17-803 Empirical Methods Spring 2024

Research Project Kickoff Presentations

Tuesday, February 20

Project Title

Student / Team member names

Example Slide 1

Problem - gap - hook / Research questions

Maecenas luctus tempus augue, at ultricies nulla. Pellentesque ac tortor. Quisque aliquet interdum odio, a rhoncus nisl. Ut porta mi.

Example Slide 2

Overview of study design

- Which methods
- Combined how / why

Maecenas luctus tempus augue, at ultricies nulla. Pellentesque ac tortor. Quisque aliquet interdum odio, a rhoncus nisl. Ut porta mi.

Kaia Newman

Professional Programmers with ADHD: Coping Strategies, Consequences of Disclosure, and Accommodation Processes

Kaia Newman

Fixing SE accommodation processes

Many neurodivergent software engineers experience difficulties in their work. However, what these difficulties are (especially for ADHD software engineers), how they can be addressed, and how they are addressed *now*, is not well understood.

However, we do know that ADHD accommodations processes in software workplaces are:

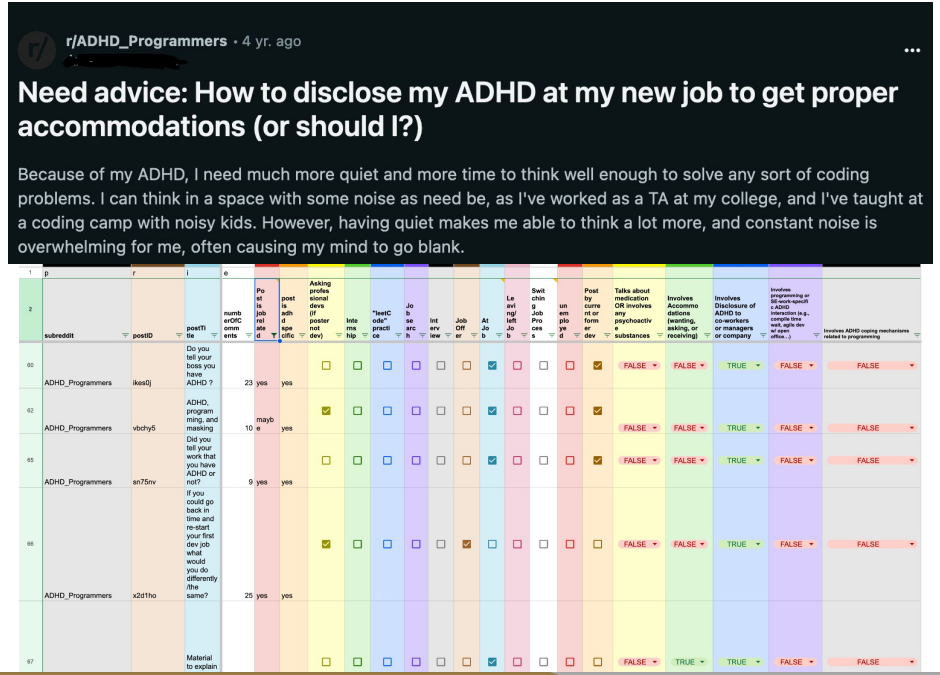
- Opaque: it is not clear what accommodations a company has available without disclosing one's neurodivergence
- A risk to the disclosee in terms of being exposed to stigma or discrimination, much of which can still occur despite legal protections
- Vague: managers, and even ADHD software engineers themselves, are not trained to understand what accommodations may help with ADHD software work impacts

What can we do about it in a lightweight, cheap way that helps everyone involved?

Qualitative Methods and RQs

RQ1: What challenges do ADHD software engineers face at work?
RQ2: What coping strategies and accommodations do ADHD software engineers use to address these challenges?
RQ3: What are some of the consequences of disclosure of ADHD in software workplaces (positive and negative)?

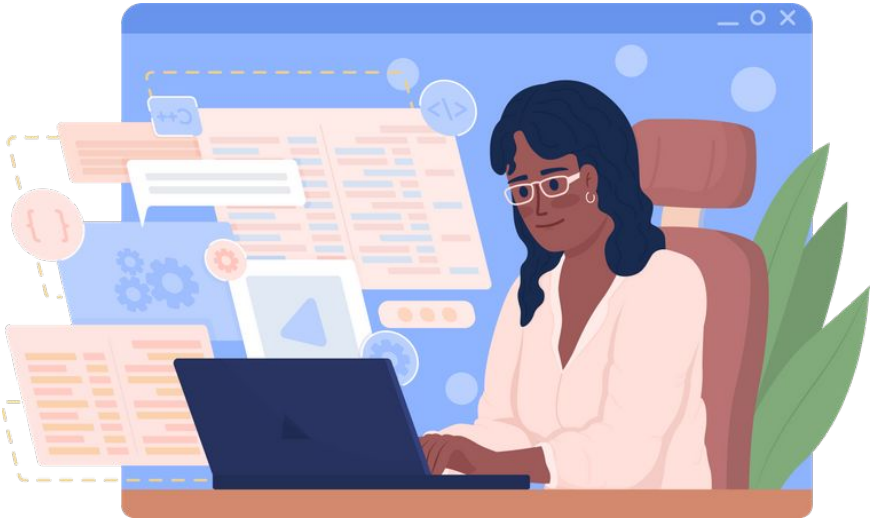
- *Scrape r/ADHD_Programmers using relevant keywords*
- *Manually check to see if posts concern professional programming or these RQs*
- *Thematically analyze a random sample of posts from each "bin" of coping strategies, challenges, disclosure, and accommodations*



Quantitative Methods and RQs

RQ1: Which strategies/accommodations may help with which work impacts?
RQ2: What are the perceived efficacy, feasibility, and scope of the most salient strategies/accommodations?
RQ3: Do neurotypicals and other neurodivergent people think these strategies/accommodations would be helpful/unhelpful for them?

- *Using themes from the qualitative analysis, construct a survey using the most popular strategies/accommodations/impacts*
- *Can triangulate, ask for Likert scale efficacy, gather needed demographics data, and ask if accommodations would be or are feasible at their companies*
- *Deploy survey at CMU, other universities, online, and in SE workplaces*



Catarina Gamboa

Challenges in *Adopting and Using* LiquidHaskell

Catarina Gamboa

Reliability has become more crucial, nowadays.

Type systems are useful to find bugs before runtime.

Research shows Liquid Types in improve reliability further.

```
{-@ head :: [a] -> a @-}
head (x:_) = x
```

However,

We don’t know about the real **adoption** of Liquid Types or the **challenges** developers face when adopting them.

Rq.1 How are developers adopting LiquidHaskell in their projects?

Rq.2 What are the issues developers are facing with LiquidHaskell?

Methodology

Quantitative analyses

repos using LiquidHaskell
different users
projects last

Qualitative analyses

Project scopes/LH use cases
Types of issues mentioned

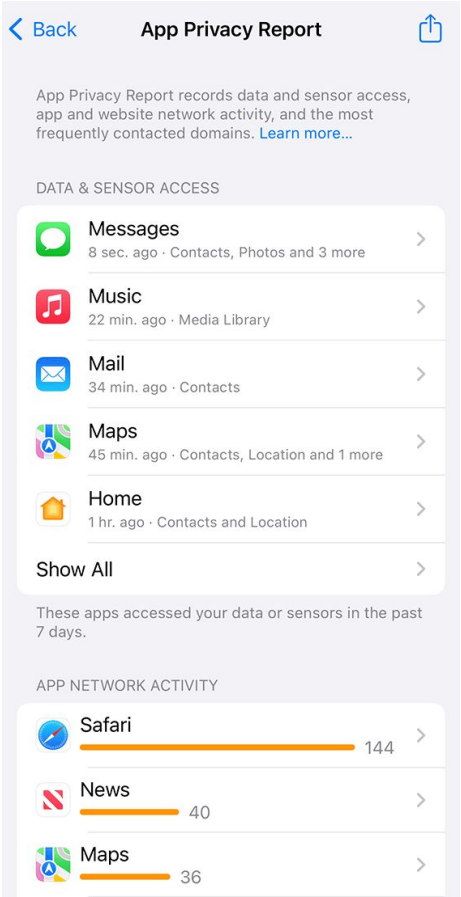
Xiaoyuan (Owen) Wu

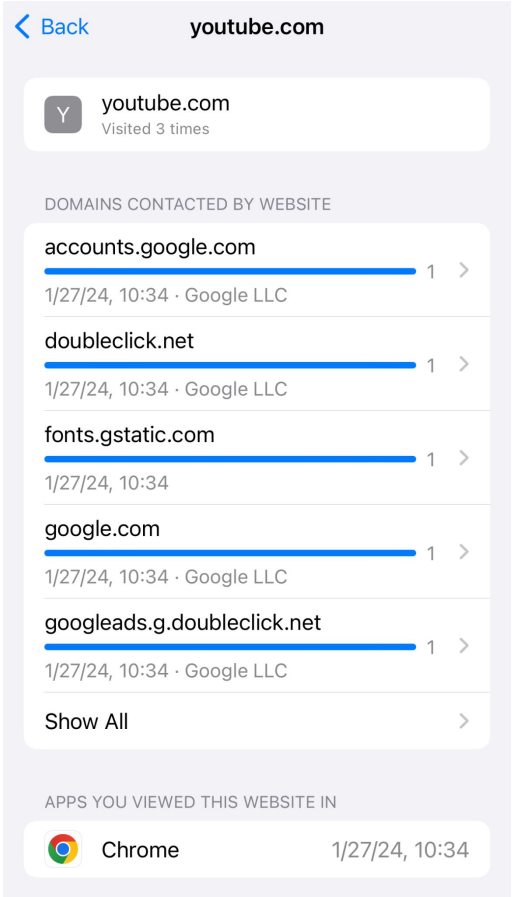
App Privacy Report: User Perceptions and Online Tracking Implications

Xiaoyuan (Owen) Wu

Background

Released in December 2021, the iOS App Privacy Report (privacy report) aims to provide users with "a more complete picture of how the apps [they] use treat [their] data[1]."





[1] Apple. About app privacy report. https://support.apple.com/en-us/102188

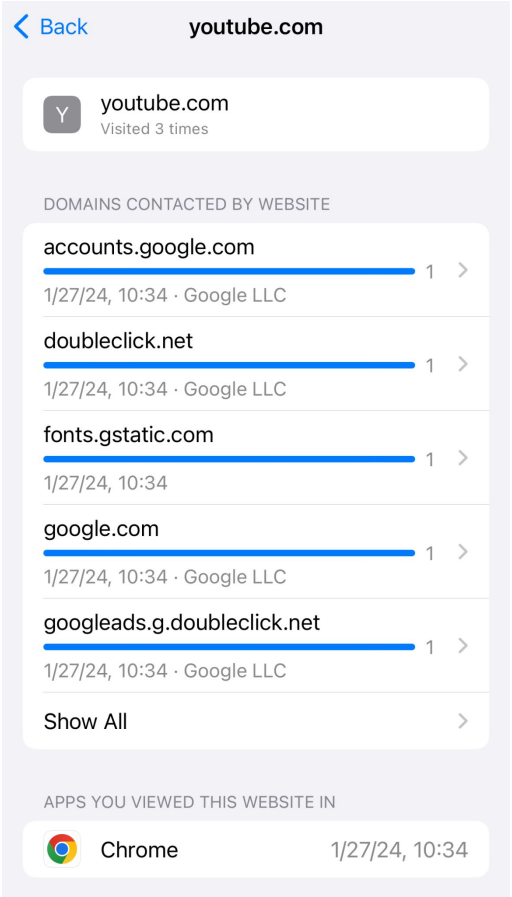
Research Questions

- RQ1 (Qualitative): Does the privacy report help users identify and understand overpermissioning, cross-app tracking, or third-party data collection happening on their phone?
- RQ2 (Qualitative): What are users' attitudes toward the information they learn from the privacy report? Are there intentions to change how they use their phone?
- RQ3 (Quantitative): Through the privacy report, how prevalent is online tracking on mobile phones?

Quantitative Analysis

Collect participants' privacy report and analyze:

1. Amount of domains contacted
2. Functionality vs. tracking
   E.g., *fonts.gstatic.com* vs. *doubleclick.net*
3. First vs. third-party
   E.g., youtube vs. facebook contacting *google.com*
4. Apps vs. websites
   E.g., Youtube App vs. youtube.com



Taylor McCampbell

The State of Cyber Security Education in Pennsylvania

Taylor McCampbell

Problem(s), Gap, Hook

Problem 1 - Critical shortage of cyber security professionals in the United States.

Problem 2 - High teacher attrition rates, low test scores, funding issues, behavioural issues, grade inflation, etc.

Gap - **No research** on Pennsylvania 9-12 cyber security course access. **No statistics** from Penn Department of Education. **No surveys** of teacher/student readiness to implement/learn the content.

Hook - This project will survey the state of cyber security education in Pennsylvania to be used for **filling gaps in our public education system**. In filling these holes, the **gap between the number of cyber security professionals and available jobs will begin to close**.

Research Questions

RQ1: What percentage of Pennsylvania high schools offer cyber security courses?

RQ2: What percentage of Pennsylvania high schools offer pathways to cyber security courses?

RQ3: To what extent are students ready to begin building a technical foundation in cyber security?

RQ4: To what extent are teachers equipped to offer cyber security courses?

Methodology

Quantitative

- What cyber security courses are offered, how many courses are offered, how many teachers are certified to teach them
- Scraping websites, emails, phone calls

Qualitative

- Interviews with STEM teachers on student/teacher readiness.
- Good mix of large/small districts
- Thematic analysis of interview transcripts

Jeffrey Chen

Sustainable Open Source Community

Feb 15th, 2024

Jeffrey Chen(weigenc)

Overview

- Reuse of open source artifacts in software ecosystems are important
- Reports of stress and burnout among open source developers are increasing
- Explore sustainability challenges in open source community

```
import argparse
from warcio.archiveiterator import ArchiveIterator
import pyspark
import re
import os
```

Problem Gap Hook

Problem: How to maintain a dynamic open source community

Gap 1 From the productor: The paid contributors and volunteers are losing
Gap 2 From the consumer: There is a lack of tools to help developers quickly evaluate the quality, stability, and suitability of open source projects.

Hook: Discover indicators that affect open source projects during surveys and research, and display them with visual visualizations to help maintain the community

Research Question

From productor's perspective:
RQ 1: What are the main reasons for volunteer contributors to drop out of open source projects?
RQ 2: Why some contributors will continue to maintain the open source project?
RQ 3: In what situations do volunteer contributors experience stress?
RQ 4: Which past interventions, such as contribution guidelines and code of conducts, have been successful in retaining contributors and easing transitions?

From developer's perspective:
RQ1: In the current development, what indicators are the most important
RQ2: Which tools or resources can best help them evaluate the quality, stability, and suitability of the project

Methodology

Qualitative
- User Interview & Survey

Quantitative
- Some data from own prototype (In development)
- Some data from the existed API (like github, snyk, libhunt)

dependencies situation of **express**

Dependency Name	Last Commit Time
accepts	2022-02-02T23:46:50Z
array-flatten	2023-12-07T08:04:06Z
body-parser	2023-02-22T01:25:54Z
content-disposition	2021-12-10T22:50:59Z
content-type	2023-01-29T19:21:49Z
cookie	2023-11-07T04:58:58Z
cookie-signature	2023-04-12T23:19:54Z
debug	2023-06-04T11:10:21Z
depd	2021-11-12T06:32:53Z
encodeurl	2019-01-01T02:50:37Z

Elizabeth Gilbert

Challenges for Tooling Developers

•••

Elizabeth Gilbert

WebAssembly (Wasm) bytecode is growing beyond the browser!

Compilation target for many languages.

Originally created to run bytecode in the browser at high speed, now growing to new domains and use cases.

Lacking in tooling, but we want to change that.

We want to create an awesome ecosystem for developing Wasm tooling (debuggers/dynamic analyses).

But first, how should we build it?

What challenges do dynamic analysis developers face?

What technologies/formats/protocols are used to alleviate these challenges?

- What features/types of support/integrations would be helpful?

What do dynamic analysis developers do to work around these challenges? (hacks)

What platforms/PLs support analysis development well?

- What features in different PLs/platforms are good/helpful?
- What are they missing?

Discovering challenges and their remedies for a tooling development ecosystem.

Phase 1 - Survey		Phase 2 - Interviews
High-level, quantitative	**Insight**	Low-level, qualitative
Bytecode dynamic analysis developers	**Population**	Survey subgroup
Weighted challenges and remedies	**Result**	Motivations, desires, nuance
Prioritization of features	**Contribution**	How to make features impactful

Harrison Green

How do Hackers Hack?

Harrison Green

Human-based Binary Exploitation

Problem: Expert humans demonstrate a great capacity to *understand* and *exploit* computer programs, yet modern automated techniques are nowhere near as capable.

Gap: Zero studies on human binary exploitation; a few studies on reverse-engineering focusing specifically on decompilation.

Hook: Binary exploitation is one of many complex tasks where human *creativity* and *ingenuity* somehow surpasses computational methods in navigating high dimensional search spaces. Observing how humans perform binary exploitation will not only guide the development of program analysis tools–critical for the development of secure software, but may also shed light on the more general question of how humans *think about* and *solve* complex problems.

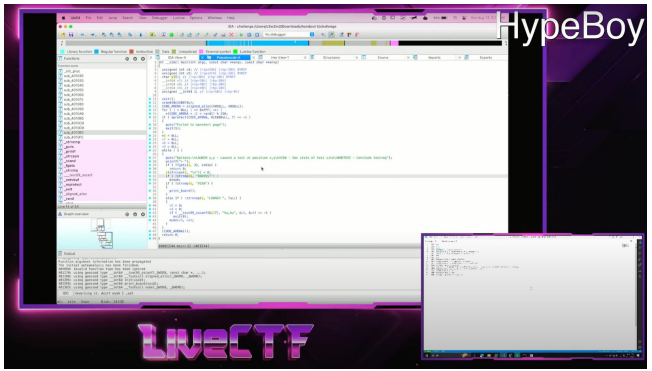
Research Questions

RQ1: What techniques/tools do humans use to understand programs?

RQ2: How do humans develop and debug exploits?

RQ3: Where are the bottlenecks / pain-points in this process? Where do people get stuck?

Methods



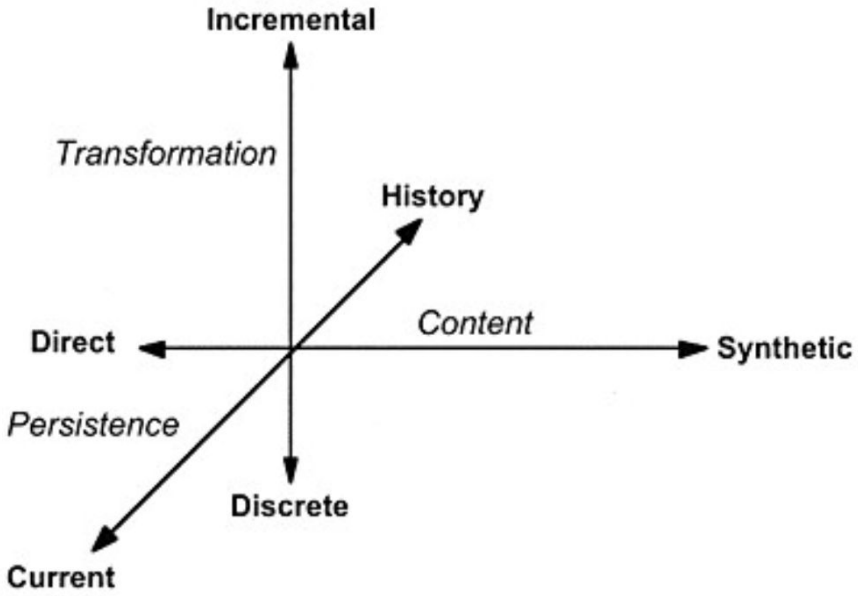
1. Qualitative/quantitative analysis of 1v1 CTF matches

- LiveCTF (at DEFCON) and Pwny Racing
- Screen recordings of top hackers competing to solve binary exploitation problems

2. Qualitative interview/survey

- Interview CTFers
- Understand how people approach solving problems / the thought process / what tools they use / where they get stuck

Hemant Gouni / Long Nguyen



An Empirical Study of Domain-Specific Debuggers

Hemant Gouni / Long Nguyen

Classifying domain-specific debugging logic

We want to understand how people write and debug programs.

There exist zero studies on the role played by custom debugging logic.

Debugging is an understudied problem relative to its importance to software engineering, and a rigorous study of debugging logic would advance our ability to write better programs.

RQ1: What kinds of debugging logic is already written– and how often?

RQ2: Does the time required to maintain custom debugging logic exceed the time saved?

Is any of this actually useful?

Qualitative

- Taxonomy of debugging logic
  - What parts of computation/data need visualization?
  - What capabilities/features does the custom debugging logic have?

Data gathered from corpus of open-source projects

Quantitative

- Does it *actually* save time? Do people *perceive* that it saves time?

Data from observational + interview studies

Luís Gomes

Developers, Drawings and Code: Creating generative Sketch to Code tools

Luís F. Gomes

Problem Definition

Developers productivity could be improved if **generative code tools understand sketched mental models**. ML workflows and data visualization are examples of mental models that can be directly transferred to a sketch.

In-IDE tools for this domain are rare and **little information about how to leverage sketching practices to generate code** is available.

In this project we explore **how AI-based sketch to code tools can be created and used** to improve their acceptance and usability by ML programmers and data scientist.

Research Questions

RQ1: What **patterns/characteristics** are commonly **observed** in ML developers and Data Scientists sketches to represent the same concept? E.g. A neural network, a transformer, a plot.

RQ2: How do different **sketching patterns** impact the **accuracy** of code generation? Explorative/Constructivist E.g. Arrows to separate step? Left to right? Top to bottom?

RQ3: How do developers **perceive the usefulness** of in-IDE sketch to code generation tools? What developers like, dislike and perceive as useful to implement.

Study Methodology

01

Qualitative
User Study + Interview
Users sketch and use the tool to generate code, providing **feedback about perceived usefulness and things to improve**.

02

Quantitative
Tool Benchmarking
Compare task **completion rate** (correct subtasks) and **generation accuracy** (user modified code).

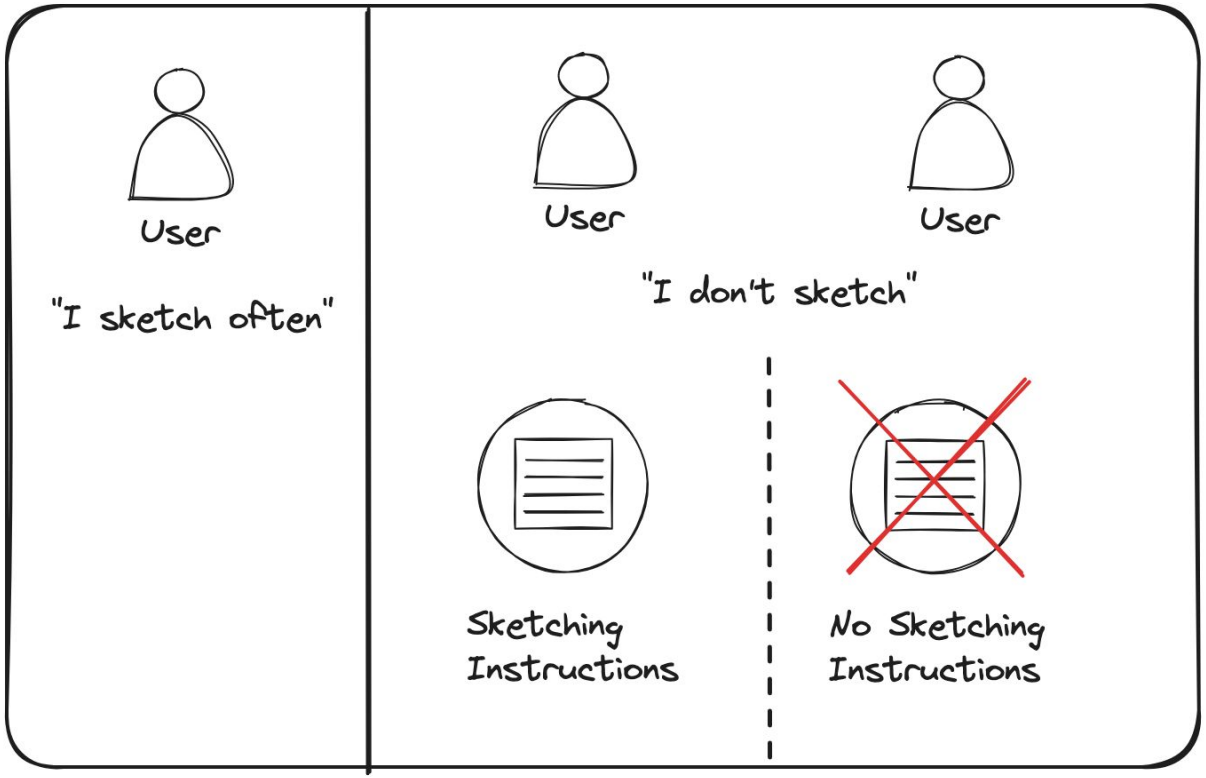
Study Design

User perform ML tasks:

1. **Sketch ML workflow**
2. **Generate Code from Sketch**
3. **Modify/use generated code**

3 groups:

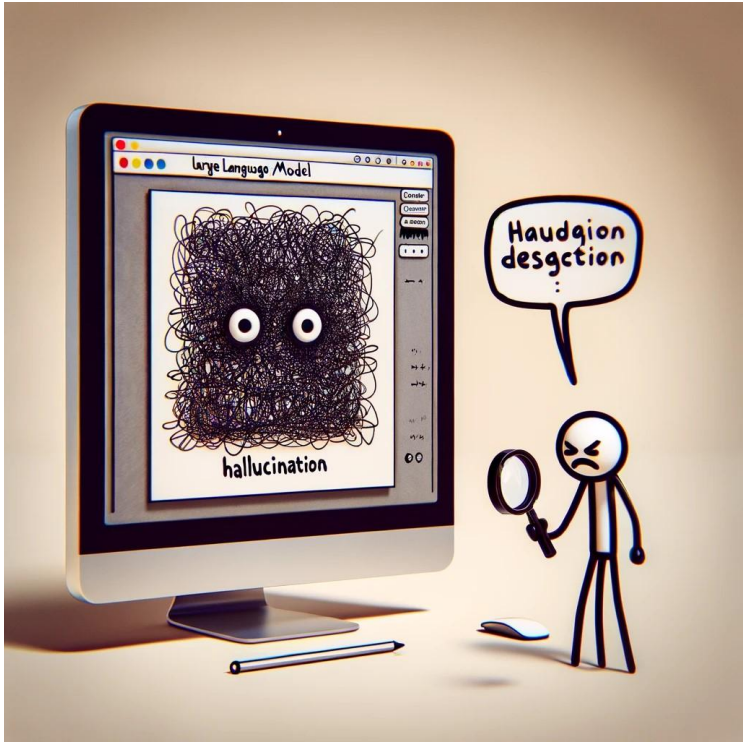
- Sketch
- NoSketch + Instructions
- NoSketch + NoInstructions



Yining She

Mitigating Hallucination in LLMs: An Empirical Comparison of Prompting Strategies

Yining She



Problem & Gap

Large Language Models are very powerful and can be adapted to many tasks.

A significant drawback for LLMs is their tendency to "**hallucinate**" - generating content that appears factual but is ungrounded.

Example:

Me: Given x+1 = 11, x=?

Chatgpt: x=10 because (1) x = 11-1 (2) 11-1= 9 (3) therefore, x=10!

Problem & Gap

Large Language Models are very powerful and can be adapted to many tasks.

A significant drawback for LLMs is their tendency to "**hallucinate**" - generating content that appears factual but is ungrounded.

Without access to datasets or the ability to modify LLM structure, many studies have proposed **prompt strategies** to mitigate LLMs' hallucination, e.g. chain-of-thought, self-reflection

*However, there is no study that **empirically compares** different prompting techniques.*

Research Questions

RQ1: What are the characteristics of generated results for each prompt strategy?

RQ2: How well can each prompt strategy perform in different text generation tasks?

RQ3: How do users select and adapt prompt strategies in practice?

RQ4: What challenges do users encounter when applying prompt strategies?

Methods

RQ1: What are the characteristics of generated results for each prompt strategy?

[Qualitative] Analyze content of each method's output and develop code scheme for them

RQ2: How well can each prompt strategy perform in different text generation tasks?

[Quantitative] Evaluate the performance of each prompt strategy across a range of tasks using established benchmarks

Methods Cont.

RQ3: How do users select and adapt prompt strategies in practice?

RQ4: What challenges do users encounter when applying prompt strategies?

[Qualitative] Conduct interviews with LLM users to gather insights into their preferences, strategies, and challenges.

Claudia Mamede

need to find a way to add "interpretability" to the title

A Comparative Analysis of LLM-based Chain of Thought and Human Decision-Making in Vulnerability Detection

Claudia Mamede

Problem

Transformers have achieved promising results in the field of vulnerability detection. **However, these models are black boxes so security experts avoid using them in critical scenarios where *interpretability* is needed.**

Gap

1) The definition of *interpretability* is not clear, especially in the security field.
2) LLM-based Chain of Thought is commonly used to communicate model behaviour to a diverse audience. But no one has compared the efficacy of LLM-based CoT and the traditional CoT performed by humans in a security context.

Hook

Existing research focus primarily on performance, ignoring the practical needs of those in the field.

This work provides insights for future researchers that want to develop AI-based security tools taking into consideration experts' concerns regarding interpretability.

Research Question and Methodology

1. **What are the characteristics of an interpretable LLM-based vulnerability detection, according to the security community?**

   Qualitative study with interview or survey + coding

accuracy + response time + interpretability

Existing work establishes relationship

2. **How does the efficacy of an LLM-based chain of thought compare to that of a human's in the context of vulnerability detection?**

   2.1 How does the efficacy of chain of thoughts compare across different expertise levels (e.g. novice vs experts)?

   2.2 How does the efficacy of chain of thoughts compare across different vulnerability types (e.g. xss, buffer overflow)?

   Controlled Experiment with Post-Experiment Qualitative Interviews

   Independent vars: expertise level, vulnerability type, file size(?), auxiliary reports(?)

   Dependent vars: accuracy and response time + interpretability *(Likert scale based on the codes?)*

Yuchen Shen

An Empirical Study on the Real-time Guided Interview Elicitation Tool

Yuchen Shen

Study Overview

Model (Interview Assistant): A backend Masked Language Model (MLM) that generates related concepts for a chosen word of interest (clicked by interviewer). E.g: Kitchen - big, dilapidated, fireplace, stove, etc.

RQ1: Does the Guided Interview Elicitation Tool improve requirement elicitation?

Method: Recruit 8 interviewers (trained) and 32 interviewees, randomly break into Control (C) and Test (T) groups to test the tool via interviews on 4 different directory service topics.

Topic IDs:	apartment=1	restaurant=2	hiking=3	clinic=4
Interviewer	**round1**	**round2**	**round3**	**round 4**
T1	1	4	2	3
T2	2	1	3	4
T3	4	3	1	2
T4	3	2	4	1
C1	3	1	2	4
C2	2	3	4	1
C3	4	2	1	3
C4	1	4	3	2

Analyze the resulting interview transcripts to evaluate whether there is an improvement of requirement elicitation on the T group:

Qualitative: Analyze the transcripts (manual & GPT) to discover and count the number of requirement-related concepts.
Quantitative: Use Statistical Measurements to test statistical significance of the C vs T group elicitation results.

Additional Plans for Extension

Goals: 1) Improve the current speech recognition that generates real-time transcripts;

2) Add additional supports for elicitation apart from the current MLM model.

RQ2: Does the Tool Version 2.0 improve the transcription quality?

RQ3: Does the Tool Version 2.0 improve requirement elicitation?

Method: Run a smaller scale mini-study on new recruits (interviewer & Interviewee) to conduct interviews.

Analyze resulting transcripts.

Qualitative: (Manually or with GPT) Go through the transcripts to extract requirement-related concepts.

Quantitative: Test statistical significance of the above result. Additionally, use NLP measurements to measure transcription quality for the old versus new tool versions.

Hwei-Shin Harriman

Interactive Geometry Proofs (1/4) Hwei-Shin Harriman

Problem: Students who learn geometry often struggle to fully understand them. When students are tested on traditional 2-column proofs, studies have found that they do not identify the key ideas behind the proofs and the generalizability of proofs.

Gap: Studies have examined the impact of various learning interventions and teaching styles on improving students' comprehension, however, none have examined the impact of interactivity on comprehension.

Hook: Examine how 2–column proofs with added interactive elements such as highlighting and linking (see mockups) can improve student comprehension.

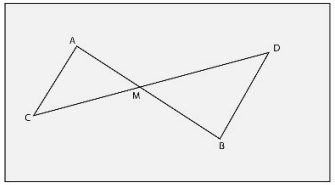
Research Questions (2/4)

- **RQ1:** Does adding interactivity improve students' reading comprehension of geometric proofs?
- **RQ2.1:** What support do teachers want to teach geometric proofs?
- **RQ2.2:** What do students misunderstand about geometric proofs?
- **RQ2.3:** What feedback to teachers have for the intervention we designed?

Example Interactive Proof (Mockup) (3/4)

Initial Construction

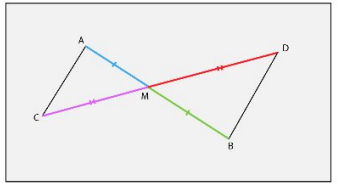
Problem: As shown, $\overline{AB}$ and $\overline{CD}$ intersect at point M, $\overline{AM} = \overline{BM}$ and $\overline{CM} = \overline{DM}$; then, must $\overline{AC}$ and $\overline{DB}$ be parallel with each other?



Statements	Reasons
$\overline{AM} \cong \overline{BM}$, $\overline{CM} \cong \overline{DM}$	Given
$\angle AMC \cong \angle BMD$	Def. of vertical angles
$\triangle AMC \cong \triangle BMD$	SAS Triangle Congruency
$\angle MAC \cong \angle MBD$	Def. of congruent polygons
$\overline{AC} \parallel \overline{DB}$	Def. of alternate interior angles

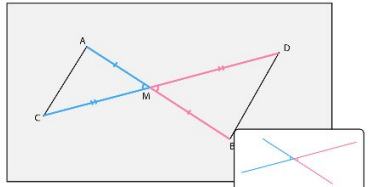
Proof Steps, HOVER over a row

Problem: As shown, $\overline{AB}$ and $\overline{CD}$ intersect at point M, $\overline{AM} = \overline{BM}$ and $\overline{CM} = \overline{DM}$; then, must $\overline{AC}$ and $\overline{DB}$ be parallel with each other?



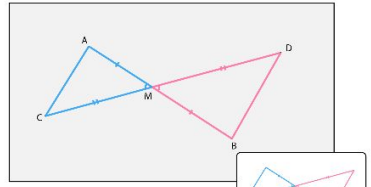
Statements	Reasons
$\overline{AM} \cong \overline{BM}$, $\overline{CM} \cong \overline{DM}$	Given
$\angle AMC \cong \angle BMD$	Def. of vertical angles
$\triangle AMC \cong \triangle BMD$	SAS Triangle Congruency
$\angle MAC \cong \angle MBD$	Def. of congruent polygons
$\overline{AC} \parallel \overline{DB}$	Def. of alternate interior angles

Problem: As shown, $\overline{AB}$ and $\overline{CD}$ intersect at point M, $\overline{AM} = \overline{BM}$ and $\overline{CM} = \overline{DM}$; then, must $\overline{AC}$ and $\overline{DB}$ be parallel with each other?



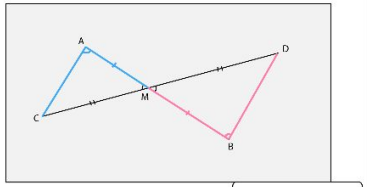
Statements	Reasons
$\overline{AM} \cong \overline{BM}$, $\overline{CM} \cong \overline{DM}$	Given
$\angle AMC \cong \angle BMD$	Def. of vertical angles
$\triangle AMC \cong \triangle BMD$	SAS Triangle Congruency
$\angle MAC \cong \angle MBD$	Def. of congruent polygons
$\overline{AC} \parallel \overline{DB}$	Def. of alternate interior angles

Problem: As shown, $\overline{AB}$ and $\overline{CD}$ intersect at point M, $\overline{AM} = \overline{BM}$ and $\overline{CM} = \overline{DM}$; then, must $\overline{AC}$ and $\overline{DB}$ be parallel with each other?



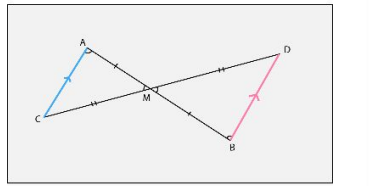
Statements	Reasons
$\overline{AM} \cong \overline{BM}$, $\overline{CM} \cong \overline{DM}$	Given
$\angle AMC \cong \angle BMD$	Def. of vertical angles
$\triangle AMC \cong \triangle BMD$	SAS Triangle Congruency
$\angle MAC \cong \angle MBD$	Def. of congruent polygons
$\overline{AC} \parallel \overline{DB}$	Def. of alternate interior angles

Problem: As shown, $\overline{AB}$ and $\overline{CD}$ intersect at point M, $\overline{AM} = \overline{BM}$ and $\overline{CM} = \overline{DM}$; then, must $\overline{AC}$ and $\overline{DB}$ be parallel with each other?



Statements	Reasons
$\overline{AM} \cong \overline{BM}$, $\overline{CM} \cong \overline{DM}$	Given
$\angle AMC \cong \angle BMD$	Def. of vertical angle
$\triangle AMC \cong \triangle BMD$	SAS Triangle Congru
$\angle MAC \cong \angle MBD$	Def. of congruent polygons
$\overline{AC} \parallel \overline{DB}$	Def. of alternate interior angles

Problem: As shown, $\overline{AB}$ and $\overline{CD}$ intersect at point M, $\overline{AM} = \overline{BM}$ and $\overline{CM} = \overline{DM}$; then, must $\overline{AC}$ and $\overline{DB}$ be parallel with each other?



Statements	Reasons
$\overline{AM} \cong \overline{BM}$, $\overline{CM} \cong \overline{DM}$	Given
$\angle AMC \cong \angle BMD$	Def. of vertical angle
$\triangle AMC \cong \triangle BMD$	SAS Triangle Congru
$\angle MAC \cong \angle MBD$	Def. of congruent po
$\overline{AC} \parallel \overline{DB}$	Def. of alternate interior angles

Methods (4/4)

Method 1: Human-studies **experiment** on American middle or high-school geometry students

- Experiment allows us to measure students' levels of reading comprehension of proofs, aka: how deeply are they understanding and internalizing the proof?
- Adapt method defined by Yang & Lin, defined a test for assessing students' reading comprehension of geometry proofs. Other papers have used or adapted this test for their own experiments on improving comprehension level.
- Use within-subject design, measure difference in scores between students who were given a static proof and students given an interactive version.

Method 2: Interviews with geometry teachers

- Interviews provide additional evidence to support claim that interactive proofs are useful in classroom settings, and collect feedback to make the artifacts more robust.
- Show them the interactive proof and ask them for their thoughts.
- Goal: get answers to RQ 2.1-2.3.

Ian Dardik

Are Compositional TLA+ Specifications Easier to Understand?

Ian Dardik

Are Compositional TLA+ Specifications Easier to Understand?

TLA+ is a formal specification language

- **Problem:** TLA+ specifications are monolithic
- **Potential solution:** Novel composition operator for TLA+
- **This project:** Easier to understand compositional specifications?

Hypothesis: TLA+ users reason about compositional specifications more easily than monolithic specifications

Qualitative Research Question

RQ1: How do TLA+ users **reason about** monolithic/compositional specifications?

Quantitative Research Questions

RQ2: **How much time** does it take for a TLA+ user **to write an invariant** that describes the relationship between particular variables in the monolithic v. the compositional method?

RQ3: **How likely** are TLA+ users **to write a correct invariant** that describes the relationship between particular variables in the monolithic v. the compositional method?

RQ4: How well can TLA+ users **predict whether a property holds** in a specification written in the monolithic v. the compositional method?

Human Subject Study (Design)

Two TLA+ specifications: monolithic and compositional

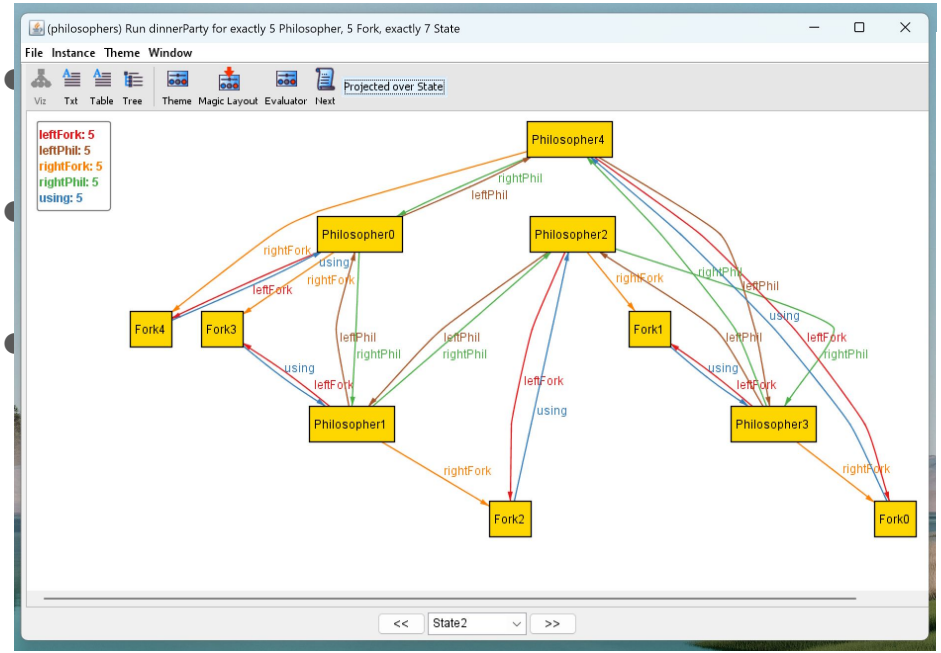
1. [Subject receives one TLA+ specification]
2. Explain what the protocol does (think aloud protocol) **(RQ1)**
3. Write an invariant that describes the relationship between [variables]? (time, check correctness) **(RQ2, RQ3)**
4. Does the specification satisfy [some property] and why? **(RQ4,RQ1)**

Yiliang (Leo) Liang

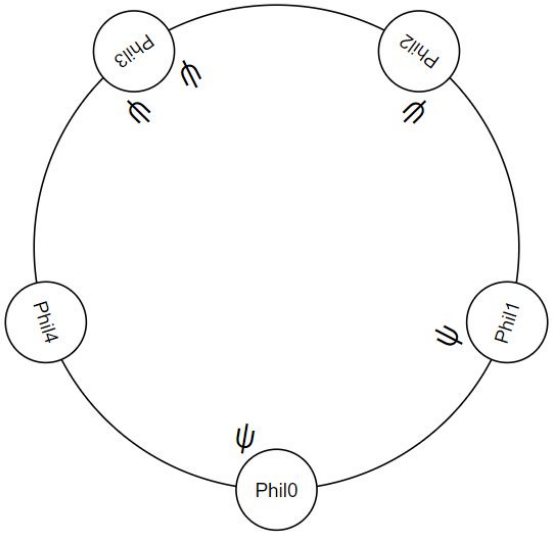
How do formal model developers use visualizations to understand and debug formal models?

Yiliang Liang

Motivation



Showing State6



Research Questions + Methods

RQ1: How do formal model developers currently use visualizations of their models?

RQ2: What kinds of visualizations do these developers want?

Qualitative study: interviews and surveys

RQ3: Do (and to what extent do) domain-specific visualizations help developers understand model behavior? *

Quantitative study: experiments

* may be contingent upon development of domain-specific visualization tool; can use a mockup too

Hao He

Pinning is Futile? On the Impact of Version Constraints in npm Dependency Management.

Hao He

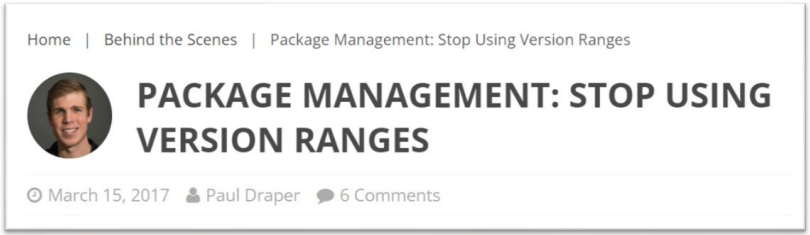
Motivation

- In software development, developers define **dependencies** and their **version constraints** in a configuration file (e.g., package.json in npm)
- Different version constraints have different trade-offs and implications

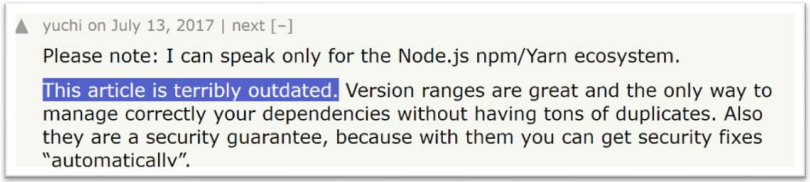
Pinning (e.g., "==1.2.3")	Floating (e.g., "1.2.3 - 2.0.0")
✅No breaking changes	❌Breaking changes
✅No malicious updates	❌Malicious updates
❌No security fixes	✅Gets security fixes automatically
❌More bloat downstream	✅Less bloat downstream

Motivation

- Developers have **contrasting philosophies** on the use of version constraints, but they have no data supporting their philosophies



https://www.lucidchart.com/techblog/2017/03/15/package-management-stop-using-version-ranges/

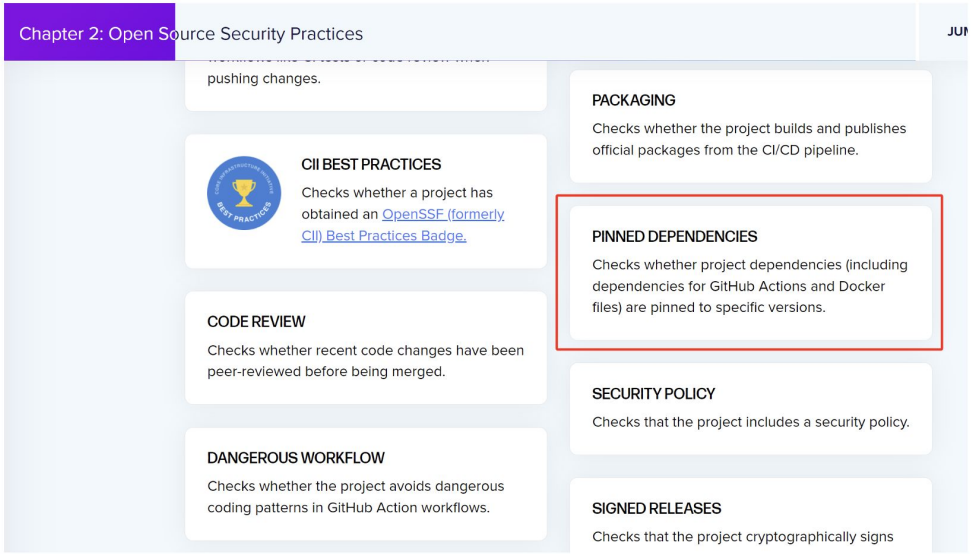


https://news.ycombinator.com/item?id=14759329

https://opensourcesecurity.io/2022/03/21/facts-vs-feelings/

Motivation

- Recent supply chain attacks have “scared” security practitioners, advocating pinning as the best practice to avoid these attacks



Research Questions

- **RQ1: (Descriptive Statistics)** How do developers specify version constraints and how do the version constraints evolve over time in the npm ecosystem?
- **RQ2: (Simulation)** How do different version constraint choices affect the attack surfaces for supply chain attacks and the cost of maintaining dependency graphs?
- **RQ3: (Network Analysis)** Are there any critical points in the npm ecosystem whose intervention can effectively minimize the risk of supply chain attacks?