BGGN 213

Course Summary & GitHub Portfolio

Class 20

Barry Grant

UC San Diego

http://thegrantlab.org/bggn213

Today’s Menu

- Summary of major learning goals,
- Course discussion and feedback (https://etherpad.wikimedia.org/p/bggn213_f25)
- Course evaluation (official)
- Final exam
  - ➡ Test structure, guidelines and rules
  - ➡ Topics and example questions
  - ➡ Exam preparation, discussion and open study
- Polish our GitHub content and publish your own website portfolios

Today's Menu

- Summary of major learning goals,
- Course discussion and feedback (https://etherpad.wikimedia.org/p/bggn213_f25)
- Course evaluation (official)
- Final exam
  - Test structure, guidelines and rules
  - Topics and example questions
  - Exam preparation, discussion and open study

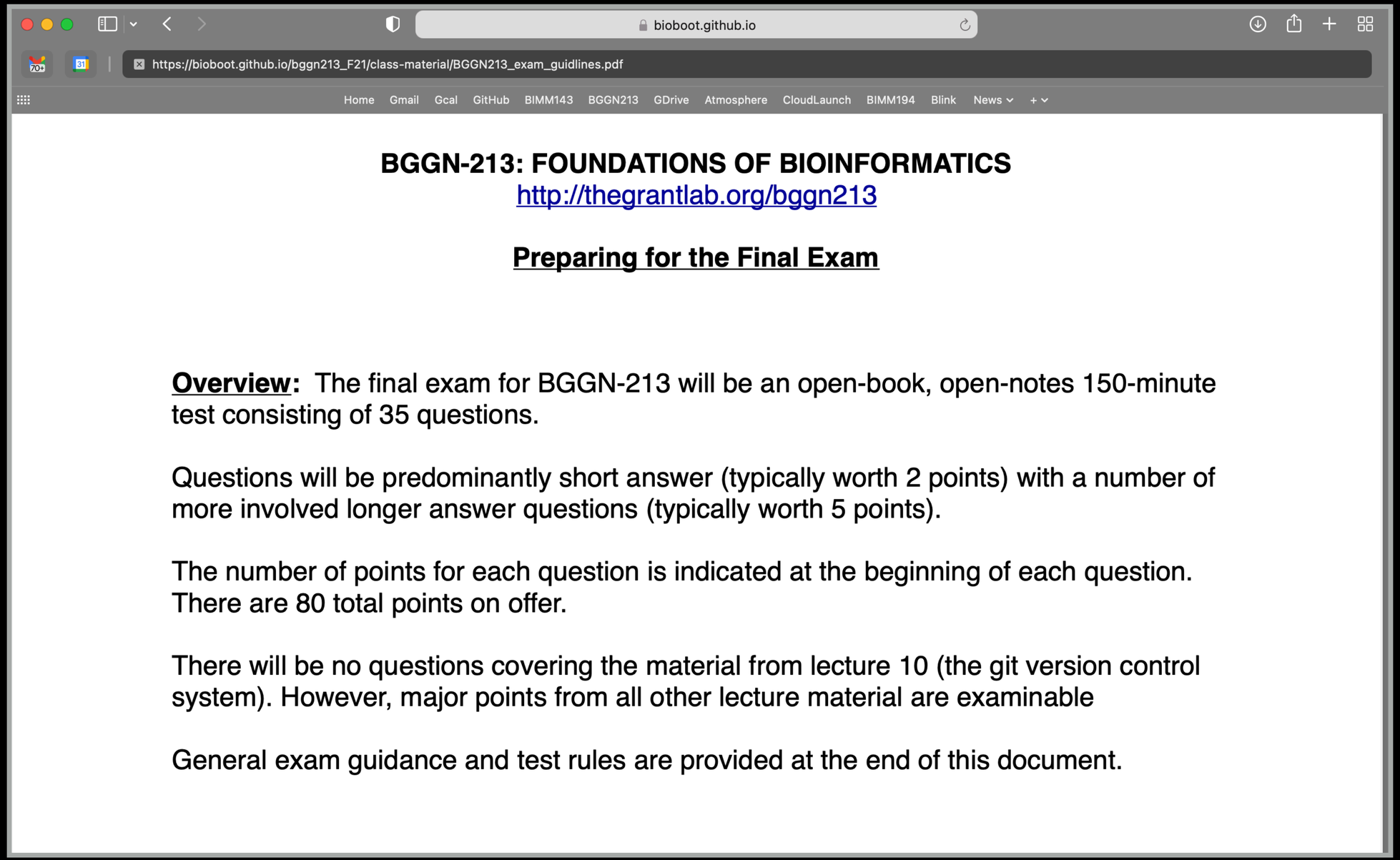
- Polish our GitHub content and publish your own website portfolios

Today's Menu

- Summary of major learning goals,
- Course discussion and feedback (https://etherpad.wikimedia.org/p/bggn213_f25)
- Course evaluation (official)
- Project troubleshooting
  - Test structure, guidelines and rules
  - Topics and example questions
  - Exam preparation, discussion and open study

- Polish our GitHub content and publish your own website portfolios

https://bioboot.github.io/bggn213_F22/class-material/BGGN213_exam_guidlines.pdf



BGGN-213: FOUNDATIONS OF BIOINFORMATICS
http://thegrantlab.org/bggn213

Preparing for the Final Exam

Overview: The final exam for BGGN-213 will be an open-book, open-notes 150-minute test consisting of 35 questions.

Questions will be predominantly short answer (typically worth 2 points) with a number of more involved longer answer questions (typically worth 5 points).

The number of points for each question is indicated at the beginning of each question. There are 80 total points on offer.

There will be no questions covering the material from lecture 10 (the git version control system). However, major points from all other lecture material are examinable

General exam guidance and test rules are provided at the end of this document.

Q1. Did you enjoy this course? (Rank in relation to others you have experienced at UCSD)?

Q2. Should this course be offered again?

Q3. If so what changes would you recommend for this course? (e.g. more/less DataCamp & Projects)

Q4. Was the course effectively organized (lecture and lab material online vs handout or TritonEd site)?

Q5. What advice would you give to another student who is considering taking this course?

Q6. Considering both the limitations and possibilities of the subject matter and the course, how would you rate the overall effectiveness of this course and instructor?

Q7. Do you agree or disagree - The course developed my abilities and skills for the subject?

Q8. On average, how many hours per week have you spent on this course, including attending classes, doing homework's and assignments?

Q9. Any other comments you would like to share?

EtherPad: https://etherpad.wikimedia.org/p/bggn213_f25

~~**Form Version:** https://tinyurl.com/bggn213-summary~~

Thank you very much!

Announcement!

Bonus:
Bioinformatics & Genomics in Industry
Live Stream Video

Enjoy a set of short open ended guest lectures from leading genomic scientists at **Illumina Inc.**, **Synthetic Genomics Inc.**, and the **La Jolla Institute for Allergy and Immunology**. Feel free to contact these scientists for networking and to have your questions about industry careers in Bioinformatics and Genomics answered.

Bonus:
GitHub Spit & Polish

bioboot.github.io/bimm143-1/

Home Gmail Gcal Bitbucket GitHub BIMM143_F18 BGGN213_S18 BIMM-194 GDocs Disqus Blink News Atmosphere Galaxy + MMTF

bioboot/bimm143_serina_f18: Serina's Fall 2018 class repository fork | Serina's BIMM 143 Class Repository | Serina's Bioinformatics Class (BIMM143,... | Bioinformatics Class BIMM-143 | Introduction to Bioinformatics (BIMM143)

Introduction to Bioinformatics

A demo site of students cool class web site

View the Project on GitHub
bioboot/bimm143-1

This project is maintained by bioboot

Hosted on GitHub Pages — Theme by orderedlist

Bioinformatics Class

This is my repository for my Bioinformatics class from UC San Diego in S18.

Index of Material

Introductory Material: Working With R

Class 5 - Basic Data Exploration and Visualization in R HTML, MD, Rmd

Class 6 - Creating R Functions

Class 7 - R Packages, working with CRAN, and working with Bioconductor

Using R and Other Tools for Bioinformatics Analysis

Class 8 - An Introduction to Machine Learning (Heirarchical Clustering)

Class 9 - Analyzing High Dimensional Datasets and Unsupervised Learning

Class 11 - Structural Bioinformatics: Analyzing Protein Structure and Function

Class 12 - Drug Discovery: Techniques and Analysis

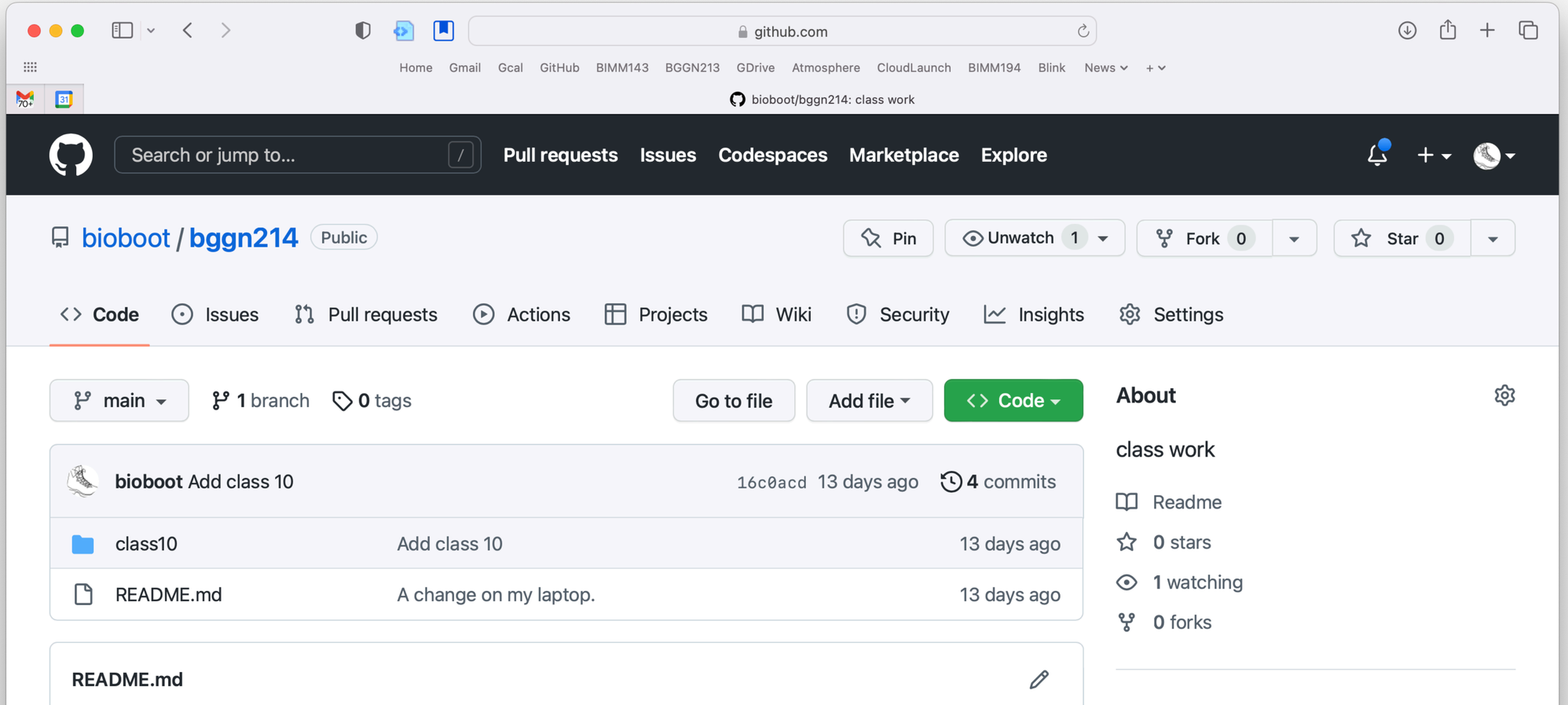
Class 13 - Genome Informatics and High Throughput Sequencing (NGS, RNA-Seq, and FastQC)

Class 14 - Transcriptomics and RNA-Seq Analysis

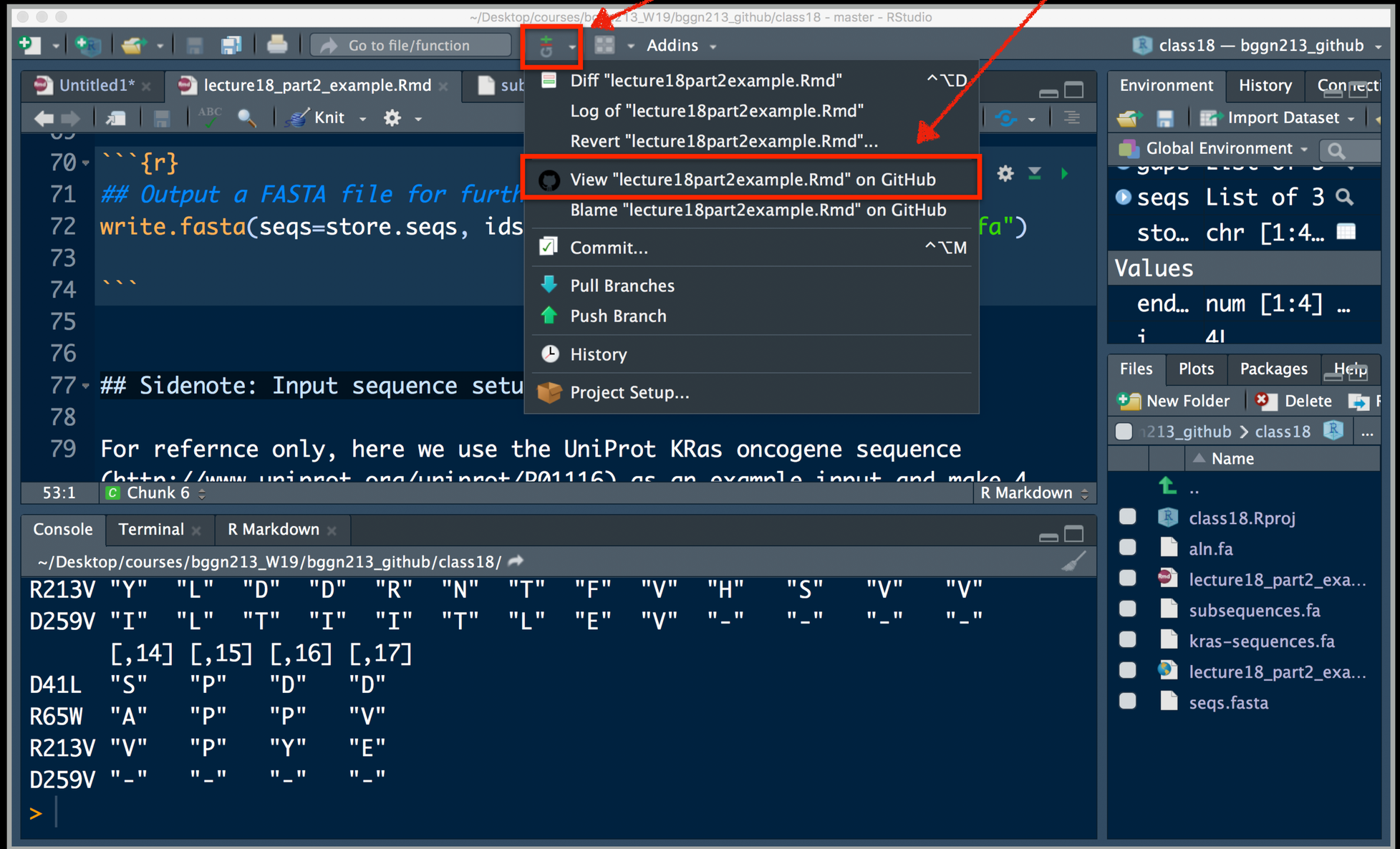
In your web browser navigate to your GitHub class repository < https://github.com/ >

Side-note:
To find the link to your GitHub repository from RStudio, open one of your past class projects and in the terminal type:

```
git remote -v
```



Or for a given GitHub tracked file click GIT icon and "VIEW on GITHUB"

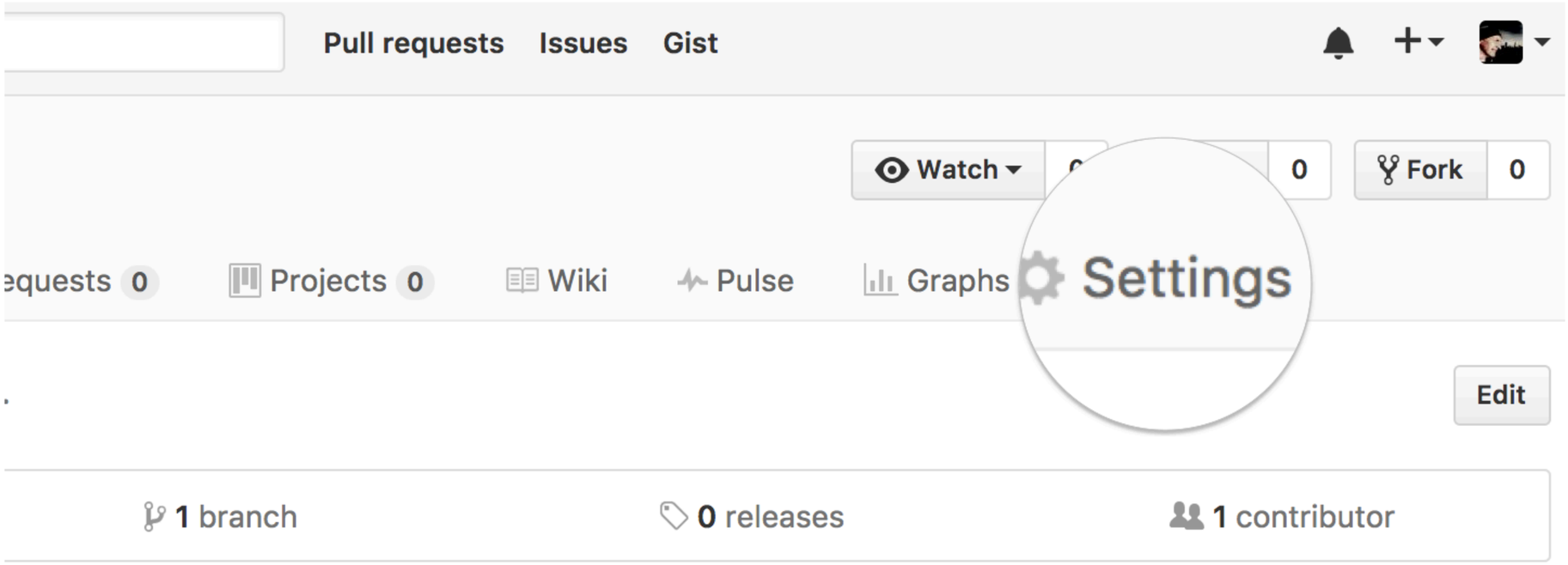


1

Repository Settings

Head over to your GitHub class repository (where you are “pushing” all your class work)

Click on the Settings tab.

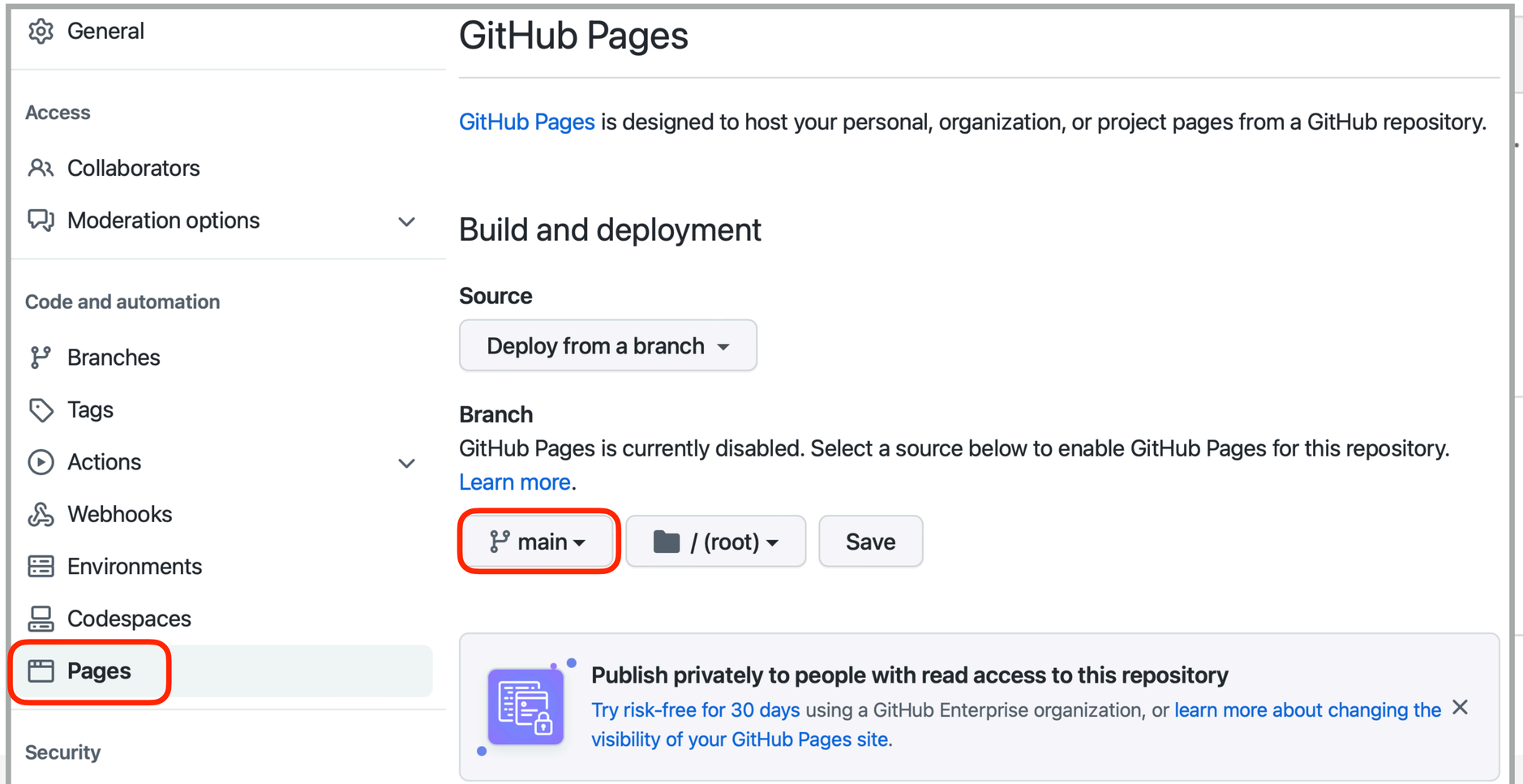


2

Theme chooser

Scroll down to the **GitHub Pages** section. ~~Press **Choose a theme**.~~

And set the Source to “main branch"

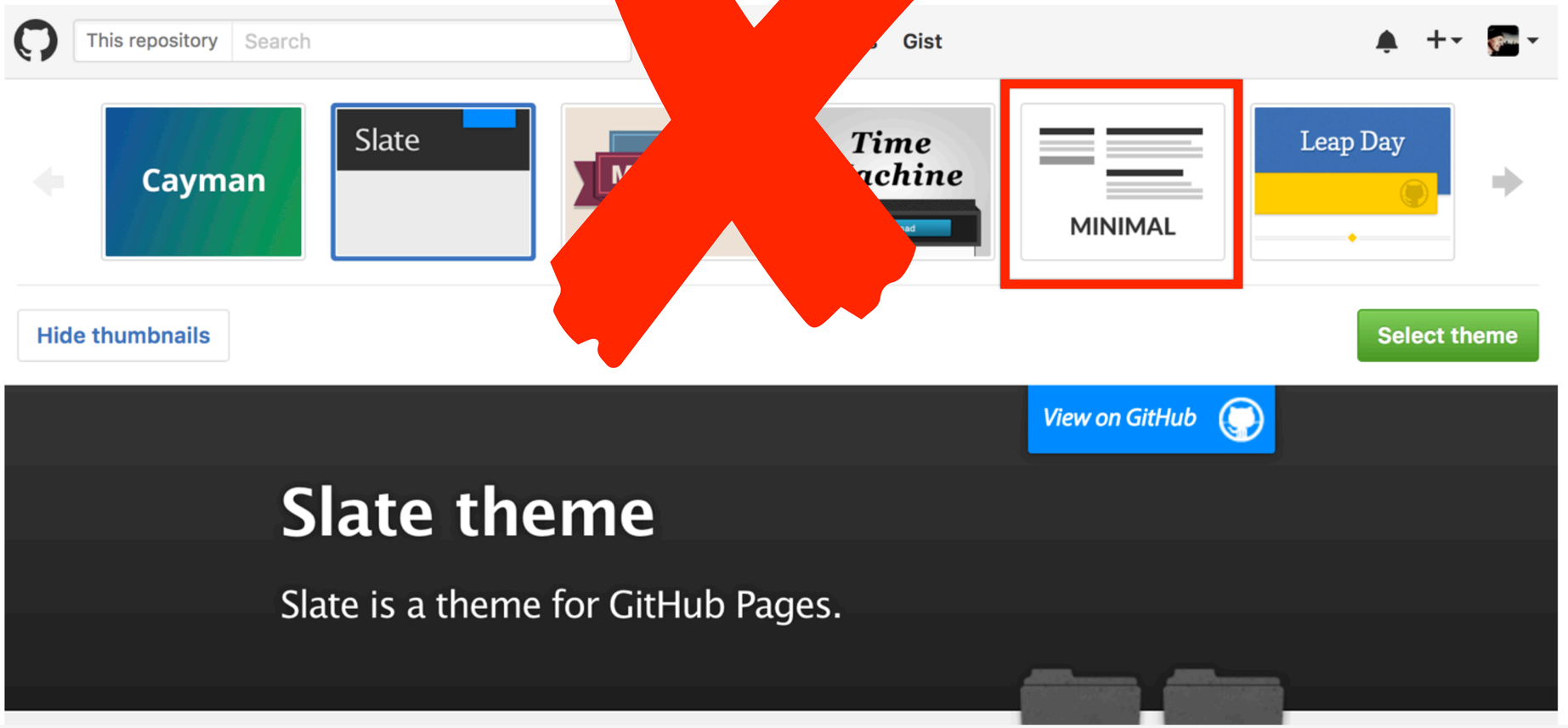


3

Pick a theme

Choose one of the themes from the carousel at the top.

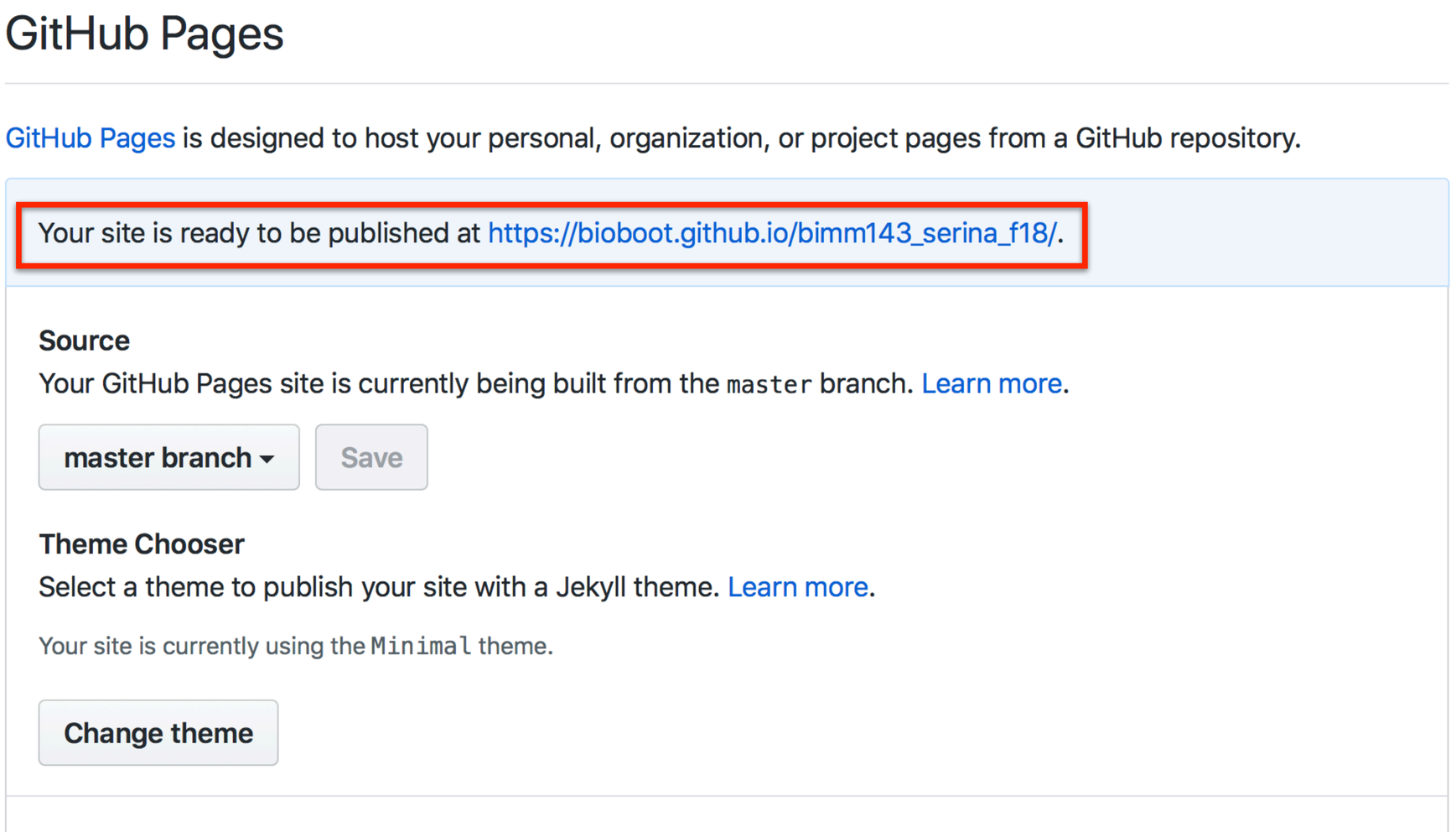
When you're done, click **Select theme** on the right.



4

Side-note:

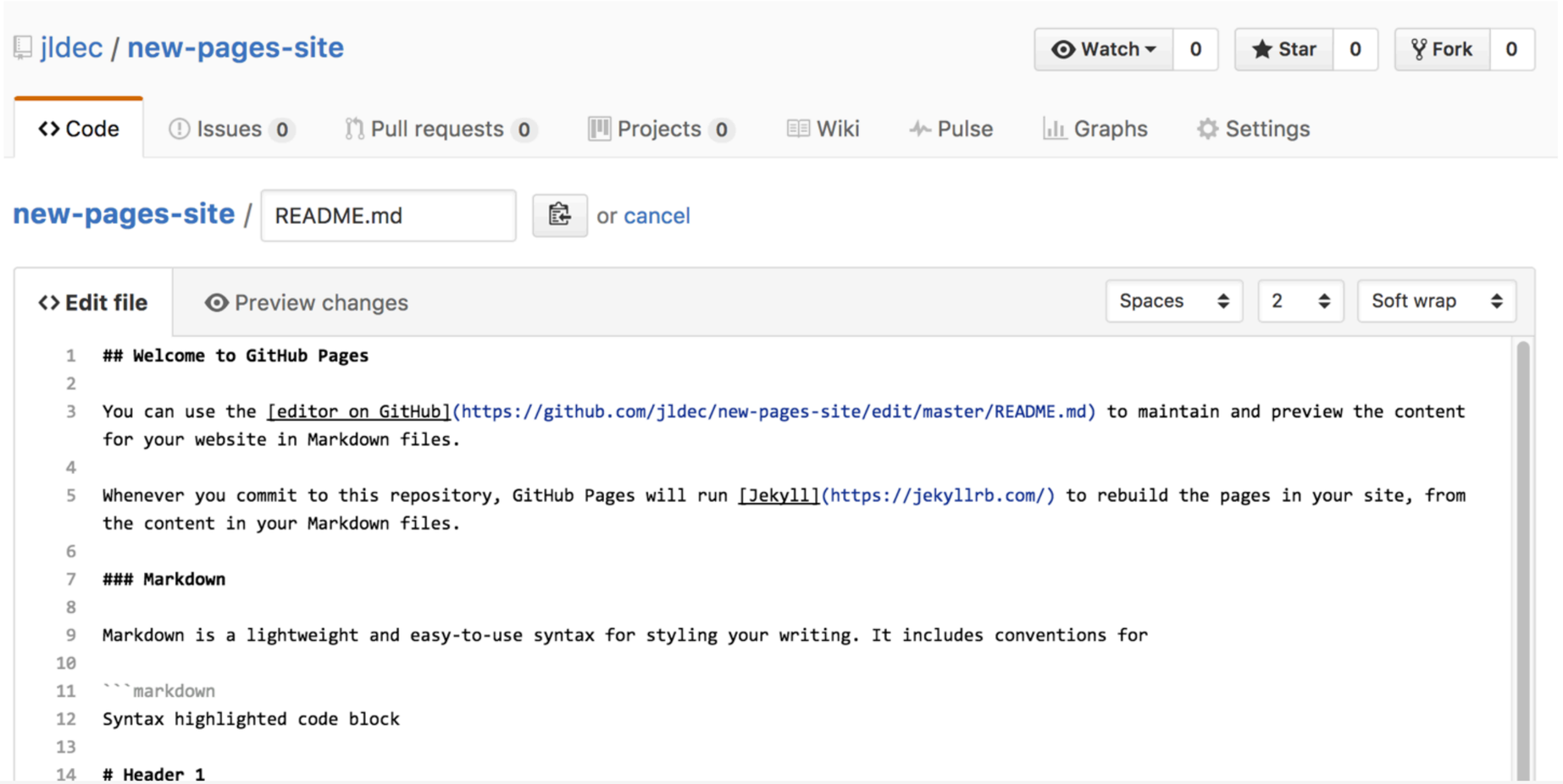
Scroll down again to the GitHub Pages section to find the link to your new website.
Open this link in a New Tab of your browser:



4

Edit content

Back on the repository main page use the GitHub online editor to add content. In particular, add links to each classes .MD file



5

Commit

Enter a commit comment and click on **Commit changes** below the editor.

```
35  ### Support or Contact
36
37  Having trouble with Pages? Check out our [documentation](https://help.github.com/categories/gi
    (https://github.com/contact) and we'll help you sort it out.
38
```

Commit changes

Add content to new pages site

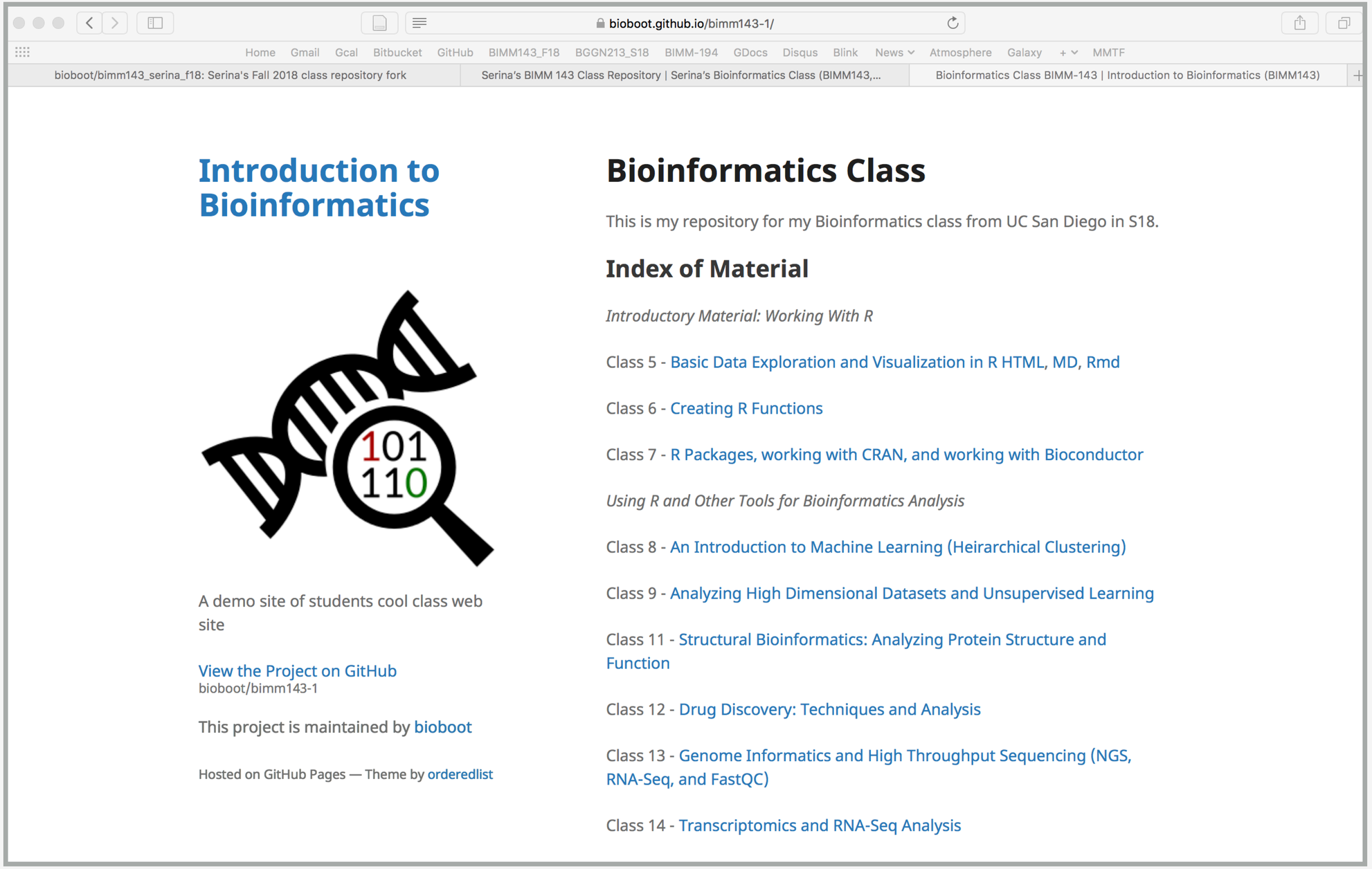
Add an optional extended description...

Commit directly to the `master` branch.

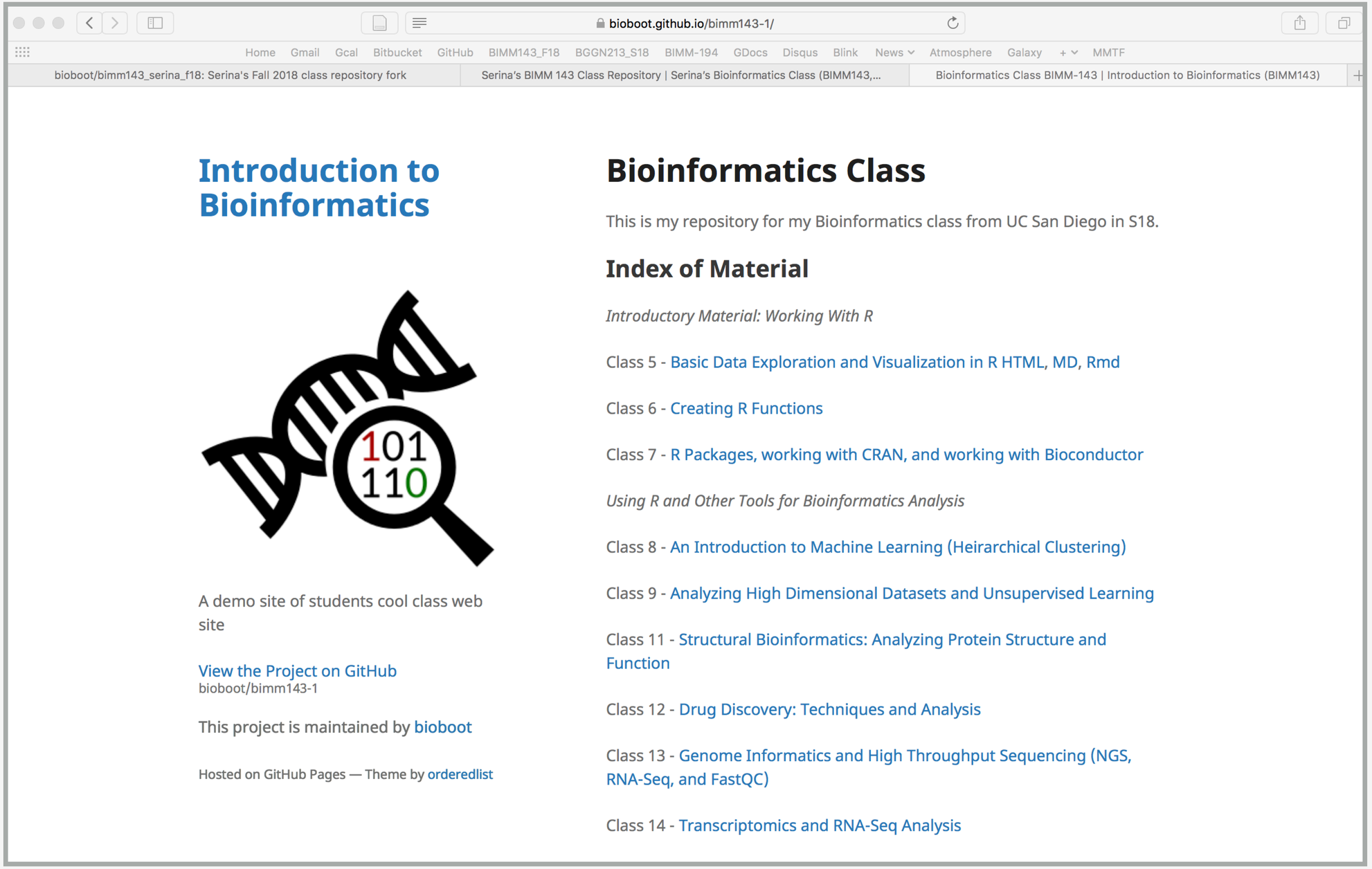
6

…and you're done!

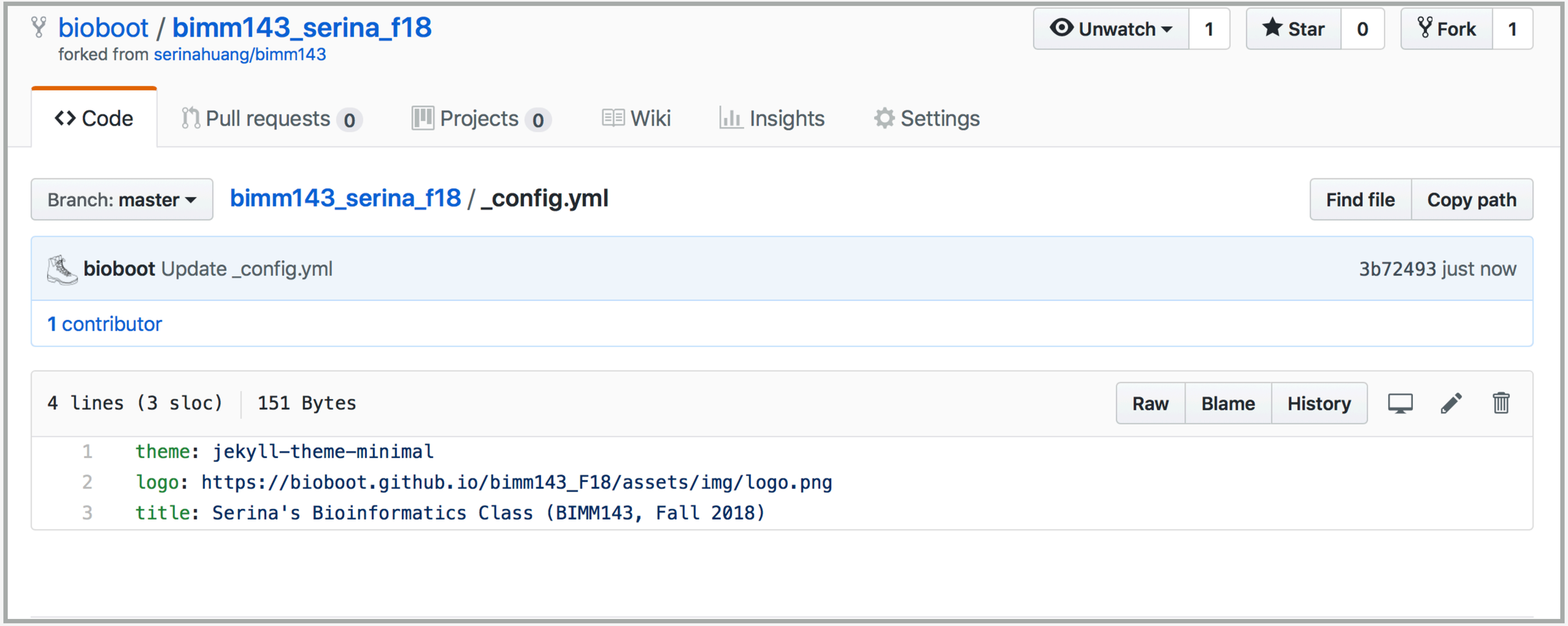
Fire up a browser and go to **http://*username*.github.io/*repository***.



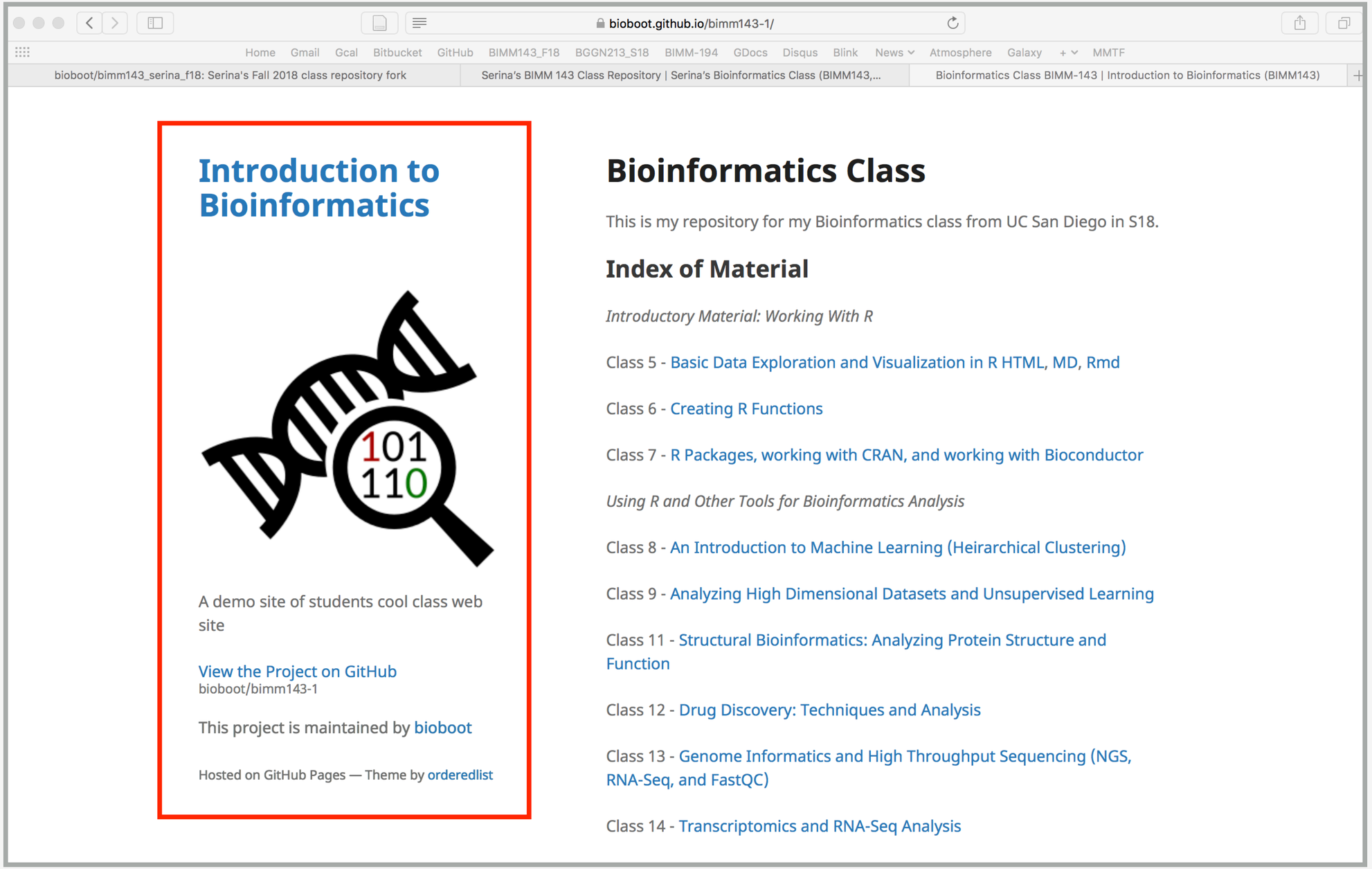
Here I: (1) forked Serina's Repo, (2) Chose the "minimal" theme, (3) Edited _config.yml (adding logo and title)



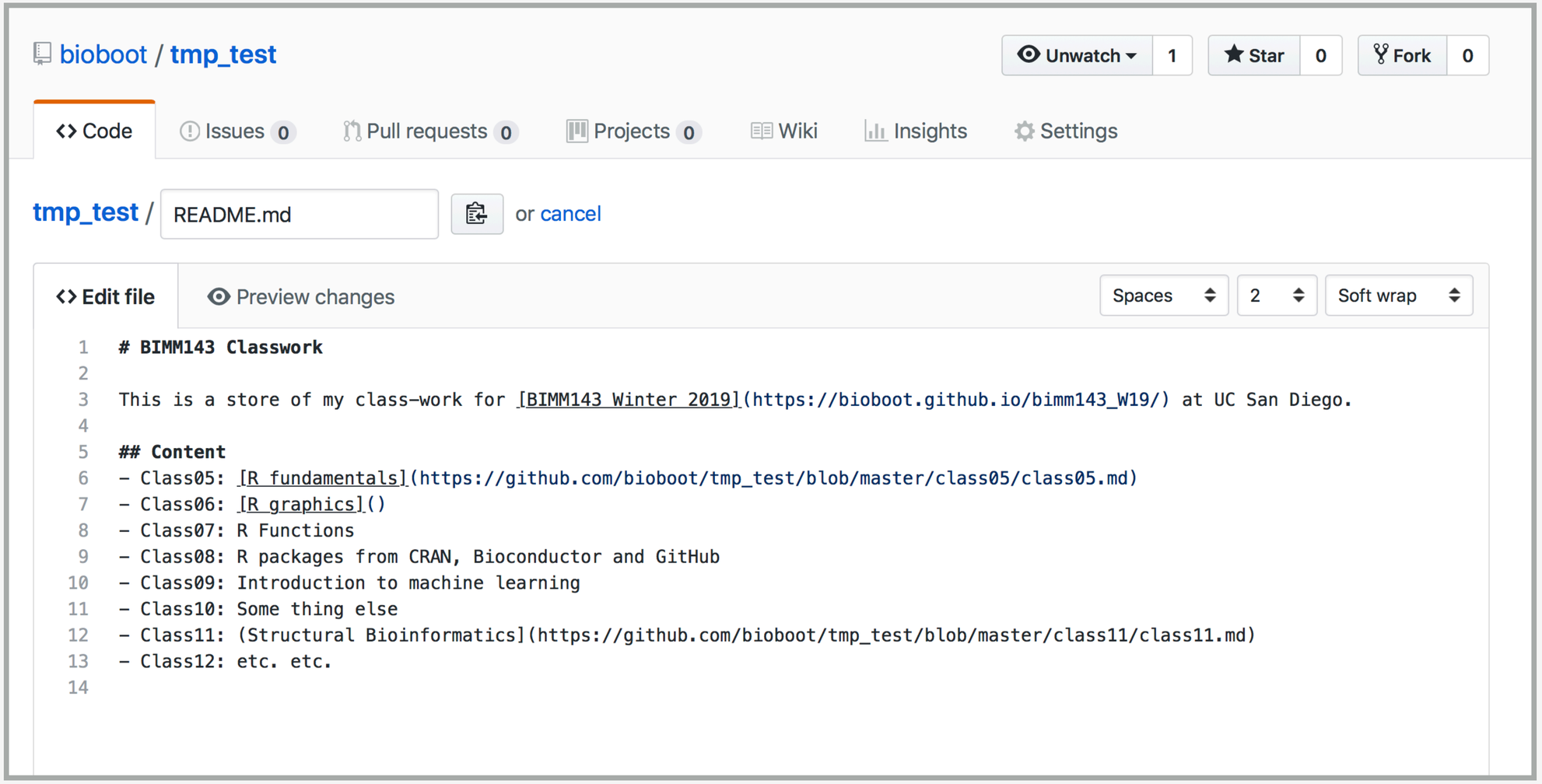
Here I: (1) forked Serina's Repo, (2) Chose the "minimal" theme, (3) Edited _config.yml (adding logo and title)



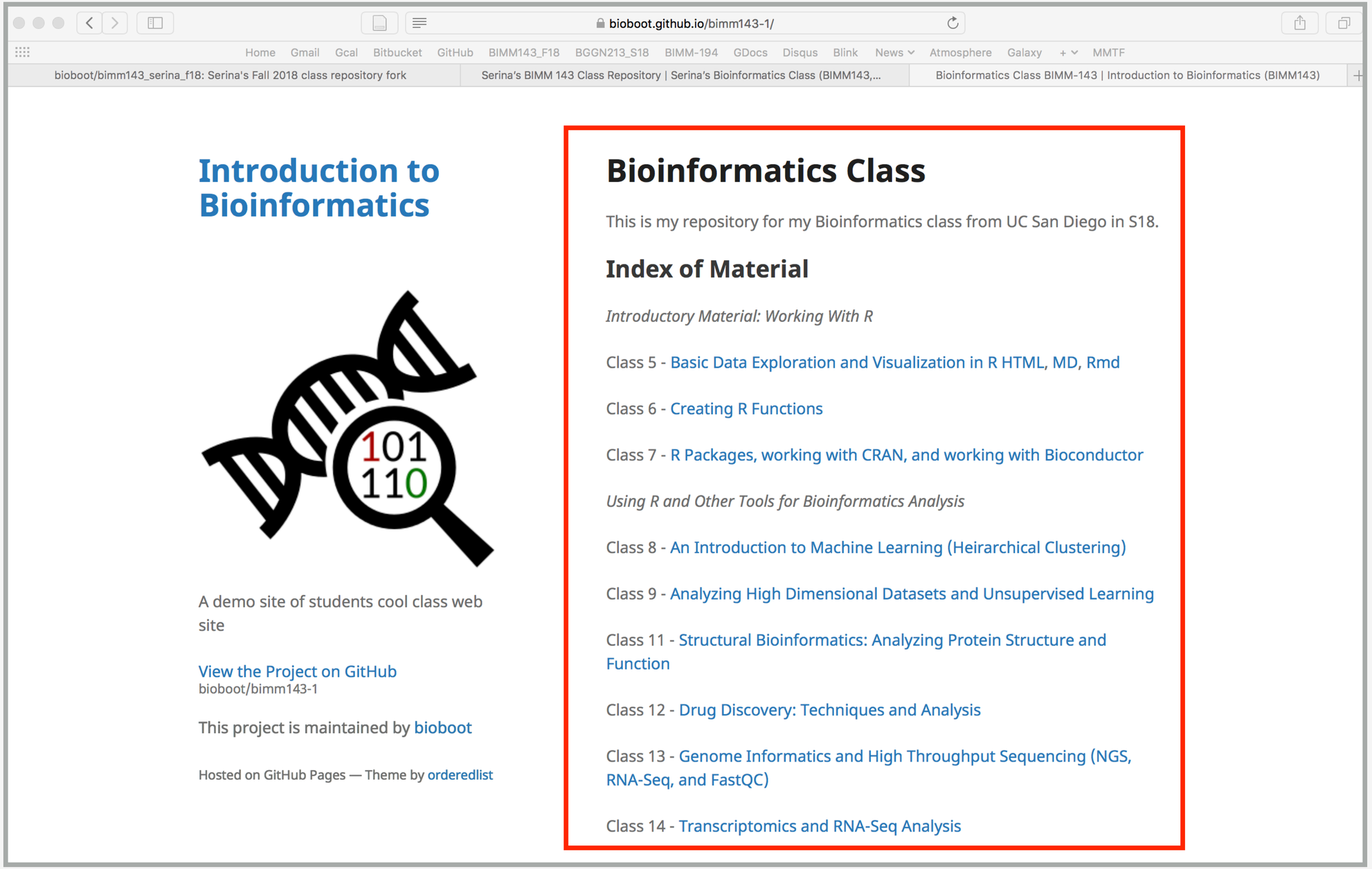
Here I: (1) Chose the "minimal" theme, (3) Edited _config.yml (adding logo and title), (4) Edited README.md



Here I: (1) Chose the "minimal" theme, (3) Edited _config.yml (adding logo and title), (4) Edited README.md



Here I: (1) Chose the "minimal" theme, (3) Edited _config.yml (adding logo and title), (4) Edited README.md



Here I: (1) Chose the "minimal" theme, (3) Edited _config.yml (adding logo and title), (4) Edited README.md

Thank you very much!

Please do fill out your CAPs evaluation (Link!) if you get a change.
It is important to the courses we offer in the future and how we teach them!

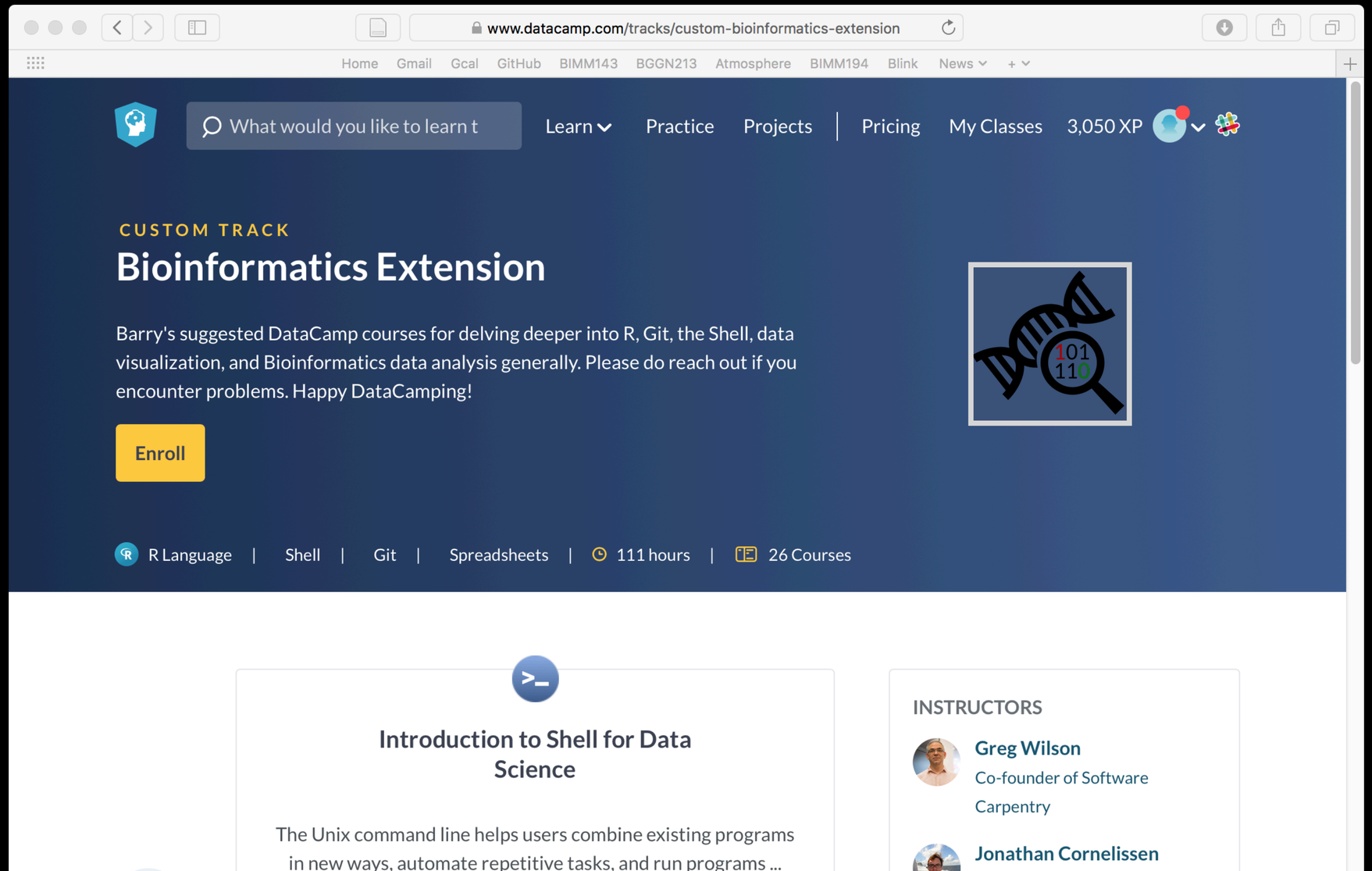
Thank you very much!

Please do fill out your CAPs evaluation (Link!) if you get a change.
It is important to the courses we offer in the future and how we teach them!

Post to GradeScope your GitHub Pages portfolio URL to GradeScope!

Going Further With DataCamp

https://bioboot.github.io/bimm143_F19/class-material/datacamp_extras.pdf



Announcement!

Bonus:
Bioinformatics & Genomics in Industry
Live Stream Video

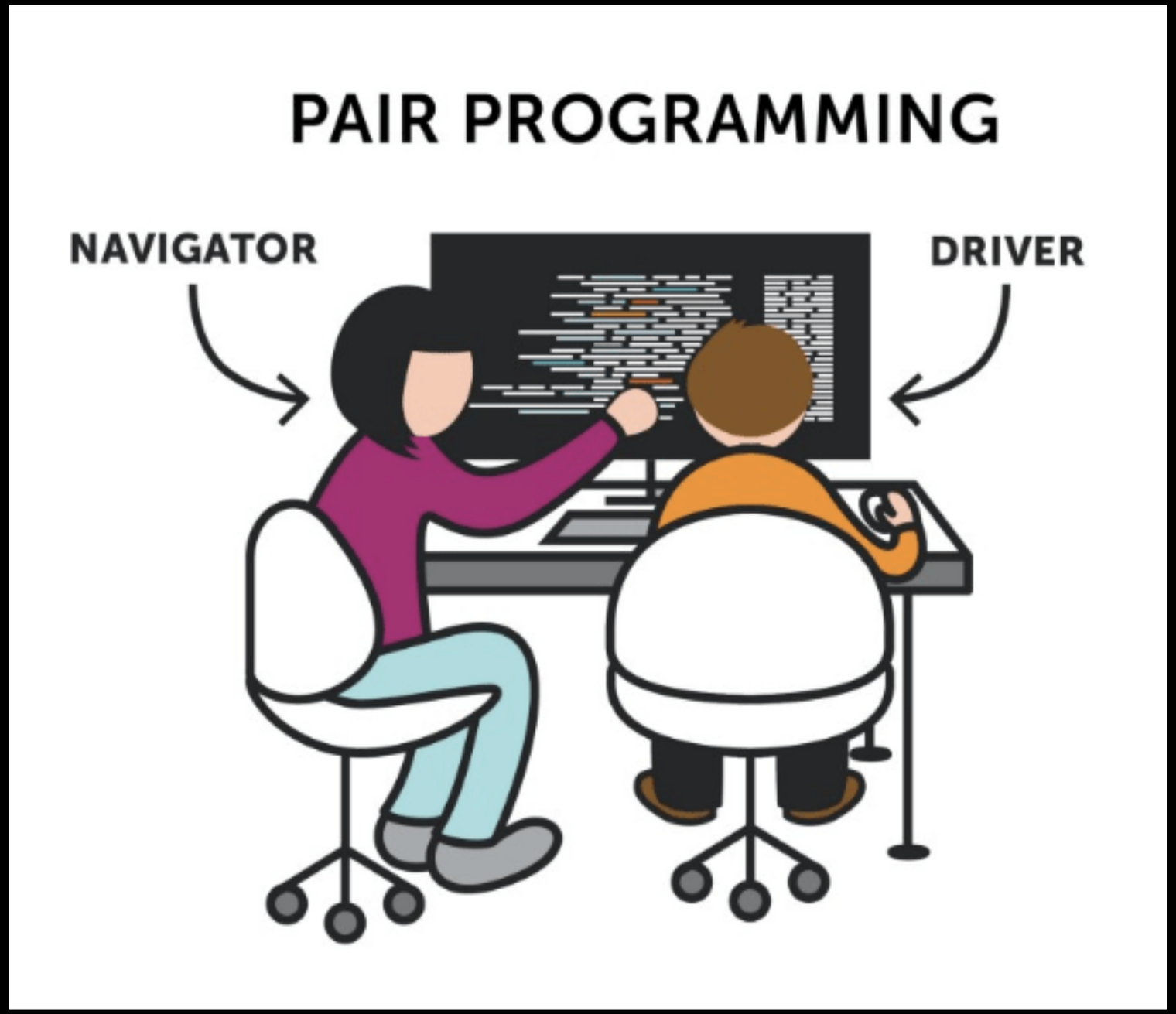
Enjoy a set of short open ended guest lectures from leading genomic scientists at **Illumina Inc.**, **Synthetic Genomics Inc.**, and the **La Jolla Institute for Allergy and Immunology**. Feel free to contact these scientists for networking and to have your questions about industry careers in Bioinformatics and Genomics answered.

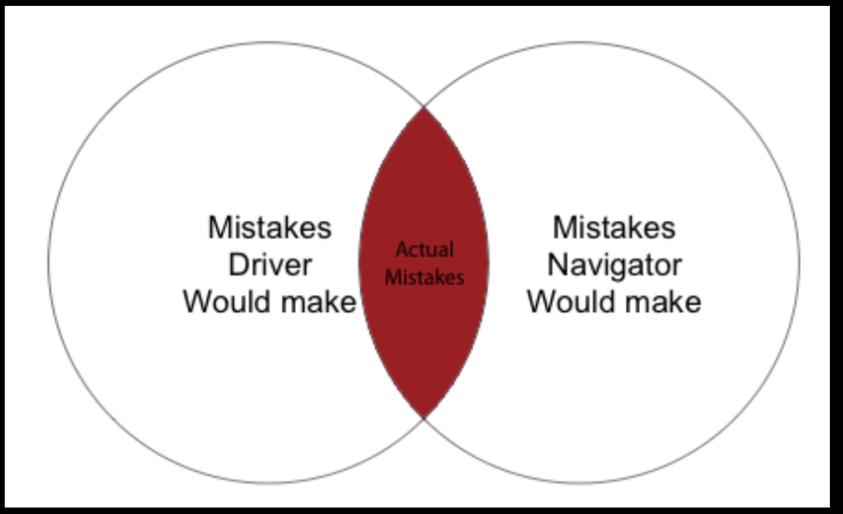
GitHub Copilot
Hands-on Lab Session
Barry Grant
UC San Diego
http://thegrantlab.org

GitHub Copilot

GitHub Copilot

Your AI “pair programmer”





Source: https://unruly.co/blog/article/2019/08/27/what-is-pair-programming/

GitHub Copilot

AI powered add-on that aims to give helpful suggestions when writing code or documentation

AI pair programmer

GitHub Copilot is an AI pair programmer that offers autocomplete-style suggestions and real-time hints for the code you are writing by providing suggestions as “ghost text” based on the context of the surrounding code

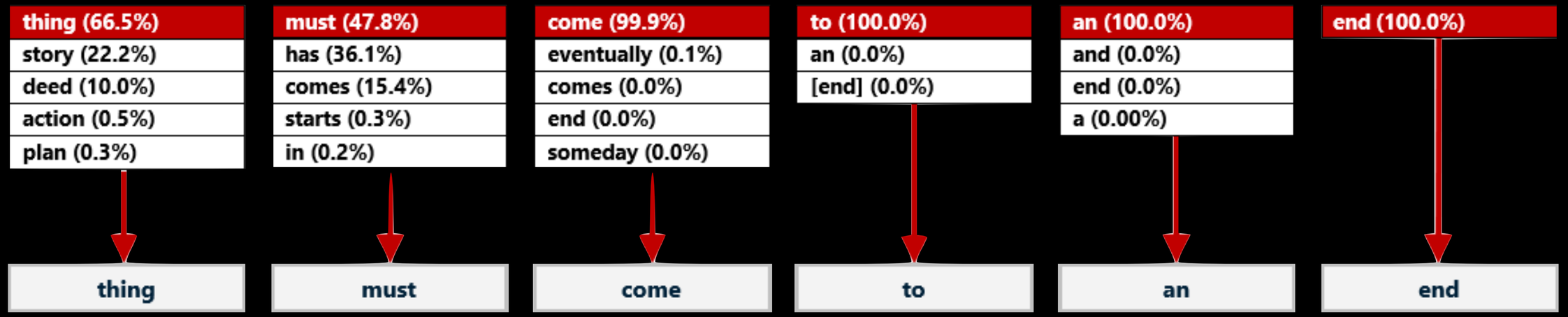
What can GitHub Copilot do?

- It provides autocomplete suggestions while you code.
- You can get suggestions by typing code or describing it in natural language (i.e. code comments or markdown text).
- Copilot analyzes your file (and related files), offering suggestions in your editor.
- It uses OpenAI Codex to derive context from your code and comments, and then suggests new lines or entire functions.

Generative AI Models

Generative AI

- For text generation, Generative AI just wants to predict the next word/ token/string!
- I might ask ChatGPT: > “Complete the sentence every good…”



OpenAI

Generative AI

GPT-3

Prompt:

```
Write a tagline for an ice cream shop.
```

Response:

```
We serve up smiles with every scoop!
```

Codex

Prompt:

```
Table customers, columns = [CustomerId, FirstName, LastName, Company, Address, City, State, Country, PostalCode]

Create a SQL query for all customers in Texas named Jane
query =
```

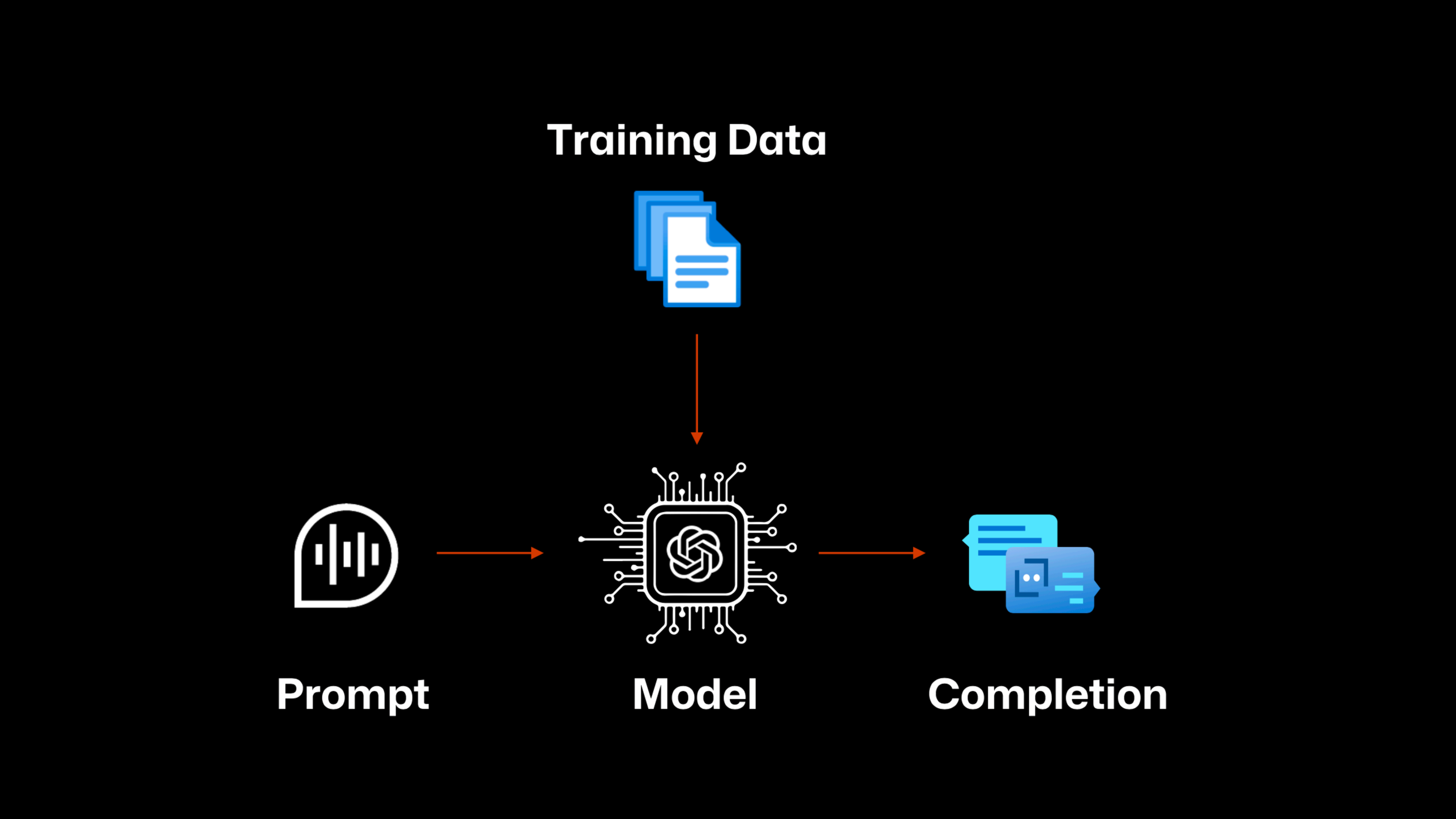
Response:

```
SELECT *
FROM customers
WHERE State = 'TX' AND FirstName = 'Jane'
```

DALL·E

Prompt: A white Siamese cat

Response:


Training Data
Prompt
Model
Completion

Generative AI can:

Generate text, images and code

Different models are trained on different corpuses, depending on the application.

Generate “humanlike” output

What is a likely continuation of the prompt, given the training data?

Extract information

The continuation is likely to be similar to text frequently represented in the training data.

Create novel content

Text, images and code not contained in its training set. Translations. “Creative” works.

Generative AI is not:

Intelligent

It's just a predictive system, designed to give a likely continuation of the prompt given the training data.

Deterministic

Run the same prompt. Get back a different response (probably).

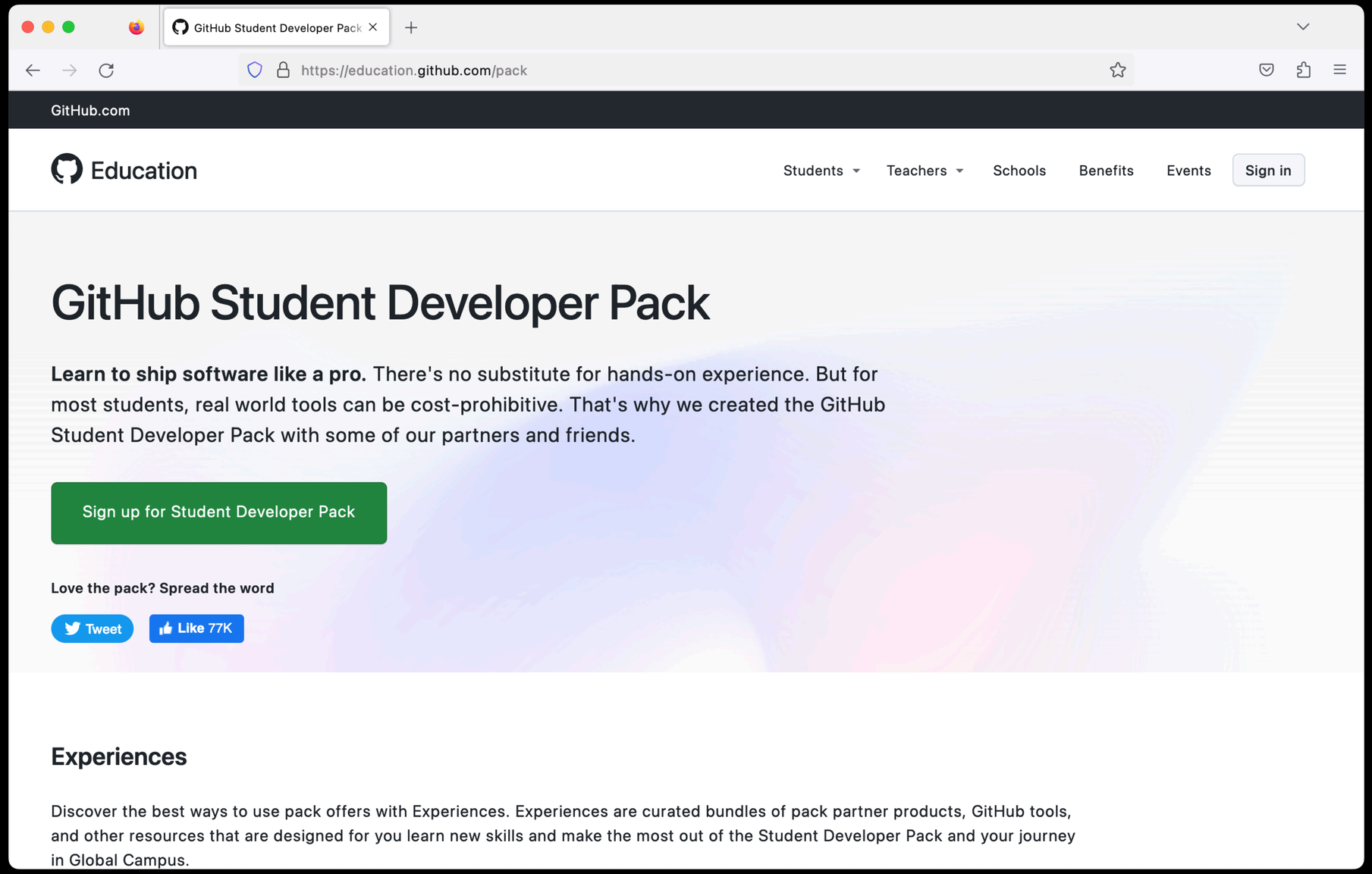
Trustworthy

It can "hallucinate" facts and confidently assert them to be true.

TODO

Generative AI does not:

- **Learn**
  The model is fixed at the time of its training.
- **Contain all of the information of its training set**
  Think: a blurry jpeg of its training data.
- **Include verbatim copies of its training data**
  But it can generate stuff that looks like it.


GitHub Student Developer Pack
https://education.github.com/pack
GitHub.com
Education
Students
Teachers
Schools
Benefits
Events
Sign in
GitHub Student Developer Pack
Learn to ship software like a pro. There's no substitute for hands-on experience. But for most students, real world tools can be cost-prohibitive. That's why we created the GitHub Student Developer Pack with some of our partners and friends.
Sign up for Student Developer Pack
Love the pack? Spread the word
Tweet
Like 77K
Experiences
Discover the best ways to use pack offers with Experiences. Experiences are curated bundles of pack partner products, GitHub tools, and other resources that are designed for you learn new skills and make the most out of the Student Developer Pack and your journey in Global Campus.

GitHub Student Developer Pack

https://education.github.com/pack

Tools

Security and Monitoring

Explore industry standard tools that keep your code safe, your permissions secure and your infrastructure monitored so that you can confidently scale your brilliant ideas.

Offers in this bundle

GitHub · Blackfire · Honeybadger · AstraSecurity · +3

Learn more about Security and Monitoring >

Primer

Primer: Copilot

What is Copilot and how do you get started using it? Use this Experience as a Primer on Copilot. We share tutorials, guides and templates to learn how to use AI in your development process.

Offers in this bundle

Github Copilot · GitHub Codespaces · Visual Studio Code

Learn more about Primer: Copilot >

Tools

Career Readiness

Whether you're approaching graduation or just want to get a head start, it's never too early to prepare for your future in the industry.

Offers in this bundle

Educative · InterviewCake · DailyBot · GitHub Pages · +2

Learn more about Career Readiness >

Tools

Intro to Web Dev

Everything you need to build your next website.

Tools

Hackathon in the Cloud

Learn how to organize, promote, and communicate

Tools

Aspiring Creatives

Working on a creative project? Develop your design

Request a discount - GitHub Ed | Virtual Event Kit - GitHub Educa | Intro to GitHub Flow - GitHub Ed | Data Science & Machine Learnin | Primer: Copilot - GitHub Educati

https://education.github.com/discount_requests/application

GitHub.com

Education

Students Teachers Schools Benefits Events Global Campus

Home / Benefits / Benefits application

Get your GitHub benefits

Learn and teach using real-world developer tools

Select your academic status *

Teacher Student

Benefits for Students

STUDENT

FREE GitHub Pro while you are a student

STUDENT

Valuable GitHub Student Developer Pack partner offers

STUDENT

GitHub Campus Expert training for qualified applicants

To qualify for student benefits, you must:

- Have a GitHub account.
- Be at least 13 years old.

Quick Start Guide

https://docs.github.com/en/copilot/quickstart

Additional Learning Resources

- **Video series from GitHub:**
  - https://learn.microsoft.com/en-us/shows/introduction-to-github-copilot/
- **Blog post series:**
  - https://github.blog/2023-06-20-how-to-write-better-prompts-for-github-copilot/
- **RStudio Docs:**
  - https://docs.posit.co/ide/user/ide/guide/tools/copilot.html

The End!
http://thegrantlab.org

Thank You!
http://thegrantlab.org