Software Archaeology and Anthropology

17-313 Fall 2025

Foundations of Software Engineering

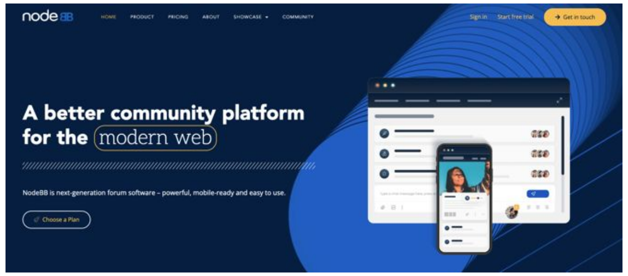
https://cmu-17313q.github.io

Eduardo Feo Flushing

Administrivia

- Slack
  - Please add a profile picture.
  - Ask questions in `#general` or `#technical-questions`.
    - Please use threads.
    - Use the search tool.
- Office hours can be found on the course home page: http://cmu-17313q.github.io

Project P1


A better community platform for the modern web

NodeBB is next-generation forum software – powerful, mobile-ready and easy to use.

- **P1A**: Checkpoint due next Sunday (August 31st)
  - Only 5% of total P1 points – meant to ensure you start on time
- **P1B**: Due Sunday, September 7th
  - Refactor a javascript file to improve its quality
  - It will be posted tonight
  - Start early

Archaeology vs Anthropology

Artifacts

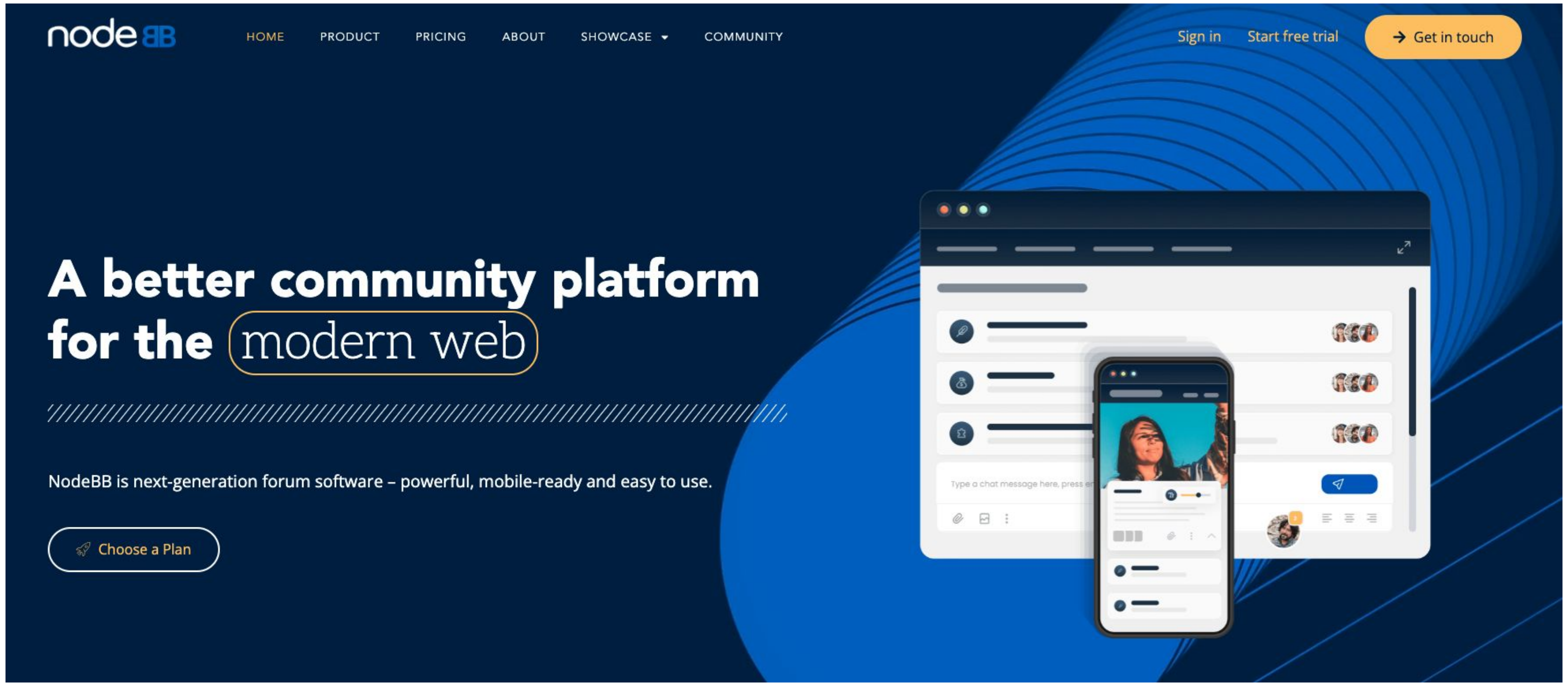
Human societies and culture

Learning Goals

- Understand and scope the task of working with a new and complex piece of existing software
- Appreciate the importance of configuring an effective IDE
- Contrast different types of code execution environments: local, remote, application, and library-based
- Enumerate both static and dynamic strategies for understanding and modifying a new codebase

Context: big old pile of code

- … do something with it!



A better community platform for the modern web

NodeBB is next-generation forum software – powerful, mobile-ready and easy to use.

You will never understand the entire system!

Challenge: How do I tackle this codebase?

WHY IS THIS STRUCTURE HERE?
WHERE COULD THIS BRIDGE POSSIBLY LEAD?
THIS SIGN DOESN'T HELP ME MUCH.
//THIS IS BRIDGE

Challenge: How do I tackle this codebase?

- Leverage your previous experiences (languages, technologies, patterns)
- Consult documentation, whitepapers
- Talk to experts, code owners
- Follow best practices to build a working model of the system

Bad news: There are few helpful resources!

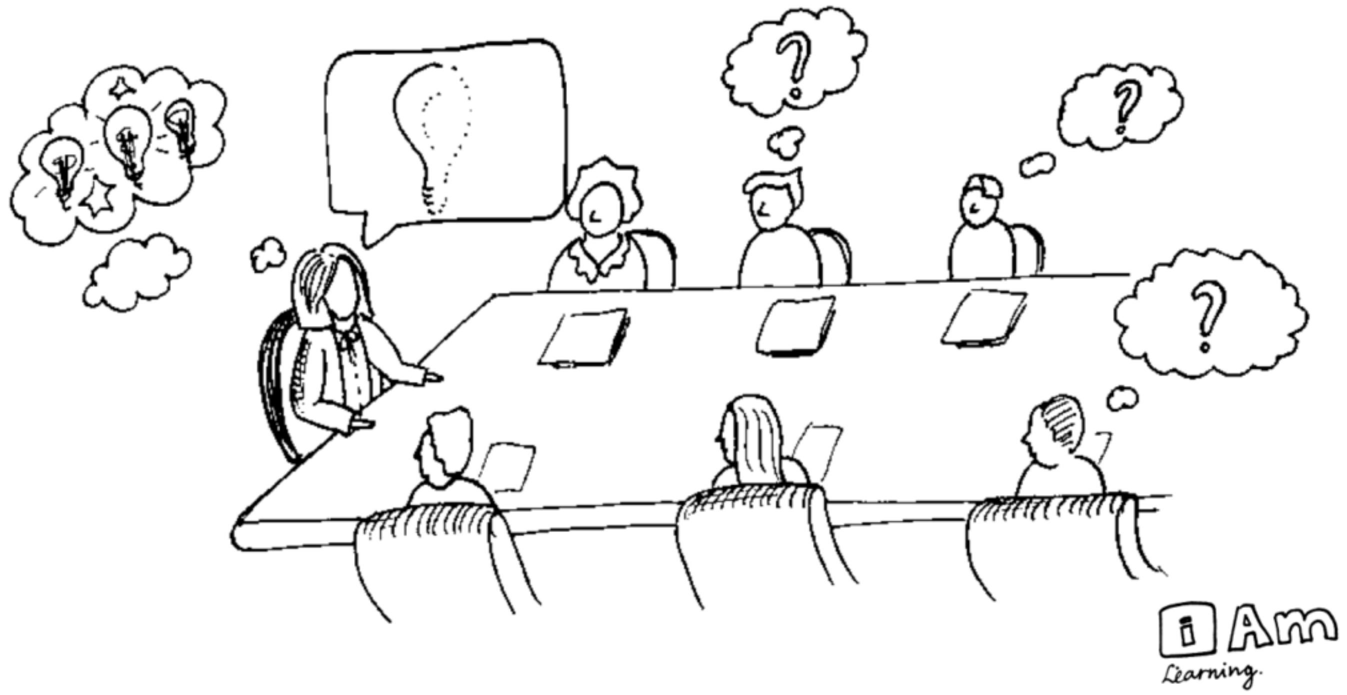
- **Working Effectively with Legacy Code.**
  Michael C. Feathers (2004)
- **Re-Engineering Legacy Software.**
  Chris Birchall (2016)
- **The Legacy Code Programmer's Toolbox.**
  Jonathan Boccara (2019)

START INVESTING
Wealthsimple Magazine
MENU
MONEY & THE WORLD
Robert C. Martin Series
WORKING EFFECTIVELY WITH LEGACY CODE
Michael C. Feathers
Re-Engineering LEGACY SOFTWARE
Chris Birchall
THE LEGACY CODE PROGRAMMER'S TOOLBOX
PRACTICAL SKILLS FOR DEVELOPERS WORKING WITH LEGACY CODE
JONATHAN BOCCARA
FOREWORD BY KEVLIN HENNEY

The Code That Controls Your Money

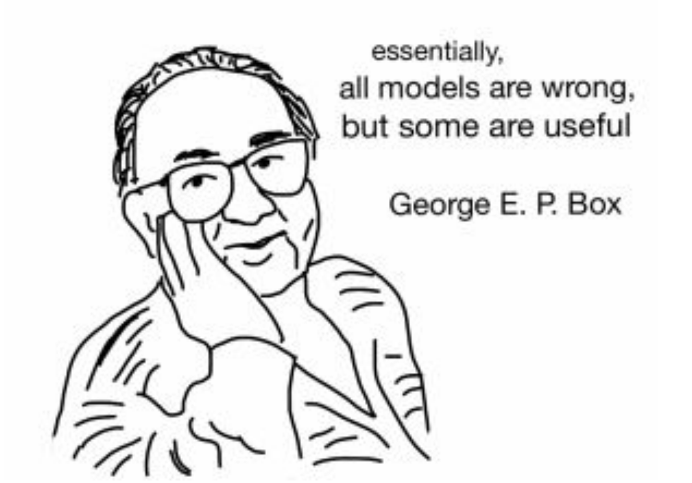
COBOL is a coding language older than Weird Al Yankovic. The people who know how to use it are often just as old. It underpins the entire financial system. And it can't be removed. How a computer language controls the financial life of the world.

Why? Because of Tacit Knowledge

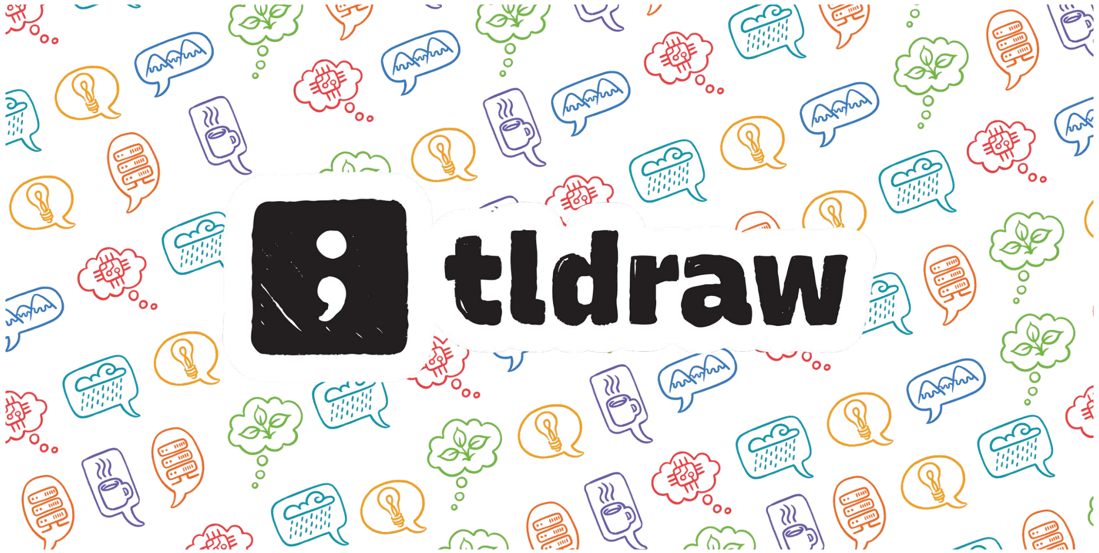


Today: How to tackle codebases

- Goal: Develop and test a working model (or set of hypotheses) about how part of a system works
- Working model: an understanding of the pieces of the system (components), and the way they interact (connections)
- Focus: Observation, probes, and hypothesis testing
  - Helpful tools and techniques!



Live Demonstration: tldraw



https://github.com/tldraw/tldraw

Academic Honesty

- Standard Collaboration Policy
- In group work, be honest about contribution of group members; do not cover for others
- Unless explicitly prohibited, you may use generative AI (e.g. ChatGPT) to help you write your prose and code. You are responsible for its correctness. Be sure to attribute the content to the service you used. (let us know if you have concerns about teammate's work)
- DO NOT submit participation sheets for people who are not in class. This is considered an academic integrity violation

Steps to Understand a New Codebase

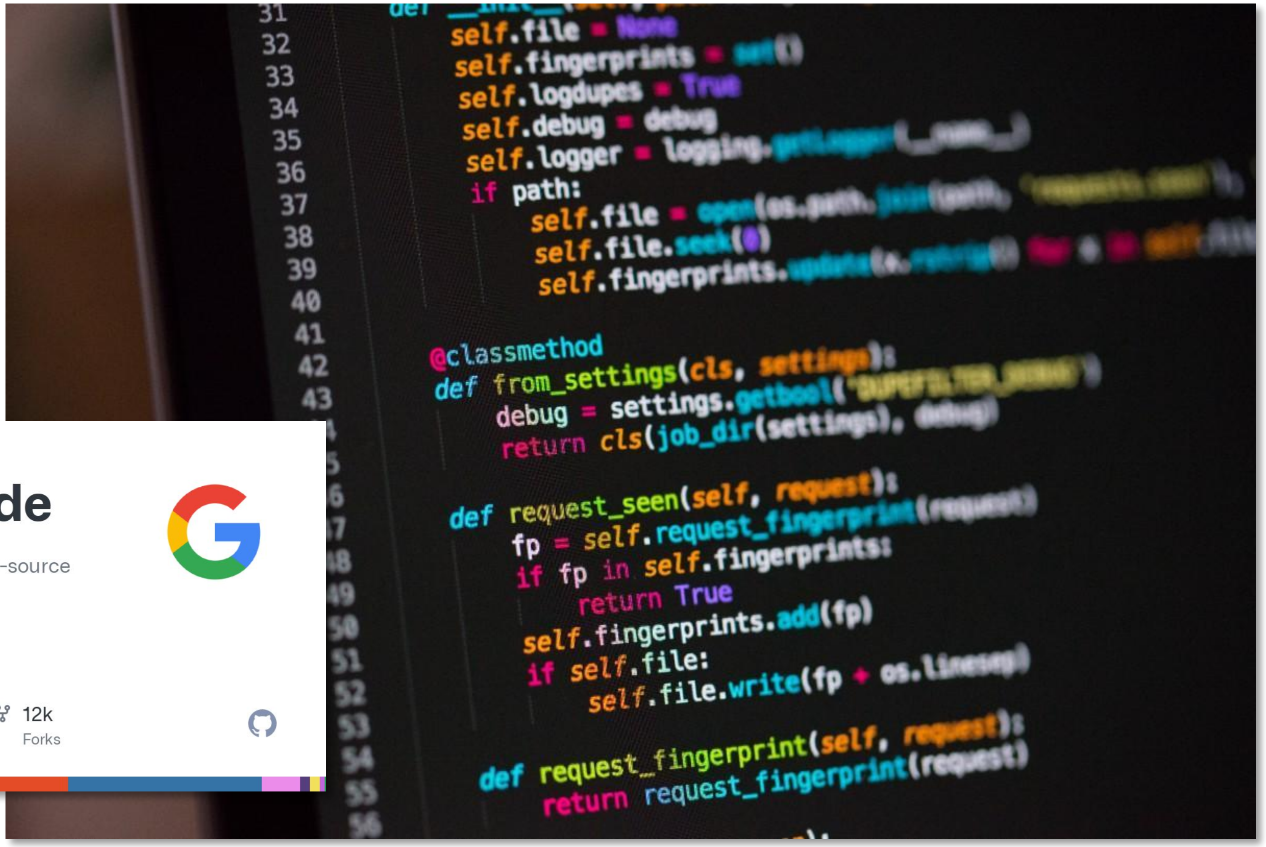
- Look at README.md
- Clone the repo.
- Build the codebase.
- Figure out how to make it run.
- What do you want to mess with?
    - Clone and own
- Traceability - Attach a debugger
    - View Source
    - Find the logs.
    - Search for constants (strings, colors, weird integers (#DEADBEEF))

Participation Activity

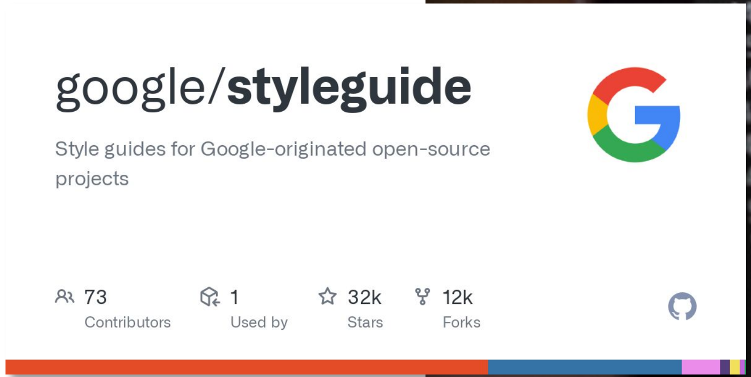
- Take out a piece of paper.
- Write down one pro and one con about trying to understand a new codebase by compiling and building it vs. just reading the code.
- Pair with your neighbor and discuss your answers. Do you agree?
- Share with the class!
- Write your own andrewID on the paper, leave it at the end of class.

Observation: Software is full of patterns

- File structure
- System architecture
- Code structure
- Names
- …



```
        self.file = None
        self.fingerprints = set()
        self.logdupes = True
        self.debug = debug
        self.logger = logging.getLogger(__name__)
        if path:
            self.file = open(os.path.join(path, 'requests.seen'), 'a+')
            self.file.seek(0)
            self.fingerprints.update(x.rstrip() for x in self.file)

    @classmethod
    def from_settings(cls, settings):
        debug = settings.getbool('DUPEFILTER_DEBUG')
        return cls(job_dir(settings), debug)

    def request_seen(self, request):
        fp = self.request_fingerprint(request)
        if fp in self.fingerprints:
            return True
        self.fingerprints.add(fp)
        if self.file:
            self.file.write(fp + os.linesep)

    def request_fingerprint(self, request):
        return request_fingerprint(request)
```



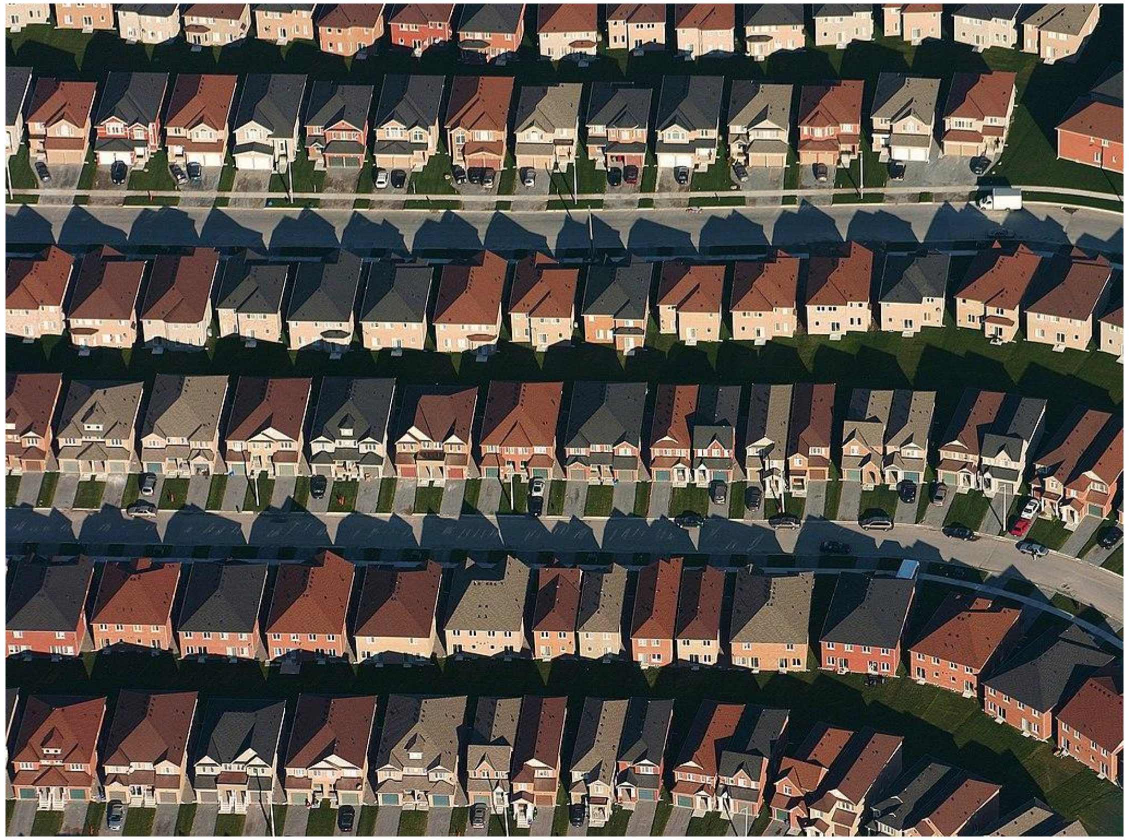
google/**styleguide**

Style guides for Google-originated open-source projects

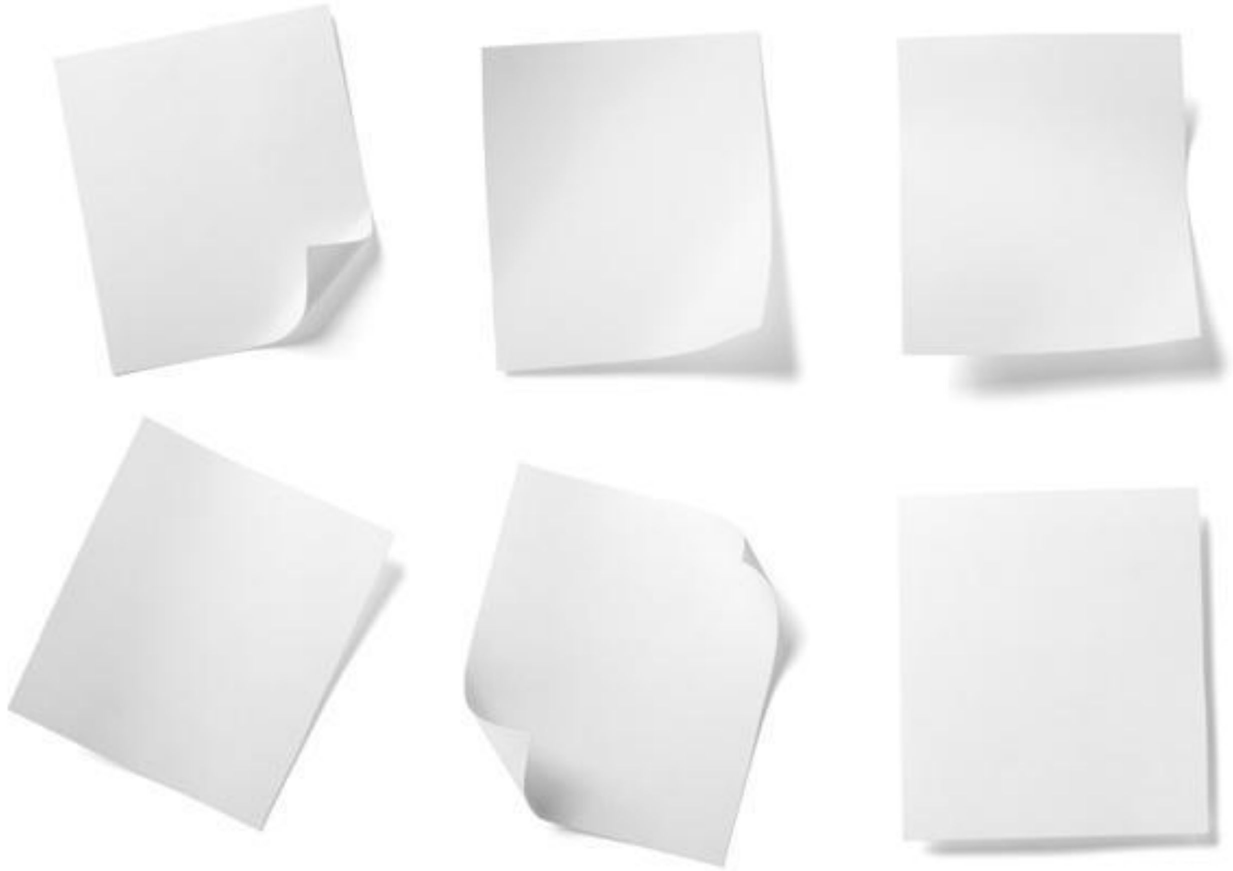
73 Contributors · 1 Used by · 32k Stars · 12k Forks

Observation: Software is massively redundant

- There’s always something to copy/use as a starting point!



Observation: Code must run to do stuff!



Observation: If code runs, it must have a beginning…

The Beginning: Entry Points

- Locally installed programs: run cmd, OS launch, I/O events, etc.

- Web apps server-side: Browser sends HTTP request (GET/POST)

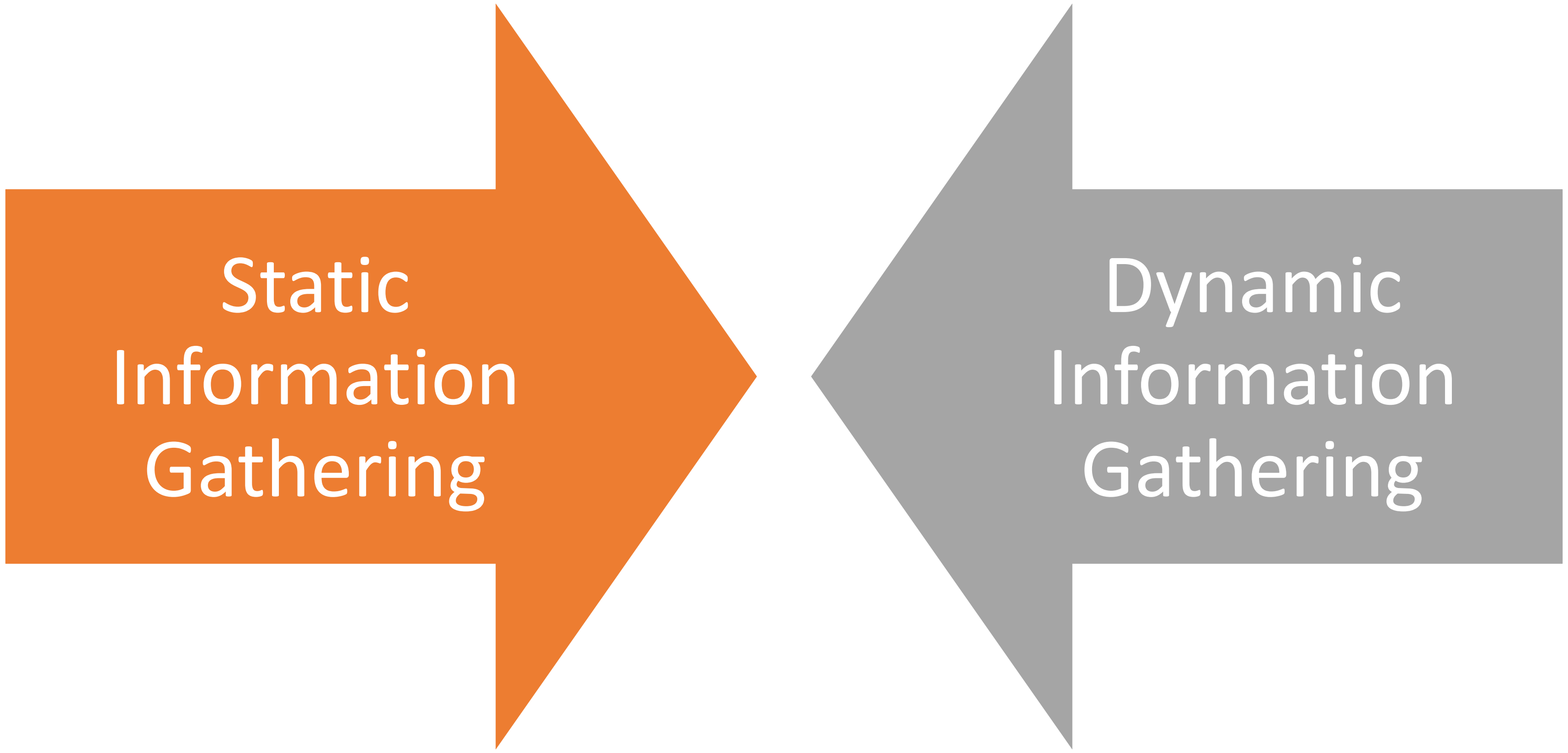
- Web apps client-side: Browser runs JavaScript, event handlers

Can running code be Probed/Understood/Edited?

Transparent	Translucent		Opaque	
Source code built locally	Binaries running locally		Server-side apps running remotely	
	Open source	Closed source	Open source	Closed source
(P+U+E)	(P+U)	(P)	(U)	(Talk to NSA)

Creating a model of unfamiliar code

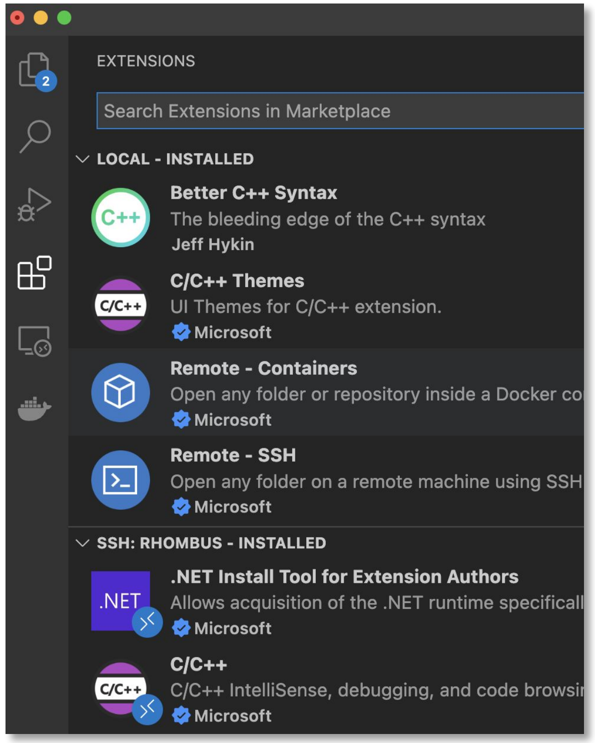
Source code built locally


Static Information Gathering
Dynamic Information Gathering

Static Information Gathering

- Basic needs:
  - Code/file search and navigation
  - Code editing (probes)
  - Execution of code, tests
  - Observation of output (observation)
- Many choices here on tools! Depends on circumstance.
  - grep/find/etc.   Knowing Unix tools is invaluable
  - A decent IDE
  - Debugger
  - Test frameworks + coverage reports
  - Google (or your favorite web search engine)
  - ChatGPT or LLaMA

Static Information Gathering: Use an IDE!
Real software is too complex to keep in your head



EXTENSIONS

Search Extensions in Marketplace

LOCAL - INSTALLED

Better C++ Syntax
The bleeding edge of the C++ syntax
Jeff Hykin

C/C++ Themes
UI Themes for C/C++ extension.
Microsoft

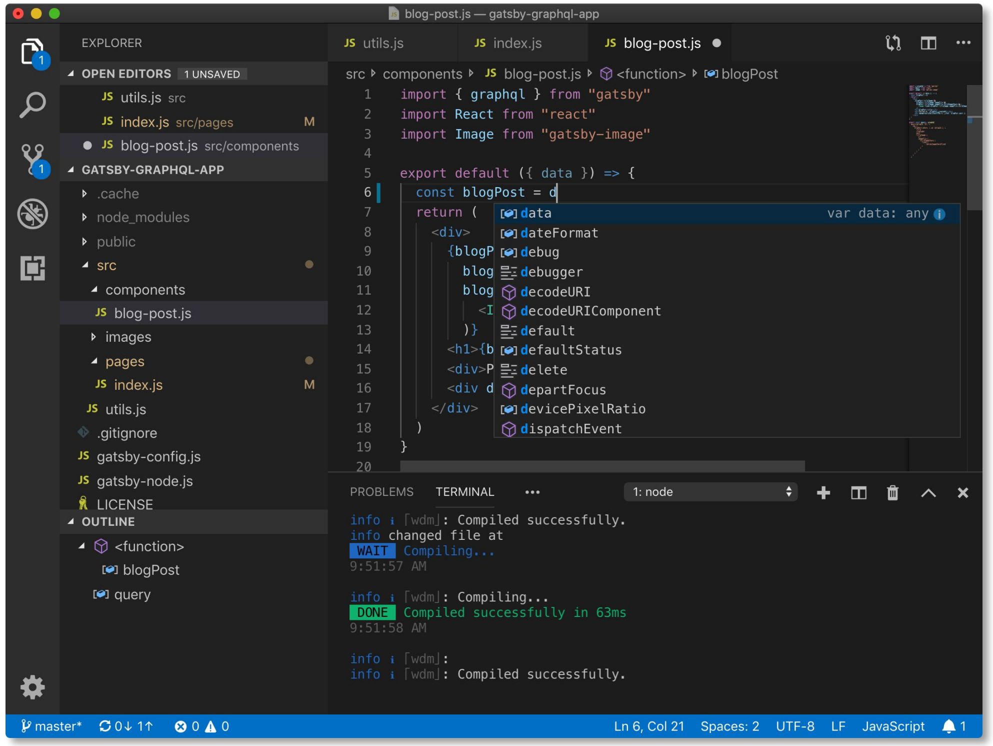
Remote - Containers
Open any folder or repository inside a Docker co
Microsoft

Remote - SSH
Open any folder on a remote machine using SSH
Microsoft

SSH: RHOMBUS - INSTALLED

.NET Install Tool for Extension Authors
Allows acquisition of the .NET runtime specificall
Microsoft

C/C++
C/C++ IntelliSense, debugging, and code browsi
Microsoft



blog-post.js — gatsby-graphql-app

EXPLORER

OPEN EDITORS 1 UNSAVED
- utils.js src
- index.js src/pages M
- blog-post.js src/components

GATSBY-GRAPHQL-APP
- .cache
- node_modules
- public
- src
  - components
    - blog-post.js
  - images
  - pages
    - index.js M
  - utils.js
- .gitignore
- gatsby-config.js
- gatsby-node.js
- LICENSE

OUTLINE
- <function>
  - blogPost
  - query

utils.js | index.js | blog-post.js ●

src › components › blog-post.js › <function> › blogPost

```
import { graphql } from "gatsby"
import React from "react"
import Image from "gatsby-image"

export default ({ data }) => {
  const blogPost = d
  return (
    <div>
      {blogP
        blog
        blog
          <I
      )}
      <h1>{b
      <div>P
      <div d
    </div>
  )
}
```

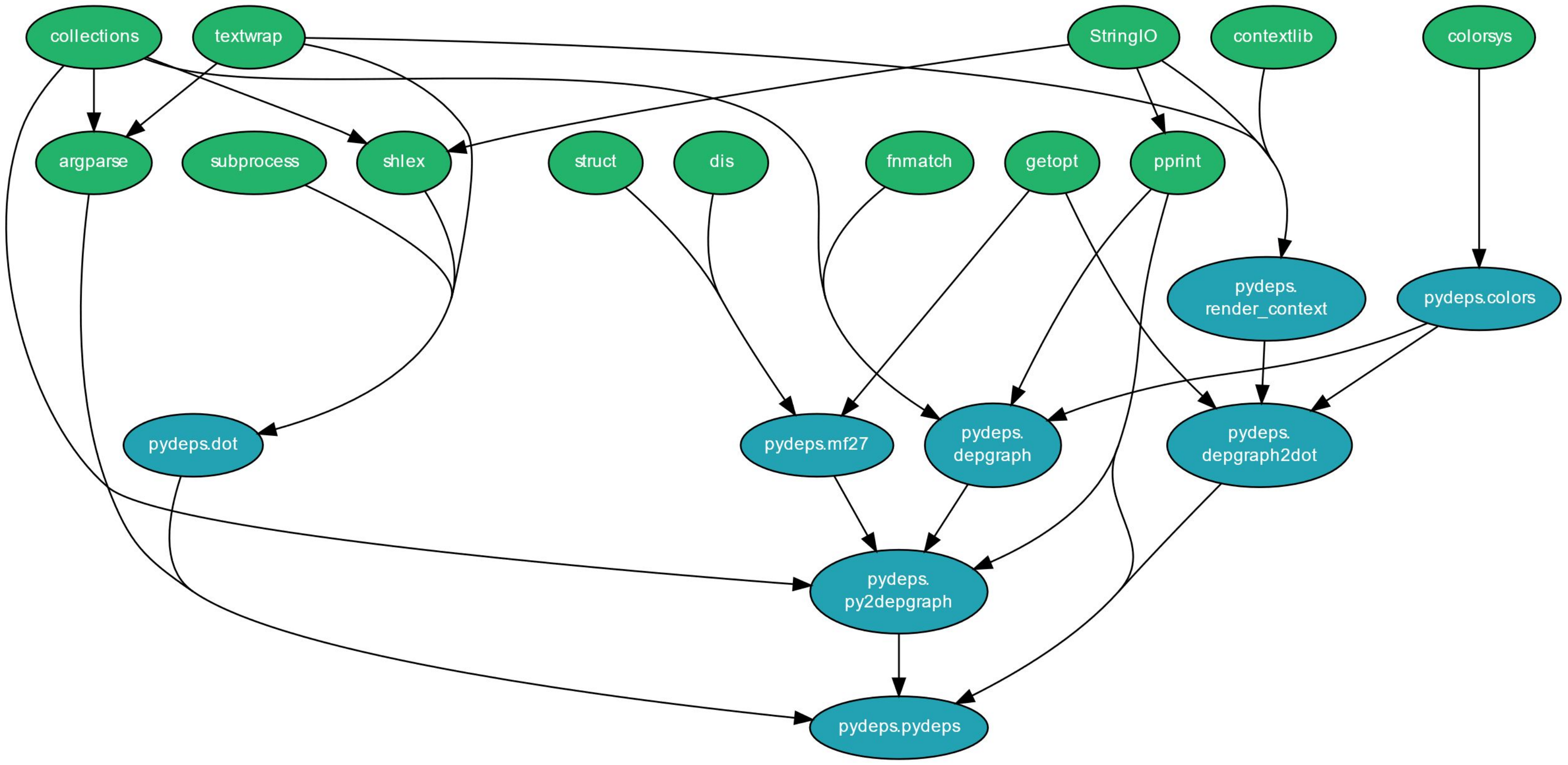
```
data                    var data: any
dateFormat
debug
debugger
decodeURI
decodeURIComponent
default
defaultStatus
delete
departFocus
devicePixelRatio
dispatchEvent
```

PROBLEMS TERMINAL ··· 1: node

```
info ℹ ｢wdm｣: Compiled successfully.
info changed file at
 WAIT  Compiling...
9:51:57 AM

info ℹ ｢wdm｣: Compiling...
 DONE  Compiled successfully in 63ms
9:51:58 AM

info ℹ ｢wdm｣:
info ℹ ｢wdm｣: Compiled successfully.
```

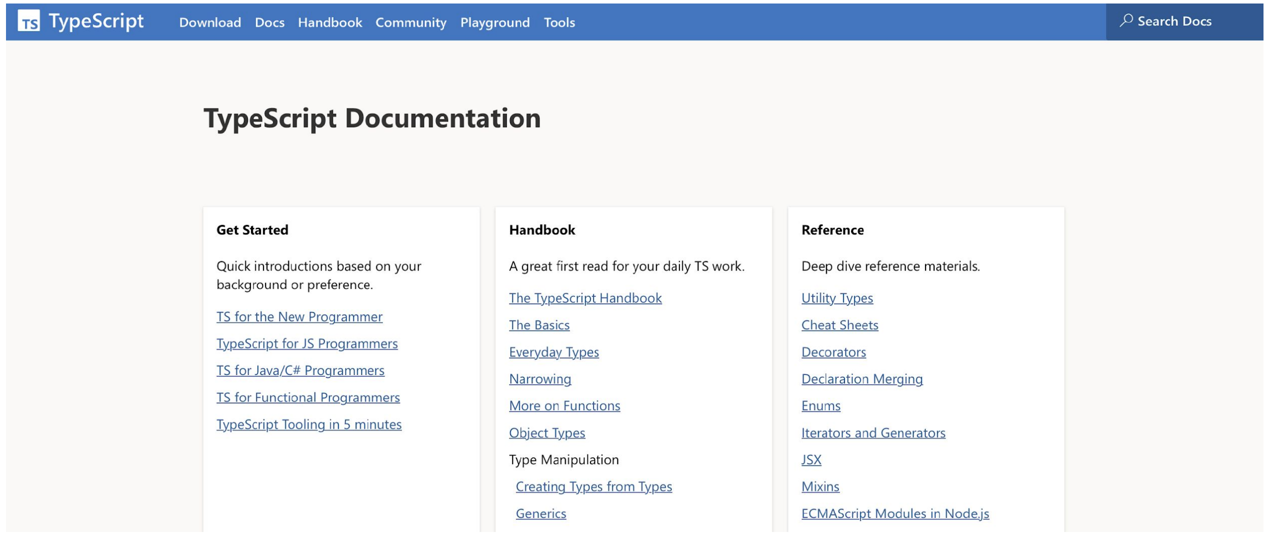
master* 0↓ 1↑ ⊗ 0 ⚠ 0 Ln 6, Col 21 Spaces: 2 UTF-8 LF JavaScript 🔔 1

Dependency maps



Consider documentation and tutorials judiciously

- Great for discovering entry points!
- Can teach you about general structure, architecture (more on this later in the semester)
- Often out of date.
- As you gain experience, you will recognize more of these, and you will immediately know something about how the program works
- Also: discussion boards; issue trackers



TypeScript Documentation

Get Started

Quick introductions based on your background or preference.

TS for the New Programmer
TypeScript for JS Programmers
TS for Java/C# Programmers
TS for Functional Programmers
TypeScript Tooling in 5 minutes

Handbook

A great first read for your daily TS work.

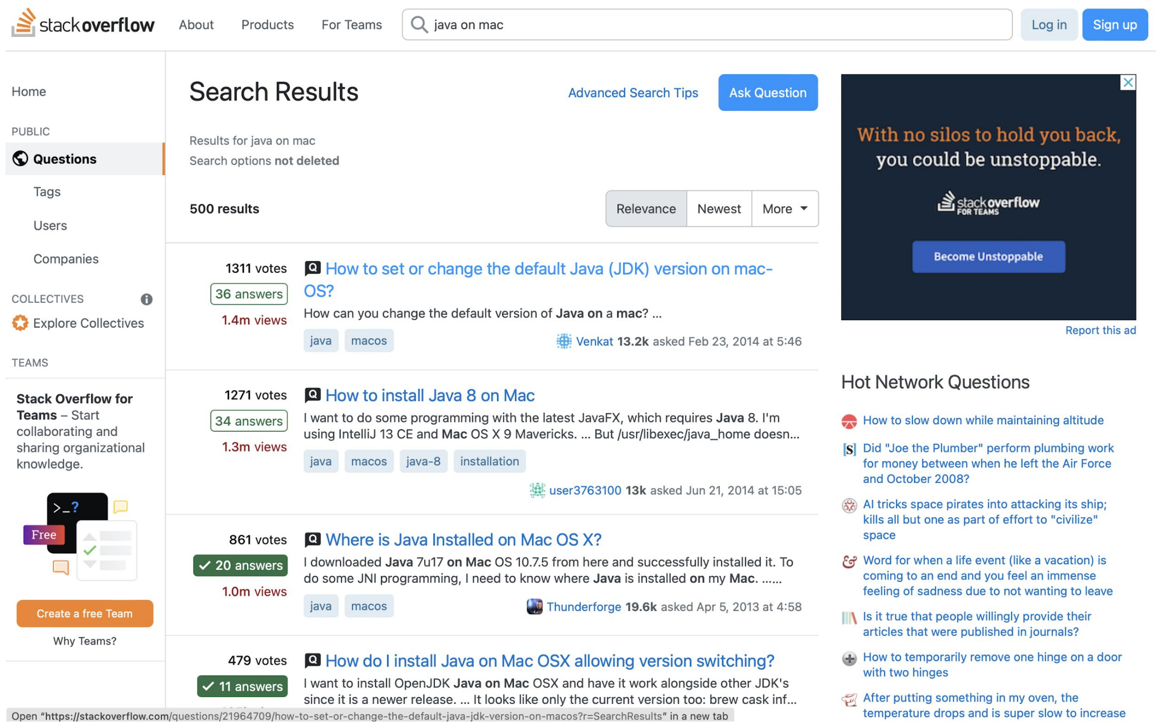
The TypeScript Handbook
The Basics
Everyday Types
Narrowing
More on Functions
Object Types
Type Manipulation
Creating Types from Types
Generics

Reference

Deep dive reference materials.

Utility Types
Cheat Sheets
Decorators
Declaration Merging
Enums
Iterators and Generators
JSX
Mixins
ECMAScript Modules in Node.js

Discussion Boards and Issue Trackers



stackoverflow About Products For Teams java on mac Log in Sign up

Home

PUBLIC

Questions

Tags

Users

Companies

COLLECTIVES

Explore Collectives

TEAMS

Stack Overflow for Teams – Start collaborating and sharing organizational knowledge.

Free

Create a free Team

Why Teams?

Search Results

Advanced Search Tips Ask Question

Results for java on mac
Search options **not deleted**

500 results Relevance Newest More

1311 votes | 36 answers | 1.4m views
How to set or change the default Java (JDK) version on macOS?
How can you change the default version of **Java on** a **mac**? ...
java macos
Venkat **13.2k** asked Feb 23, 2014 at 5:46

1271 votes | 34 answers | 1.3m views
How to install Java 8 on Mac
I want to do some programming with the latest JavaFX, which requires **Java** 8. I'm using IntelliJ 13 CE and **Mac** OS X 9 Mavericks. ... But /usr/libexec/java_home doesn...
java macos java-8 installation
user3763100 **13k** asked Jun 21, 2014 at 15:05

861 votes | ✓ 20 answers | 1.0m views
Where is Java Installed on Mac OS X?
I downloaded **Java** 7u17 **on Mac** OS 10.7.5 from here and successfully installed it. To do some JNI programming, I need to know where **Java** is installed **on** my **Mac**. ......
java macos
Thunderforge **19.6k** asked Apr 5, 2013 at 4:58

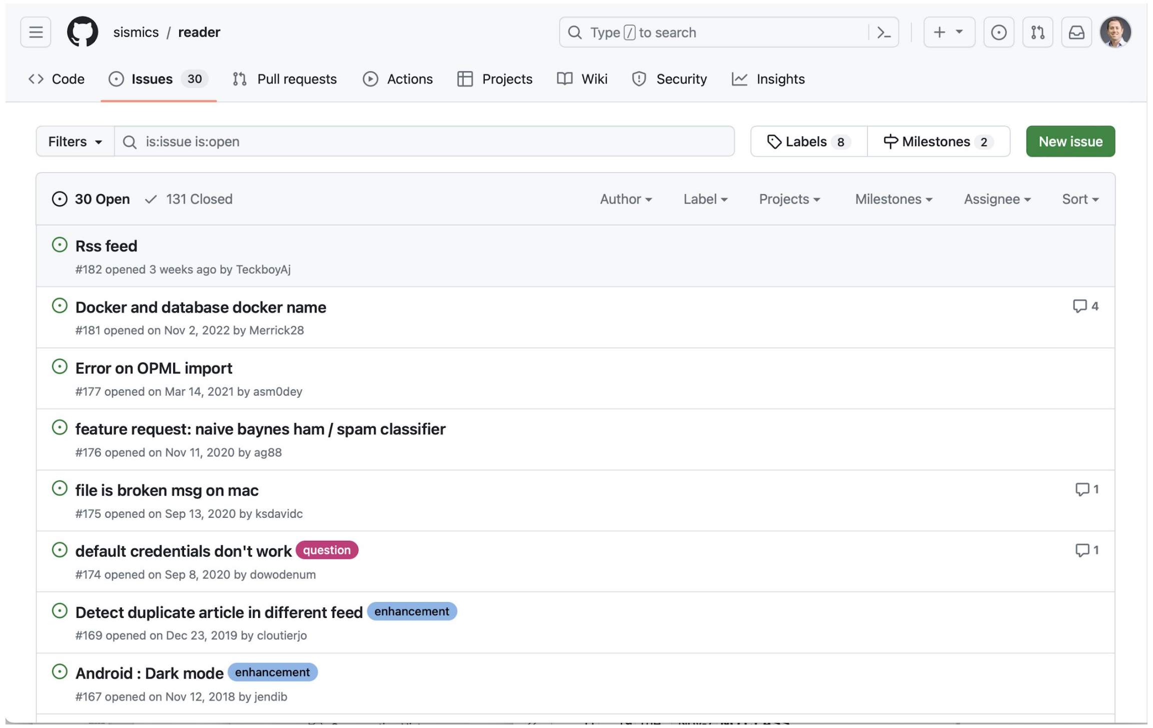
479 votes | ✓ 11 answers
How do I install Java on Mac OSX allowing version switching?
I want to install OpenJDK **Java on Mac** OSX and have it work alongside other JDK's since it is a newer release. ... It looks like only the current version too: brew cask inf...

With no silos to hold you back, you could be unstoppable.
stackoverflow FOR TEAMS
Become Unstoppable
Report this ad

Hot Network Questions

- How to slow down while maintaining altitude
- Did "Joe the Plumber" perform plumbing work for money between when he left the Air Force and October 2008?
- AI tricks space pirates into attacking its ship; kills all but one as part of effort to "civilize" space
- Word for when a life event (like a vacation) is coming to an end and you feel an immense feeling of sadness due to not wanting to leave
- Is it true that people willingly provide their articles that were published in journals?
- How to temporarily remove one hinge on a door with two hinges
- After putting something in my oven, the temperature drops and is super slow to increase

Open "https://stackoverflow.com/questions/21964709/how-to-set-or-change-the-default-java-jdk-version-on-macos?r=SearchResults" in a new tab



sismics / reader Type / to search

Code Issues 30 Pull requests Actions Projects Wiki Security Insights

Filters is:issue is:open Labels 8 Milestones 2 New issue

30 Open ✓ 131 Closed Author Label Projects Milestones Assignee Sort

- **Rss feed**
  #182 opened 3 weeks ago by TeckboyAj
- **Docker and database docker name** 4
  #181 opened on Nov 2, 2022 by Merrick28
- **Error on OPML import**
  #177 opened on Mar 14, 2021 by asm0dey
- **feature request: naive baynes ham / spam classifier**
  #176 opened on Nov 11, 2020 by ag88
- **file is broken msg on mac** 1
  #175 opened on Sep 13, 2020 by ksdavidc
- **default credentials don't work** question 1
  #174 opened on Sep 8, 2020 by dowodenum
- **Detect duplicate article in different feed** enhancement
  #169 opened on Dec 23, 2019 by cloutierjo
- **Android : Dark mode** enhancement
  #167 opened on Nov 12, 2018 by jendib

Dynamic Information Gathering
Change helps to inform and refine mental models

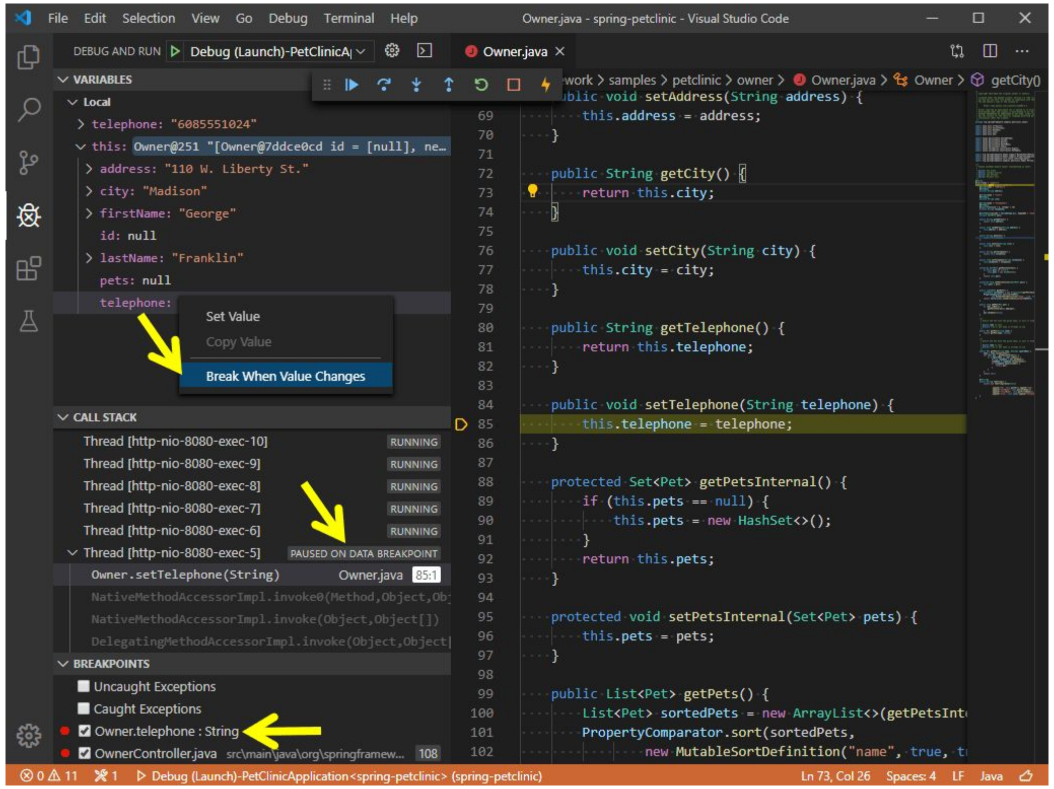
- Build it.
- Run it.
- Change it.
- Run it again.
- How did the behavior change?

How to start?

- Confirm that you can build and run the code.
  - Ideally both using the tests provided, and by hand.
- **Confirm that the code you are running is the code you built!**
- Confirm that you can make an externally visible change
- How? Where? Starting points:
  - Run an existing test, change it
  - Write a new test
  - Change the code, write or rerun a test that should notice the change
- Ask someone for help

Probes: Observe, control or “lightly” manipulate execution

- `print(“this code is running!”)`
- Structured logging
- Debuggers
  - Breakpoint, eval, step through / step over
  - (Some tools even support remote debugging)
- Delete debugging
- Chrome Developer Tools


File
Edit
Selection
View
Go
Debug
Terminal
Help
Owner.java - spring-petclinic - Visual Studio Code
DEBUG AND RUN
Debug (Launch)-PetClinicA|
Owner.java
work > samples > petclinic > owner > Owner.java > Owner > getCity()
VARIABLES
Local
Set Value
Copy Value
Break When Value Changes
CALL STACK
BREAKPOINTS

```
> telephone: "6085551024"
v this: Owner@251 "[Owner@7ddce0cd id = [null], ne…
  > address: "110 W. Liberty St."
  > city: "Madison"
  > firstName: "George"
    id: null
  > lastName: "Franklin"
    pets: null
    telephone:
```

```
Thread [http-nio-8080-exec-10]    RUNNING
Thread [http-nio-8080-exec-9]     RUNNING
Thread [http-nio-8080-exec-8]     RUNNING
Thread [http-nio-8080-exec-7]     RUNNING
Thread [http-nio-8080-exec-6]     RUNNING
Thread [http-nio-8080-exec-5]     PAUSED ON DATA BREAKPOINT
  Owner.setTelephone(String)      Owner.java 85:1
  NativeMethodAccessorImpl.invoke0(Method,Object,Ob
  NativeMethodAccessorImpl.invoke(Object,Object[])
  DelegatingMethodAccessorImpl.invoke(Object,Object|
```

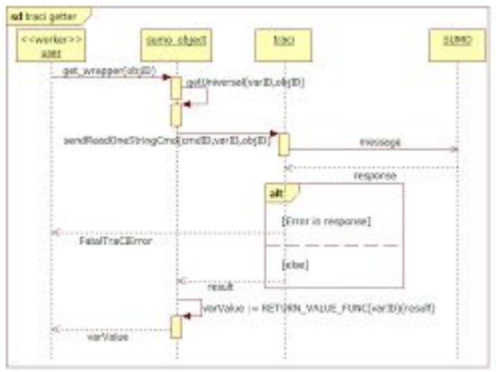
- [ ] Uncaught Exceptions
- [ ] Caught Exceptions
- [x] Owner.telephone : String
- [x] OwnerController.java src\main\java\org\springframew... 108

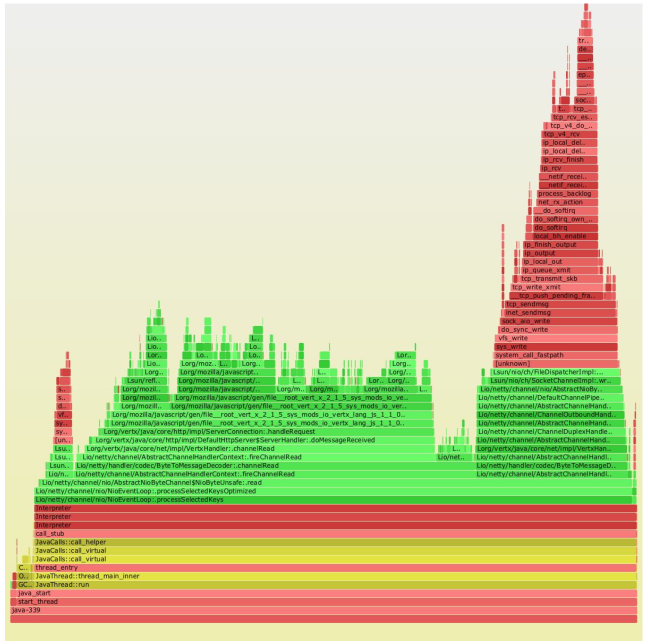
```java
ublic void setAddress(String address) {
        this.address = address;
    }

    public String getCity() {
        return this.city;
    }

    public void setCity(String city) {
        this.city = city;
    }

    public String getTelephone() {
        return this.telephone;
    }

    public void setTelephone(String telephone) {
        this.telephone = telephone;
    }

    protected Set<Pet> getPetsInternal() {
        if (this.pets == null) {
            this.pets = new HashSet<>();
        }
        return this.pets;
    }

    protected void setPetsInternal(Set<Pet> pets) {
        this.pets = pets;
    }

    public List<Pet> getPets() {
        List<Pet> sortedPets = new ArrayList<>(getPetsInt
        PropertyComparator.sort(sortedPets,
                new MutableSortDefinition("name", true, t
```

```
⊗ 0 ⚠ 11   1   Debug (Launch)-PetClinicApplication<spring-petclinic> (spring-petclinic)   Ln 73, Col 26   Spaces: 4   LF   Java
```

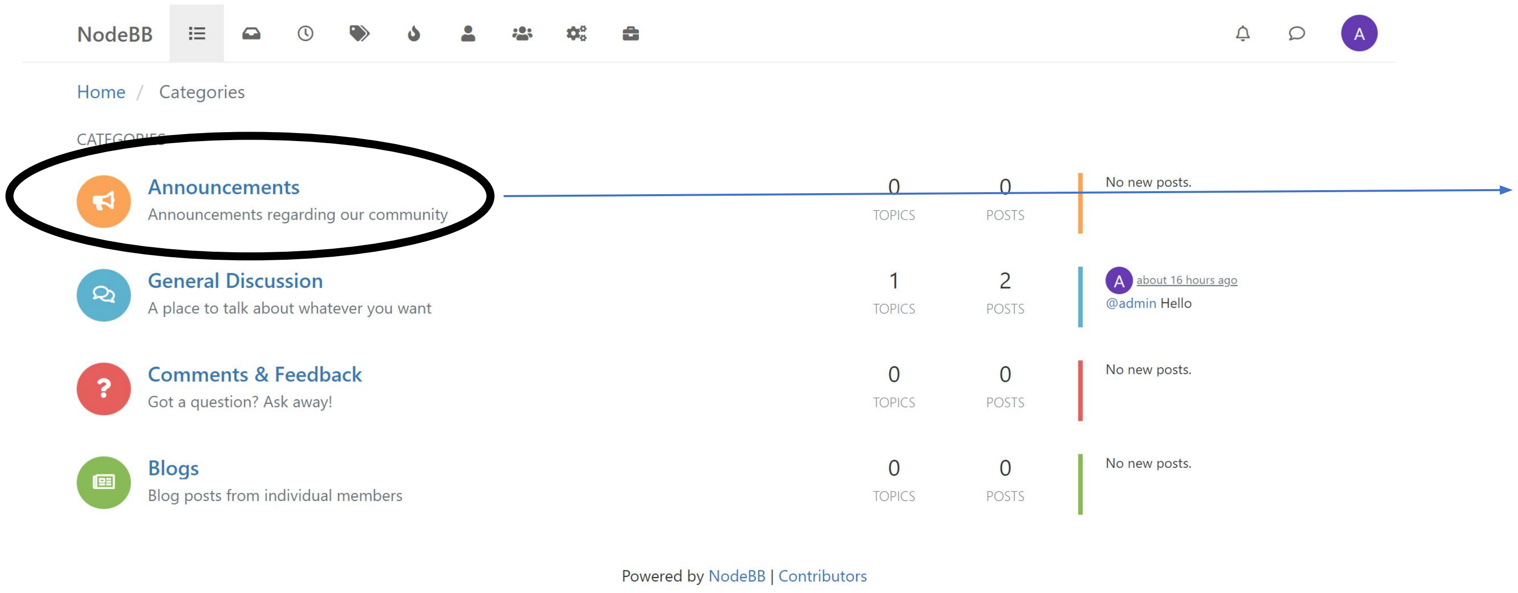
Runtime code analysis tools

- Collect runtime traces and visualize them
    - Flame graphs
    - Sequence diagrams
- Use judiciously





Tip: Find a particular thing and trace the action backward


NodeBB

Home / Categories

CATEGORIES

Announcements Announcements regarding our community	0 TOPICS	0 POSTS	No new posts.
General Discussion A place to talk about whatever you want	1 TOPICS	2 POSTS	A about 16 hours ago @admin Hello
Comments & Feedback Got a question? Ask away!	0 TOPICS	0 POSTS	No new posts.
Blogs Blog posts from individual members	0 TOPICS	0 POSTS	No new posts.

Powered by NodeBB | Contributors

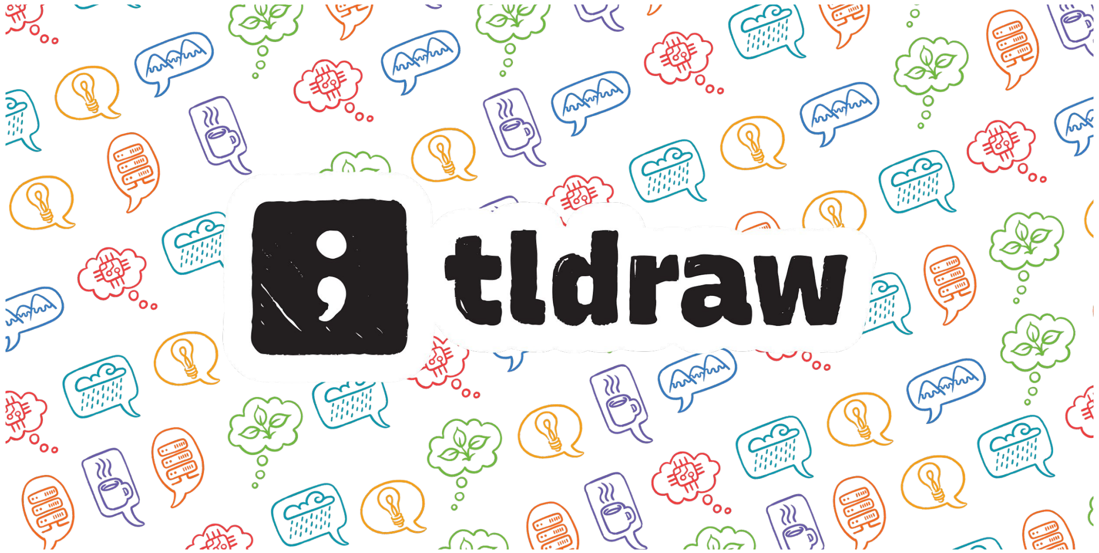
E.g.,

Where do categories come from?

How are they stored?

How are they rendered?

Let’s try some of these techniques again…



https://github.com/tldraw/tldraw

Remember…

- Reading and understanding code is one of the most important skills you should learn
- It's common to get stuck or feel overwhelmed. **Don't give up!**
- Consider yourself lucky! Things are much easier today

Learning Goals

- Understand and scope the task of taking on and understanding a new and complex piece of existing software
- Appreciate the importance of configuring an effective IDE
- Contrast different types of code execution environments including local, remote, application, and libraries
- Enumerate both static and dynamic strategies for understanding and modifying a new codebase