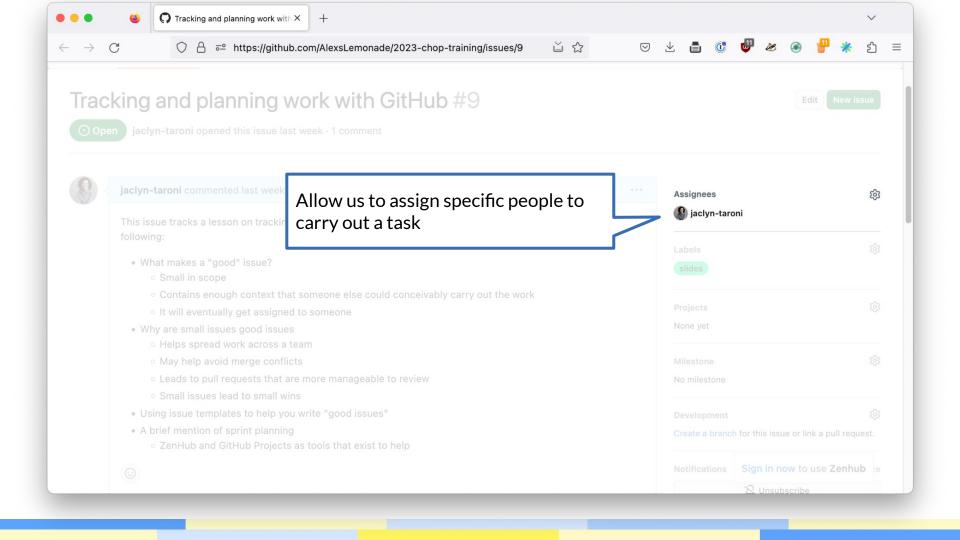
- Reduces the cognitive burden associated with tracking tasks in our head
- Allows us to have efficient conversations about how to tackle work
- Gives us the ability to surface known limitations that we don't intend to fix
- For scientific projects that are public or will be public eventually, there's the additional benefit of capturing scholarship that sometimes remains hidden

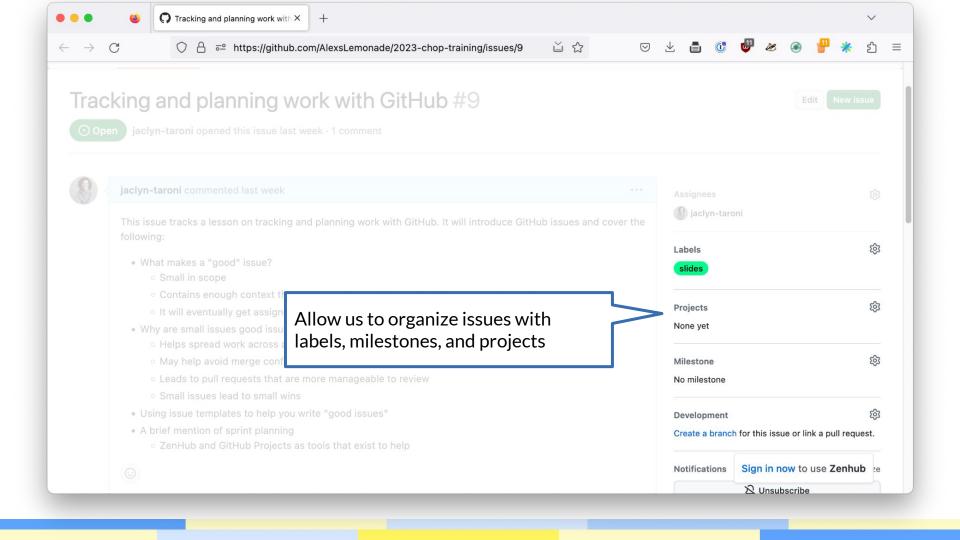


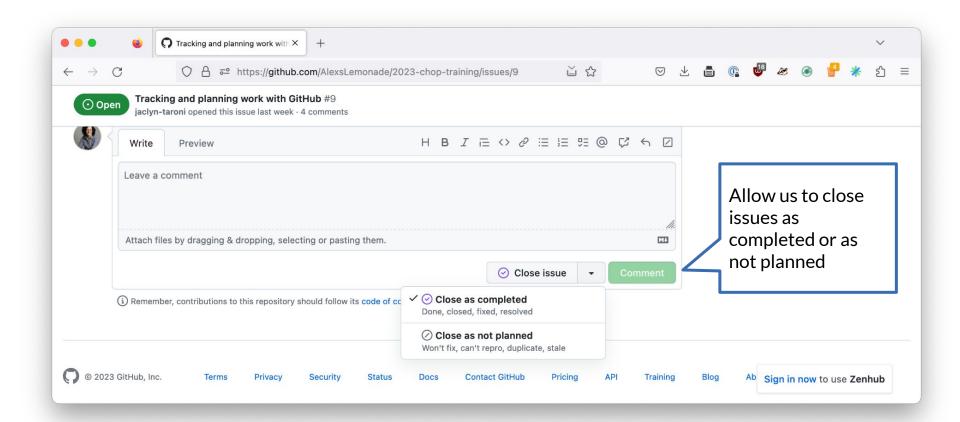
GitHub issues allow us to track issues alongside our code

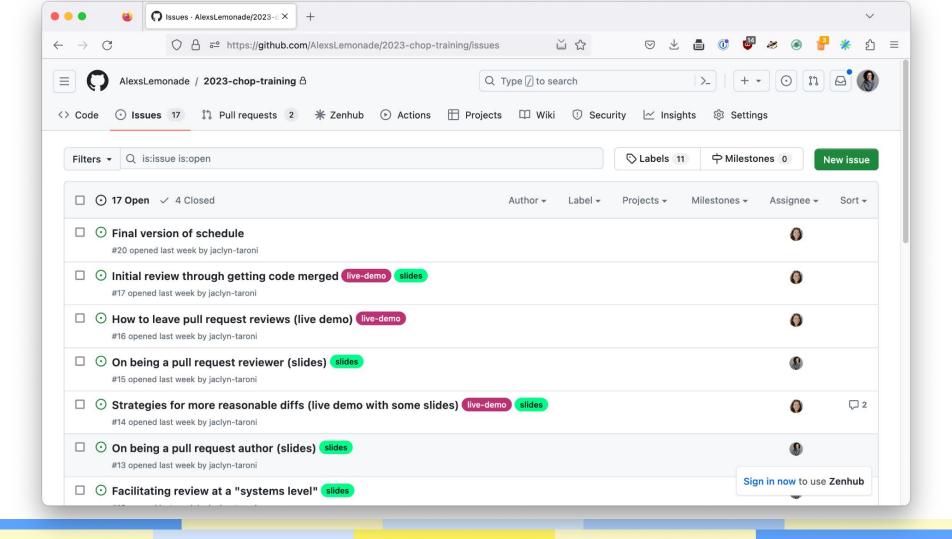


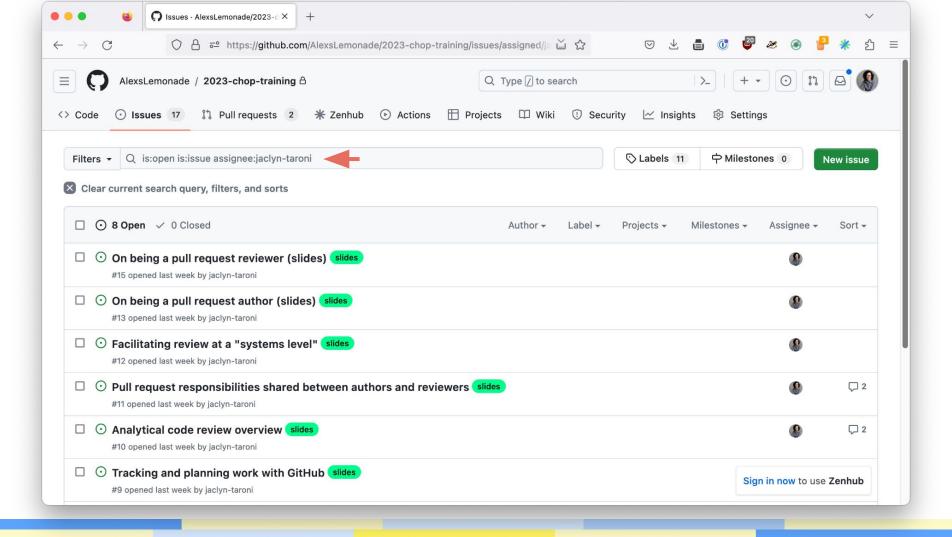


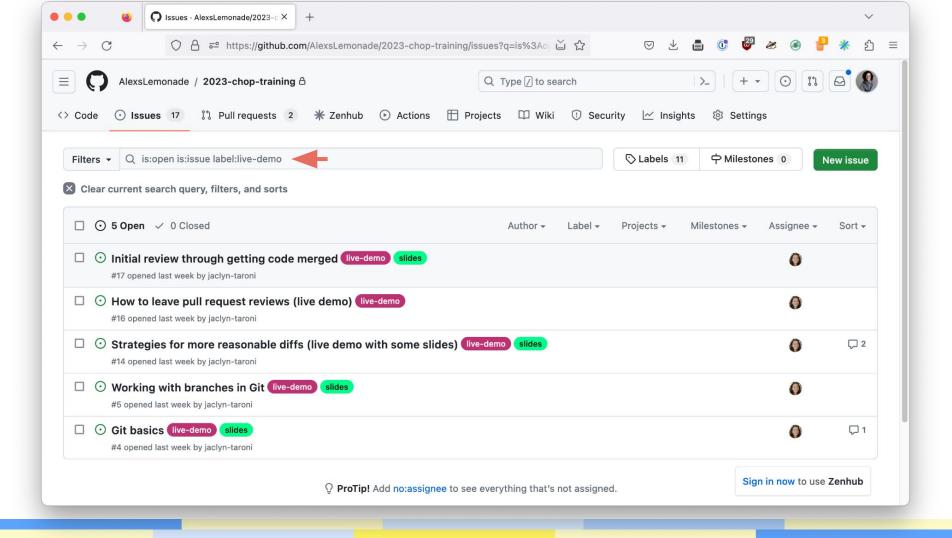


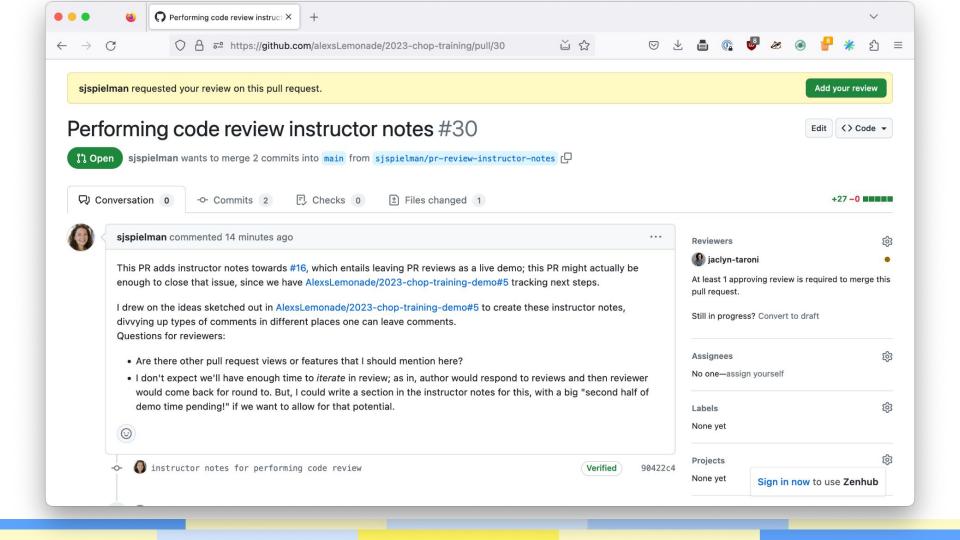










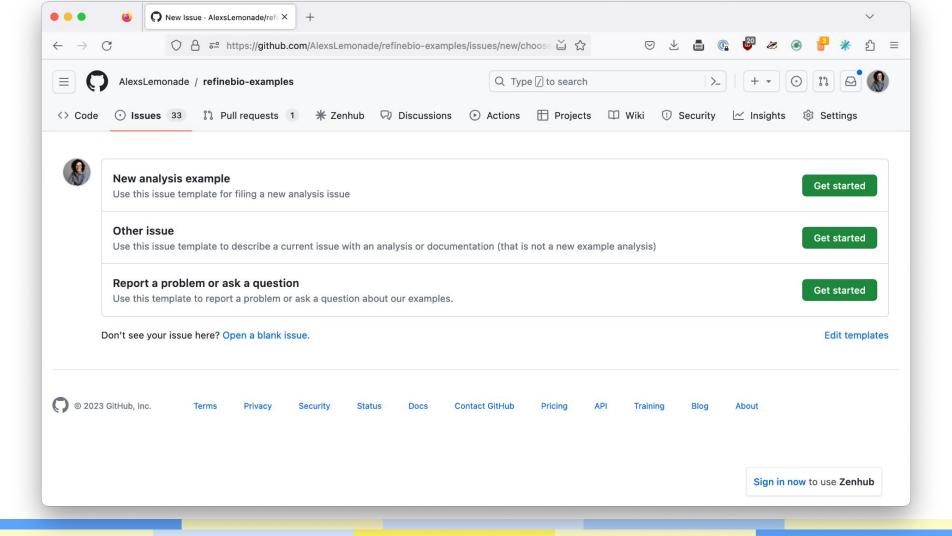


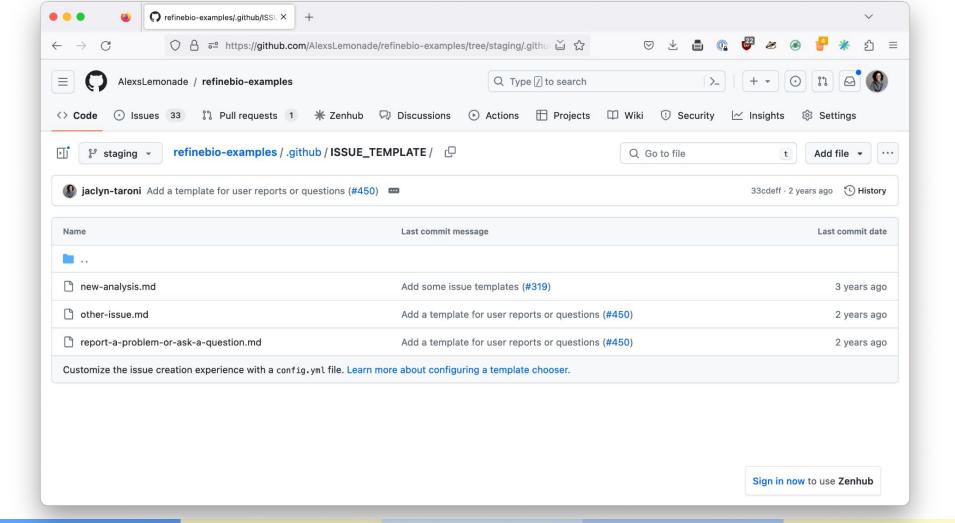


GitHub allows us to create multiple issue templates for a repository

We can create Markdown files with some features:

- Structure we find useful (e.g., a checklist or set of headers)
- A YAML header indicating things like labels and prefixes for the title





```
name: New analysis example
about: Use this issue template for filing a new analysis issue
title: 'New Analysis Example:'
labels: new analysis example
assignees: ''
### What are the goals of this new example analysis?
<!-- What should this new example show that's not currently shown in the existing samples -->
<!-- What was the background and context that lead to this new example? -->
<!-- Links to github comments or related issues are also helpful -->
### What kind of dataset will this need?
<!-- Is there a particular species, sample size, experimental model, etc. that's needed -->
### What steps should be included in this analysis?
<!-- An outline of what steps should probably be taken would be nice if possible -->
### What packages/methods do you recommend using or looking into for this analysis?
<!-- Do you have helpful references and resources we should consult for creating this example? -->
```

Milestones and projects allow you to further organize your issues

Milestones let you track groups of issues or pull requests and assign a due date. Milestones also let you view the completion percentage and find open and closed issues quickly (<u>GitHub docs</u>).

Projects can organize issues as a table, spreadsheet, or task-board and let you represent complex relationships between issues with tasklists (<u>GitHub docs</u>).

We don't use these at the Data Lab; we use something called <u>ZenHub</u> instead because it allows us to group issues in sprints and represent dependencies between issues (the latter is supposedly coming to GitHub projects in the future!).

What makes a "good" issue?

What makes a "good" issue?



- Title and labels are informative about the contents of the issue "at a glance."
- Contains enough **context** for someone onboarding to the project or returning to the project after some time to be assigned to contribute to the discussion.
- Communicates any gotchas or known unknowns.
- Records ideas about how to accomplish the task if the person filing the issue has them.
- Communicates the requirements for satisfying or closing the issue.

A good issue is focused



Roughly, the smallest unit of work something could be broken down into to result in a single, functional pull request that is manageable for review.

Some caveats



- Sometimes the author of an issue (or the person that initially comes up with an idea) can't fully flesh out the issue, either because of a time constraint or because some research is required.
 - We often file issues like that prefixed with [PLACEHOLDER] or [WIP] internally.
- Sometimes you need to have a discussion before work can begin. Filing an issue can facilitate an asynchronous discussion and helps make decisions transparent.
 - We often label these with a discussion label internally.

Why are focused issues good issues?

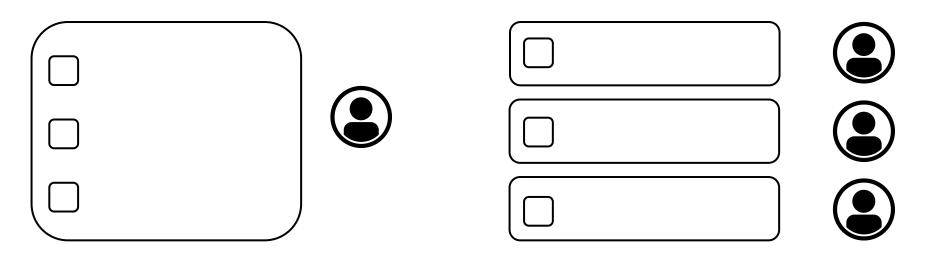
Focused issues set us up for focused pull requests



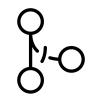
A focused pull request is a more timely and thoroughly reviewed pull request.

Focused issues allow us to spread work across the team

Without duplicating efforts. But that means you need to always assign yourself if you're working on something.



Focused issues might help us **avoid merge conflicts**



If you're combining focused issues with other practices – like planning work as a team and making sure PR branches aren't long-lived – you can assign issues in a manner that minimizes the chances that people will be touching the same parts of the codebase.

But that means you have to stick to what's on the issue!

Focused issues help our team achieve small wins



Hitting the closed as completed button feels great!

Focused issues let us see the meaningful progress we are making, which can boost the way we feel.

Using issues for planning as a team

Quick tips from Data Lab practices



The Data Lab teams plan their work (i.e., pick issues to tackle) in two week chunks ("sprints"). Here are some things that we've learned.

Make sure you **use assignments** on issues. Bonus points if you have an idea of who will review ahead of time! Review is part of everyone's workload.

You can also fall into a pattern where more folks are reviewing an individual pull request than are necessary and gum up the works that way. Planning ahead could help avoid that, too.

Quick tips from Data Lab practices



The Data Lab teams plan their work (i.e., pick issues to tackle) in two week chunks ("sprints"). Here are some things that we've learned.

It can be tempting to – or sometimes necessary to – work on issues that weren't originally planned during a set time period.

If everything you do will need to be reviewed, unplanned issues creeping in can cause more than one person's workload to balloon unexpectedly.

Quick tips from Data Lab practices



The Data Lab teams plan their work (i.e., pick issues to tackle) in two week chunks ("sprints"). Here are some things that we've learned.

We now have a meeting set up where we nominate issues for technical design discussions. Sometimes an issue isn't enough and having a synchronous discussion upfront will save both the author and reviewer time.

Folks might end up doubling back and redoing work if there isn't agreement on approach upfront!

Summary

- GitHub issues let us track and organize our tasks alongside our code, which allows us to connect pull requests to issues
- A "good issue" contains enough information to facilitate the person directly responsible for carrying out the work getting up to speed quickly
 - Templates can support us in writing good issues
- Issues that are focused in scope help us create pull requests that are manageable in size and spread work across multiple people on the team when necessary
- Using issues helps the Data Lab to plan <u>review workloads</u>, in addition to coding workloads, but we've needed to combine them with other practices