

# How to maintain software security and license compliance in a world of AI-generated code

---

OSPology LIVE Europe - Malmö 2024

Gary Armstrong

# About me

Who am I, and why am I here to talk to you about software security and license compliance?



Programmer/Project Manager  
Citibank

Senior Director of Operations & Head of Professional Services  
FossID



BEng Computer Science  
Queen's University of Belfast

Senior Manager, Open Source Audits  
Black Duck Audits, Synopsys



Gary  
Armstrong

# A little bit about FossID

*The journey of FossID...and introducing our mascot, Shinobi*



Founded by OSS experts for developers to confidently use OSS and mitigate license compliance and security risks.



2016

2021



Acquired by application security leader, Snyk, in 2021 to enhance scanning capabilities.

Re-acquired by the original founders in 2022 to focus on OSS license challenges.



2022

2024



Introducing our mascot, Shinobi.

Learn more about FossID at [sca.ninja](https://sca.ninja)

# Out of Curiosity

Interactive Quiz



Do your companies currently encourage the use of Generative AI coding tools?

A

Yes

B

Yes, but with certain restrictions

C

No

# Agenda

*What are we going to cover as part of this session?*



## ➔ Introduction

- The Rise of AI-Generated Code
- License Compliance Challenges
  - Assuring License Compliance with AI-Generated Code
- Security Risks in AI-Generated Code
  - Ensuring Security in AI-Generated Code
- Conclusion
  - Wrap-up and Takeaways
- Q&A



# Introduction

How are Generative AI tools revolutionizing software development?



Generative AI tools are transforming how developers create and enhance source code.

These tools seamlessly integrate with Integrated Development Environments (IDEs), becoming an essential part of the development workflow.

Line Autocomplete

A screenshot of the Visual Studio Code editor interface. The search bar at the top contains 'copilot-tests'. The editor shows a file named 'runtime.go' with the following code snippet:

```
21 averageRuntime := float64(totalTime) / float64(len(runs) - failedRuns)
22
23
```

Snippet Suggestion

A screenshot of the Visual Studio Code editor interface. The search bar at the top contains 'copilot-tests'. The editor shows a file named 'index.js' with the following code snippet:

```
1 // Create a clone of YouTube's video player
2 // using the YouTube API
3 // https://developers.google.com/youtube/iframe_api_reference
4
5 // Create a new instance of the YouTube player
6 // and load the video into it
7 var player;
8 function onYouTubeIframeAPIReady() {
9   player = new YT.Player('player', {
10     height: '390',
11     width: '640',
12     videoId: 'M7lc1UVf-VE',
13     events: {
14       'onReady': onPlayerReady,
15       'onStateChange': onPlayerStateChange
16     }
17   });
18 }
19
```

Chat

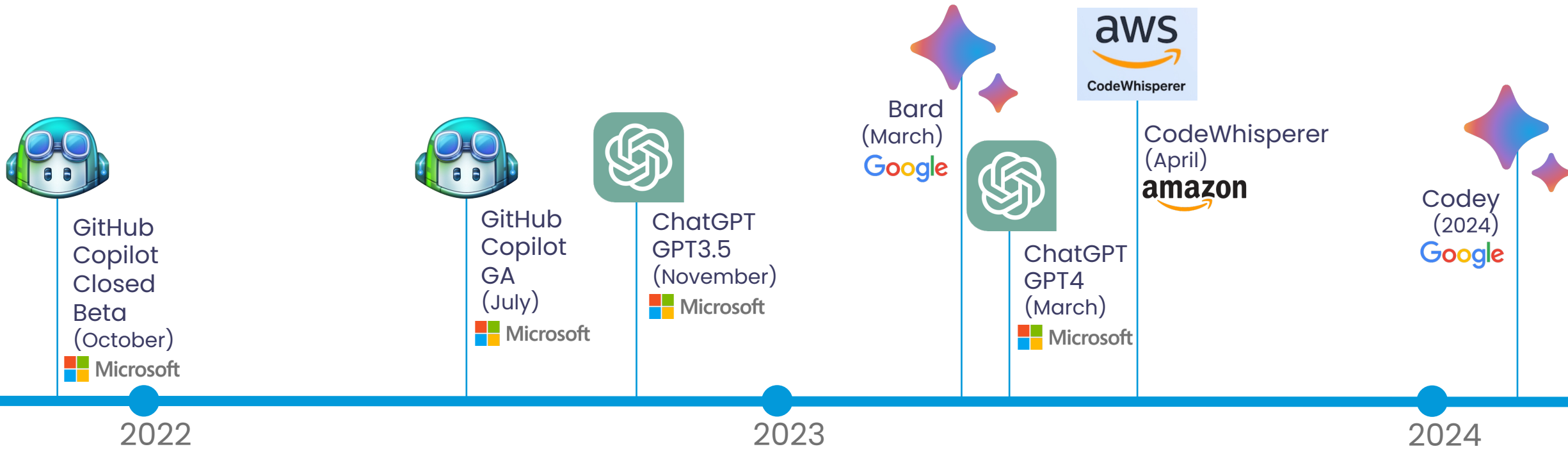
A screenshot of the Visual Studio Code chat interface. The search bar at the top contains 'copilot-tests'. The chat window is titled 'CHAT: GITHUB COPILOT' and shows a welcome message from 'GitHub Copilot' to '@euclidex'. The message reads: 'Welcome @euclidex, I'm your Copilot and I'm here to help you get things done faster. I can identify issues, explain and even improve code. You can ask generic questions, but what I'm really good at is helping you with your code. For example:'. Below the message are several suggested actions: 'Generate unit tests for my code', 'Explain the selected code', 'Propose a fix for the bugs in my code', and 'What are some other parameters that can be set using the playerVars object?'. At the bottom, there is a text input field with the placeholder 'Ask Copilot or type '/' for commands' and a send button.

# Introduction



How have Generative AI coding tools been developing over the last number of years?

Between 2021 and 2024, Microsoft (OpenAI), Google and Amazon have released their competing generative AI coding engines:



# Introduction

What are the benefits and challenges of using Generative AI



## Benefits

### Increased Productivity

- AI tools assist developers by automating repetitive tasks and providing intelligent suggestions.

### Error Reduction

- They help minimize coding errors through real-time feedback and code validation.

### Improved Code Quality

- AI-powered tools enhance code quality by suggesting best practices and optimizing code structures.

### Enhanced Collaboration

- AI tools facilitate better collaboration among team members by providing consistent code standards and shared coding practices.

## Challenges

### Over-Reliance on AI

- Developers might become overly dependent on AI tools, potentially losing critical coding skills.

### Bias in AI Algorithms

- It's important to be aware that AI tools can perpetuate existing biases present in their training data, potentially leading to biased outputs.

### Security Risks

- AI-generated code might introduce security vulnerabilities if not properly reviewed and tested.

### License Compliance

- There can be issues with license compliance, as AI-generated code might inadvertently incorporate copyrighted material or violate open-source licenses.



# Agenda

*What's next on the agenda?*

- Introduction
  - The Rise of AI-Generated Code
- ➔ License Compliance Challenges
  - Assuring License Compliance with AI-Generated Code
- Security Risks in AI-Generated Code
  - Ensuring Security in AI-Generated Code
- Conclusion
  - Wrap-up and Takeaways
- Q&A



# License Compliance Challenges

Interactive Quiz



Does copyright law protect software works?

A  
Yes

B  
No

C  
Depends on  
jurisdiction

# License Compliance Challenges



What is the definition of a copyright?

Legal right that grants the creator of an **original work**



*literary works*



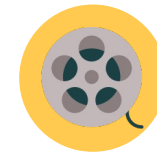
*music*



*photography*



*paintings*



*motion pictures*



*software*

**exclusive rights** to its use and distribution



*reproduce*



*perform*



*modify*



*distribute*

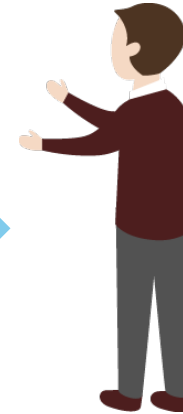
with the intention of enabling the creator to receive a **compensation**  
for a **limited period of time**

# License Compliance Challenges

What is the definition of a copyright?



Matthew



*Play it only on your radio channel*



*Pay me every time you play it*

Stephen



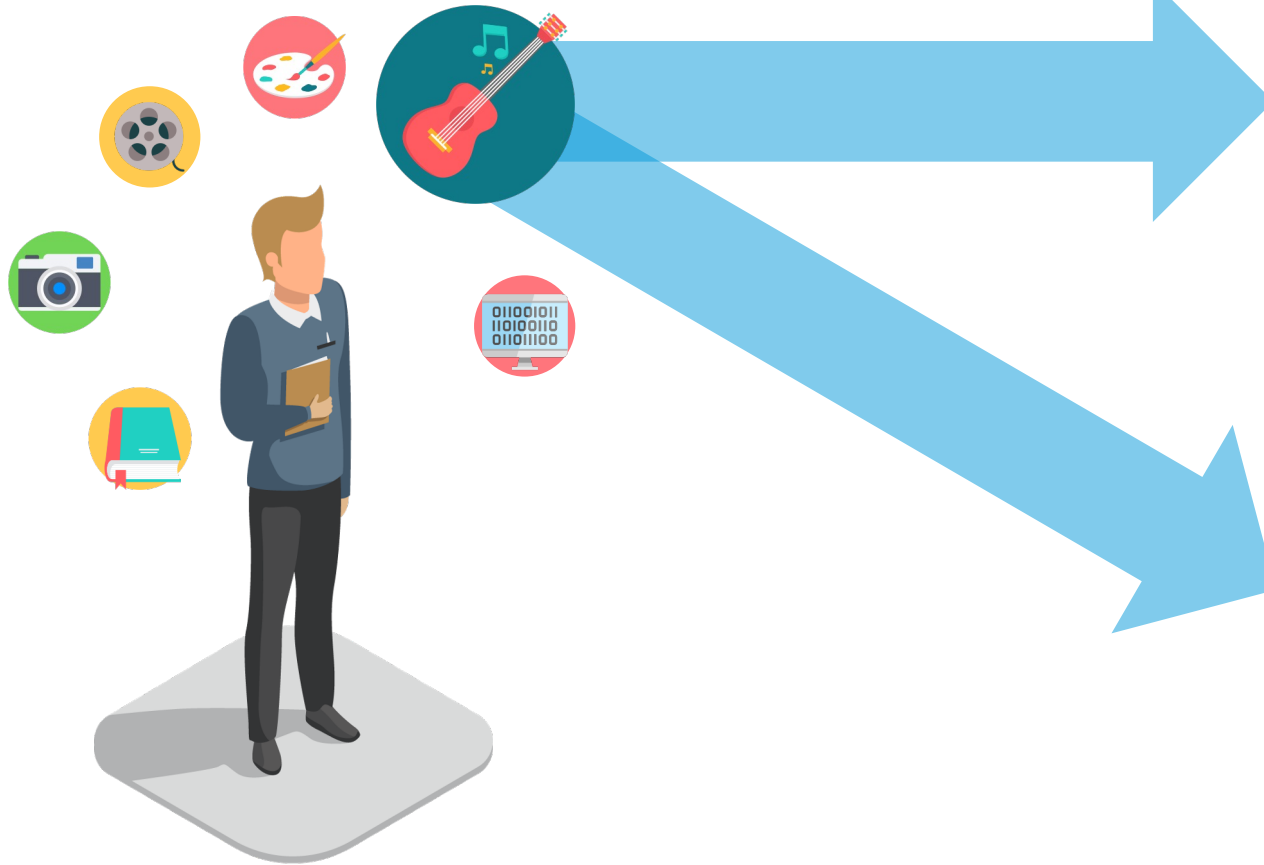
*Perform in any way you want*



*Play it anytime anywhere*



*Always credit me*



# License Compliance Challenges

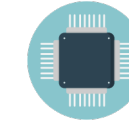
What is the definition of a copyright?



Matthew



Use it only in video games



Run it only on Intel processors



Pay me every time you sell it within your product

Stephen



Run it for any purpose



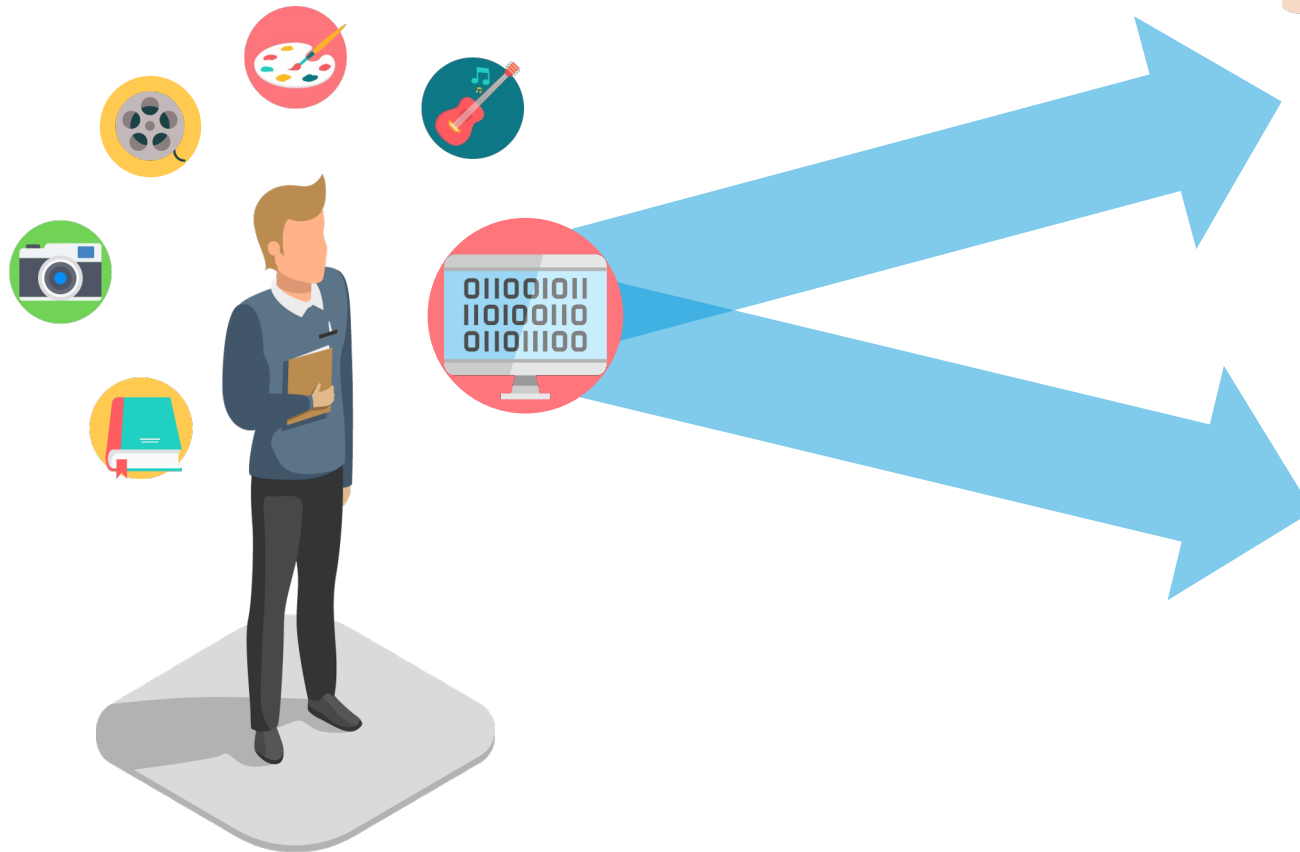
Free to modify it



Free to distribute it



Always credit the author



# License Compliance Challenges

Interactive Quiz



Which one of the following works is NOT recognized by copyright law?

A

A book written by a member of ABBA

B

A python script written by me

C

A drawing made by a monkey

# License Compliance Challenges



Interactive Quiz - Answer

Most jurisdictions, including Spain and Germany, state that **only works created by a human** can be protected by copyright.

Tuesday, November 17, 1987

## 25628 Law 22/1987 of November 11, 1987

JUAN CARLOS I  
KING OF SPAIN

All you present and understand:

Sabed: That the General Court has adopted and I have adopted the following Law:

### PREAMBLE

The need to establish a specific legal regime for intellectual property rights as a result of the enactment of the Law of January 10, 1879, value and opportunity which has so far contributed to the development of our own legal tradition.

However, the legislature of then could not provide for the profound social transformation that has occurred and, in particular, the consequences of the development of the means of dissemination of the creation works that have allowed, for the first time in history, access for most citizens to culture, but that in parallel, they have provided new modes of defrauding intellectual property rights.

These adaptation needs to the new circumstances have only been partially addressed in the internal legal system, by means of the adoption of various specific rules concerning the protection of rights in certain works, and at the international level, through conventions, in some of which Spain is a party, but without our domestic legislation being adapted to the extent necessary.

Therefore, taking into consideration the preponderant trends in the member countries of the European Community and, in particular, those of those closest to our legal tradition, a unitary or systematic nature, a new intellectual property legal regime, which aims to make the rights in works of creation real, established and effectively recognized and protected in accordance with the requirements of our time.

Within this framework, the Law is structured on two distinct sets of policy sets: the declaration of substantive rights and the regulator of the actions and procedures for the protection of such rights.

In turn, within the first set of regulations, the rights to the author are determined, on the one hand, which is the person who is carrying out the purely human task and the personal creation of the work and which, constitute the essential core of the subject matter of this Law and, on the other hand, the rights recognized to certain natural or legal persons whose intervention is indispensable to the performance or for the dissemination of such works created by the authors.

\*[https://www.wipo.int/wipo\\_magazine/en/2017/05/article\\_0003.html](https://www.wipo.int/wipo_magazine/en/2017/05/article_0003.html)

# License Compliance Challenges

*What is the future of code repositories?*



Code repositories will soon become a mix of human and AI generated code.

What does that mean in terms of copyright ownership?



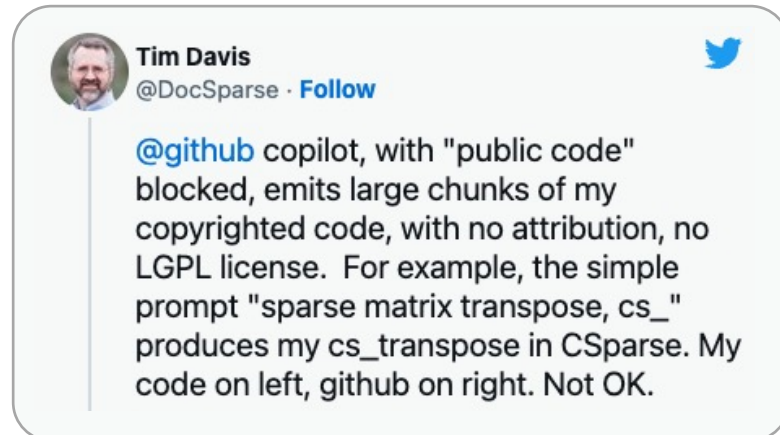


# License Compliance Challenges



*The lawsuit that could rewrite the rules of AI copyright*

- (July 2022) After its GA release, GitHub Copilot quickly became the most popular AI-driven code generation software development tool in the world.



- GitHub Copilot is called out numerous times for using copyrighted code without neither attribution nor consent.
- (October 2022) Tim Davis, Computer Science Professor at Texas University, shows a clear example of Copilot using his code without complying with the LGPL license

- (November 2022) Microsoft/GitHub/OpenAI are being sued for allegedly violating copyright law by reproducing open-source code using AI.

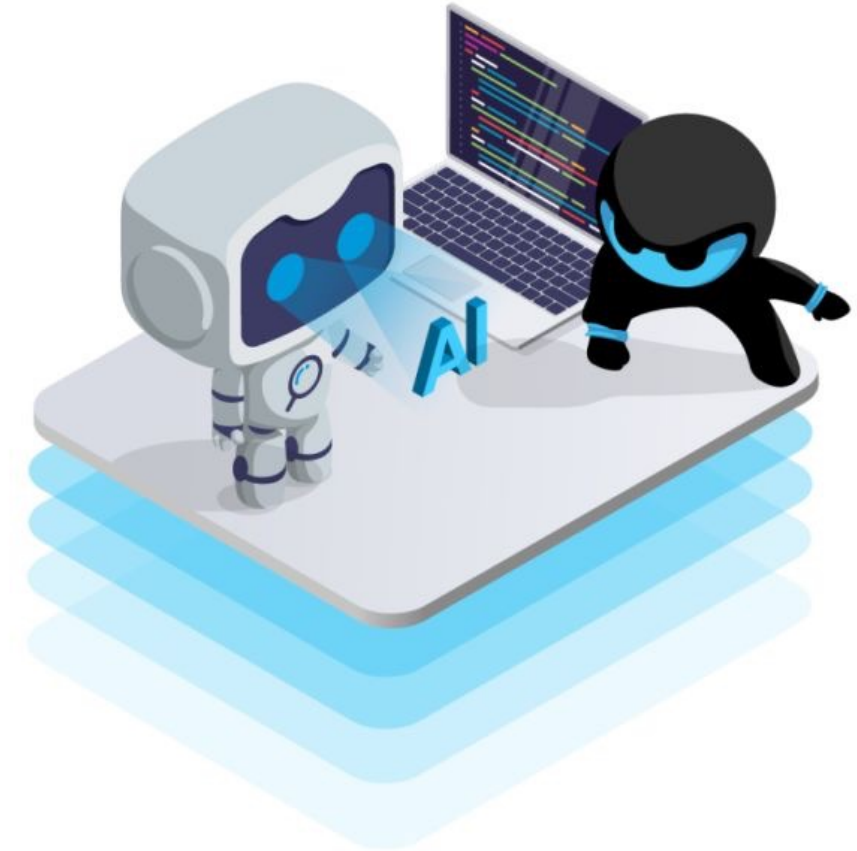
<https://www.theverge.com/2022/11/8/23446821/microsoft-openai-github-copilot-class-action-lawsuit-ai-copyright-violation-training-data>

# License Compliance Challenges

*How accurate are these claims?*



Are these claims true?



Let's run a few tests!!

# Your AI pair programmer

GitHub Copilot uses the OpenAI Codex to suggest code and entire functions in real-time, right from your editor.

Start my free trial >

Explore docs

```
sentiments.ts write_sql.go parse_expenses.py addresses.rb
1 #!/usr/bin/env ts-node
2
3 import { fetch } from "fetch-h2";
4
5 // Determine whether the sentiment of text is positive
6 // Use a web service
7 async function isPositive(text: string): Promise<boolean> {
8   const response = await fetch(`http://text-processing.com/api/sentiment/`, {
9     method: "POST",
10    body: `text=${text}`,
11    headers: {
12      "Content-Type": "application/x-www-form-urlencoded",
13    },
14  });
15  const json = await response.json();
16  return json.label === "pos";
17 }
```

Copilot

Replay

Trained on billions of lines of code, GitHub Copilot turns natural language prompts into coding suggestions across dozens of languages.

SNIPPET

```
#!/usr/bin/env ts-node
import { fetch } from "fetch-h2";
// Determine whether the sentiment of text is positive
// Use a web service
async function isPositive(text: string): Promise<boolean> {
  const response = await fetch('http://text-processing.com/api/sentiment/', {
    method: "POST",
    body: `text=${text}`,
    headers: {
      "Content-Type": "application/x-www-form-urlencoded",
    },
  });
  const json = await response.json();
  return json.label === "pos";
}
```

GO!

FossilID match

## SNIPPET

```
#!/usr/bin/env ts-node
import { fetch } from "fetch-h2";
// Determine whether the sentiment of text is positive
// Use a web service
async function isPositive(text: string): Promise<boolean> {
const response = await fetch('http://text-processing.com/api/sentiment/', {
  method: "POST",
  body: `text=${text}`,
  headers: {
    "Content-Type": "application/x-www-form-urlencoded",
  },
});
const json = await response.json();
return json.label === "pos";
}
```

GO!

## FossilID partialmatch

```
140 #!/usr/bin/env ts-node
141
142 import { fetch } from "fetch-h2";
143
144 // Determine whether the sentiment of text is positive
145 // Use a web service
146
147 async function isPositive(text: string): Promise<boolean> {
148
149   const response = await fetch('http://text-processing.com/api/sentiment/', {
150     method: "POST",
151     body: `text=${text}`,
152
153     headers: {
154       "Content-Type": "application/x-www-form-urlencoded",
155     },
156   });
157
158   const json = await response.json();
159   return json.label === "pos";
160 }
161
```

Showing 1 to 3 of 3 entries

MATCH	ARTIFACT	VERSION	AUTHOR	COMPONENT LICENSE	RELEASE DATE	FILE LICENSE			
partial	DelphiVCL4Python	91568fa611f6e6fd07e319c758420033d8db53dd	maxkleiner	PublicDomain	2022-11-05	samples/1151_sentiment_api4simple.txt	12kb	🔄	15 (100%)
partial	agsi-data	1f30dd2d0f8c52e839fadd4da70430ced40080a5	maxkleiner	N/A	2022-10-10	1155_census_data.txt	19kb	🔄	15 (100%)
partial	eriknicole-20132017	46ac8fc7c48968b8d8d6f7d23c721062dd9203eb	eriknicole-20132017	N/A	2022-11-29	sentiments.ts	475bytes	🔄	15 (100%)

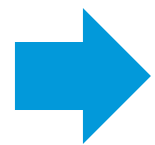
<https://github.com/maxkleiner/DelphiVCL4Python/archive/91568fa611f6e6fd07e319c758420033d8db53dd.tar.gz>
[Take me to DelphiVCL4Python maxkleiner homepage.](#)
[Download DelphiVCL4Python maxkleiner source code.](#)

Show 10 entries

Previous 1 Next

## Human oversight

- + Can GitHub Copilot introduce insecure code in its suggestions?
- + Does GitHub own the code generated by GitHub Copilot?
- + Does GitHub Copilot copy code from the training set?
- **What can I do to reduce GitHub Copilot's suggestion of code that matches public code?**



We built a [filter](#) to help detect and suppress GitHub Copilot suggestions which contain code that matches public code on GitHub.

Copilot for Individual users have the choice to enable that filter during setup on their individual accounts. For Copilot for Business users, the Enterprise administrator controls how the filter is applied. They can control suggestions for all organizations or defer control to individual organization administrators. These organization administrators can turn the filter on or off during setup (assuming their Enterprise administrator has deferred control) for the users in their organization.

With the filter enabled, GitHub Copilot checks code suggestions with its surrounding code for matches or near matches (ignoring whitespace) against public code on GitHub of about 150 characters. If there is a match, the suggestion will not be shown to you. In addition, we have [announced](#) that we are building a feature that will provide a reference for suggestions that resemble public code on GitHub so that you can make a more informed decision about whether and how to use that code, as well as explore and learn how that code is used in other projects.

Just like when you write any code that uses material you did not independently originate, you should take precautions to understand how it works and ensure its suitability. These include rigorous testing, IP scanning, and checking for security vulnerabilities. You should make sure your IDE or editor does not automatically compile or run generated code before you review it.

- Public profile
- Account
- Appearance
- Accessibility
- Notifications

Access

- Billing and plans
- Emails
- Password and authentication
- Sessions
- SSH and GPG keys
- Organizations
- Moderation

Code, planning, and automation

- Repositories
- Codespaces
- Packages
- Copilot**
- Pages
- Saved replies

# GitHub Copilot

Thank you for participating in the GitHub Copilot Technical Preview. Your extended access has ended.

## Suggestions matching public code \*

GitHub Copilot can allow or block suggestions matching public code. See [GitHub Copilot FAQ](#) to learn more.

Block

Allow GitHub to use my code snippets for product improvements \*

Allow GitHub, its affiliates and third parties to use my code snippets to research and improve GitHub Copilot suggestions, related models and product features. More information in [Privacy FAQ](#).

Save

It can take up to 30 minutes for the changes to take effect. Restart your code editor for the changes to take effect immediately.

# Your AI pair programmer

GitHub Copilot uses the OpenAI Codex to suggest code and entire functions in real-time, right from your editor.

Start my free trial >

Explore docs



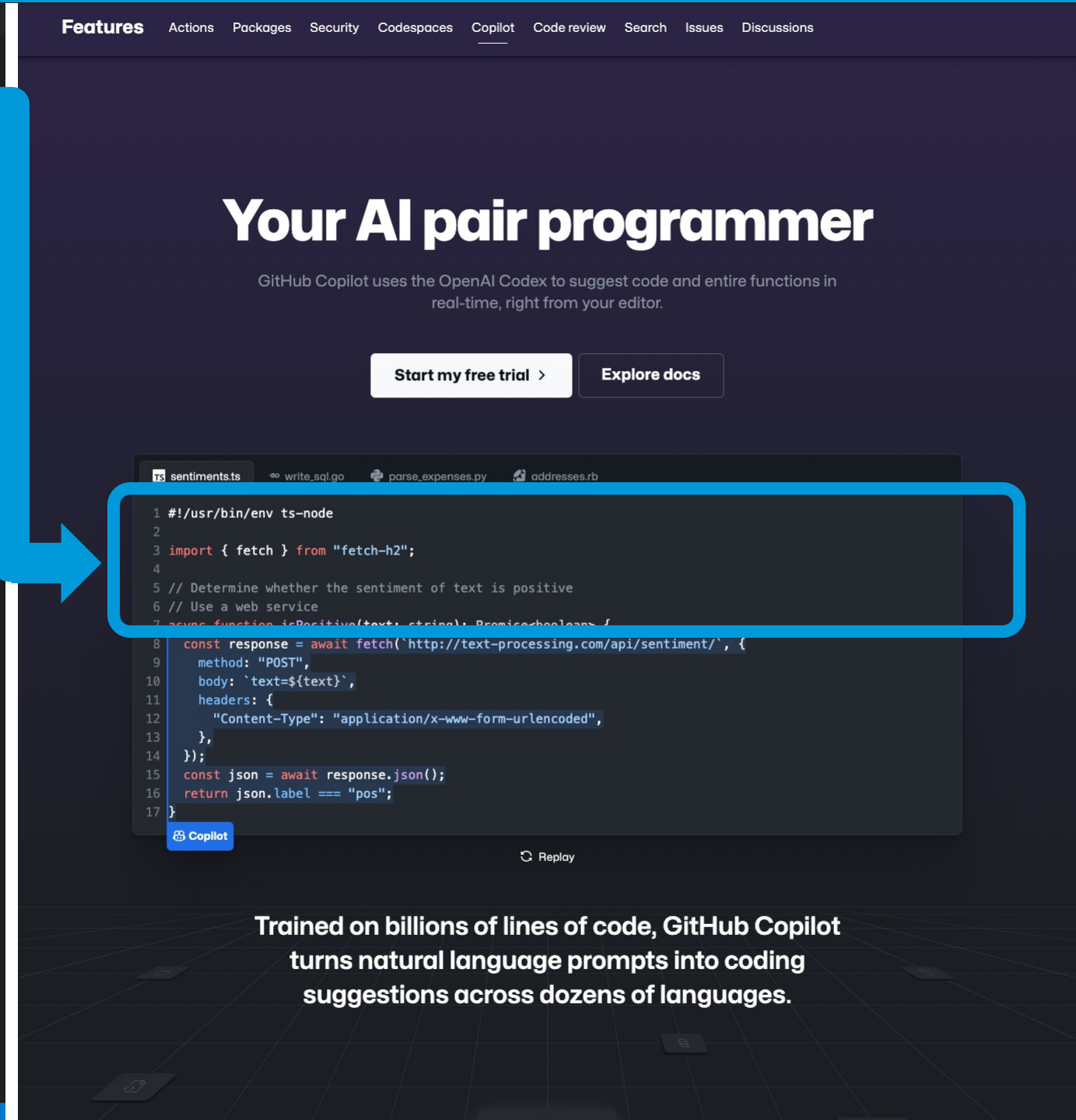
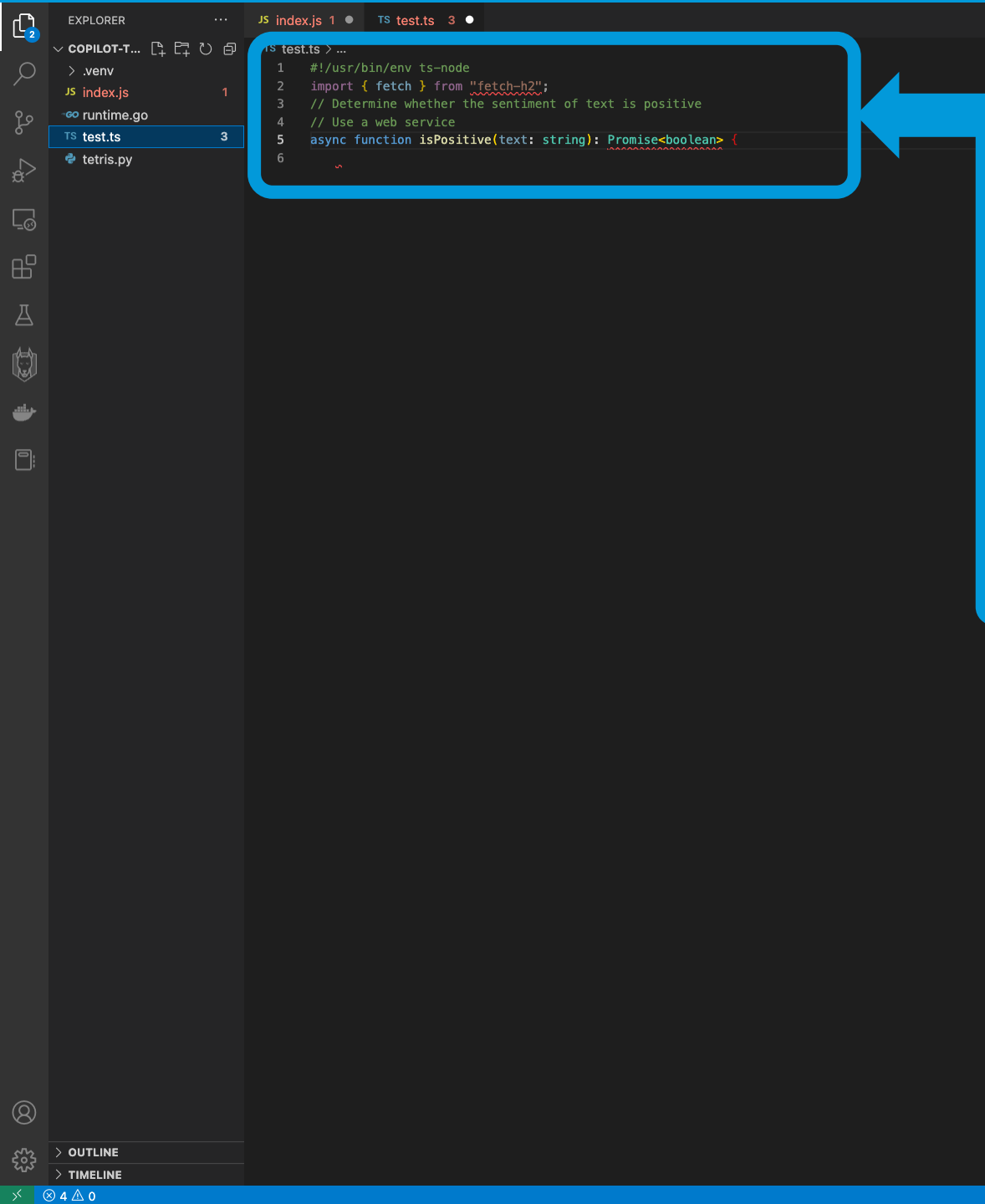
```
sentiments.ts write_sql.go parse_expenses.py addresses.rb  
1 #!/usr/bin/env ts-node  
2  
3 import { fetch } from "fetch-h2";  
4  
5 // Determine whether the sentiment of text is positive  
6 // Use a web service  
7 async function isPositive(text: string): Promise {  
8   const response = await fetch('http://text-processing.com/api/sentiment/', {  
9     method: "POST",  
10    body: `text=${text}`,  
11    headers: {  
12      "Content-Type": "application/x-www-form-urlencoded",  
13    },  
14  });  
15  const json = await response.json();  
16  return json.label === "pos";  
17 }
```

Copilot

Replay

Trained on billions of lines of code, GitHub Copilot turns natural language prompts into coding suggestions across dozens of languages.





Trained on billions of lines of code, GitHub Copilot turns natural language prompts into coding suggestions across dozens of languages.

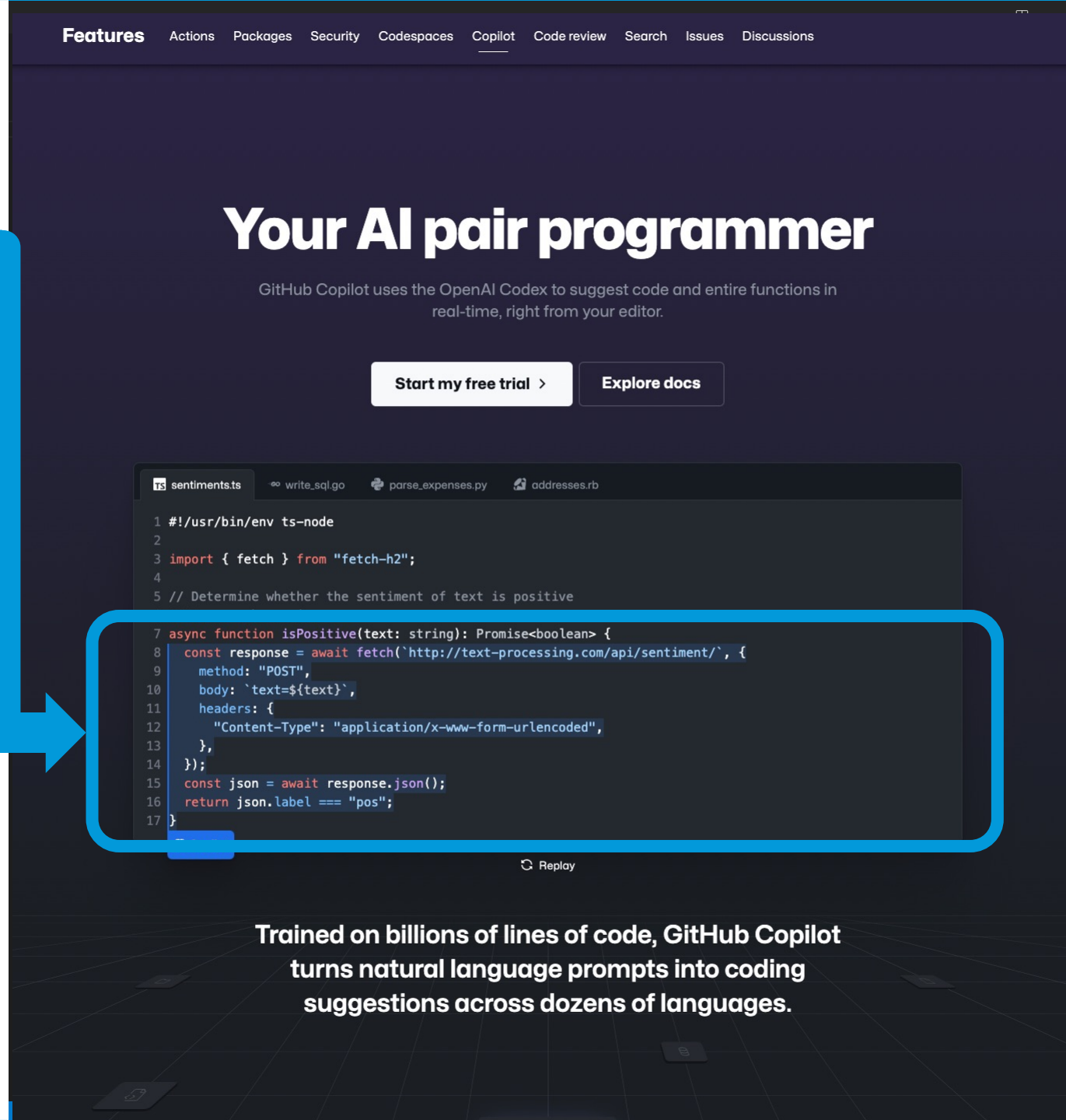
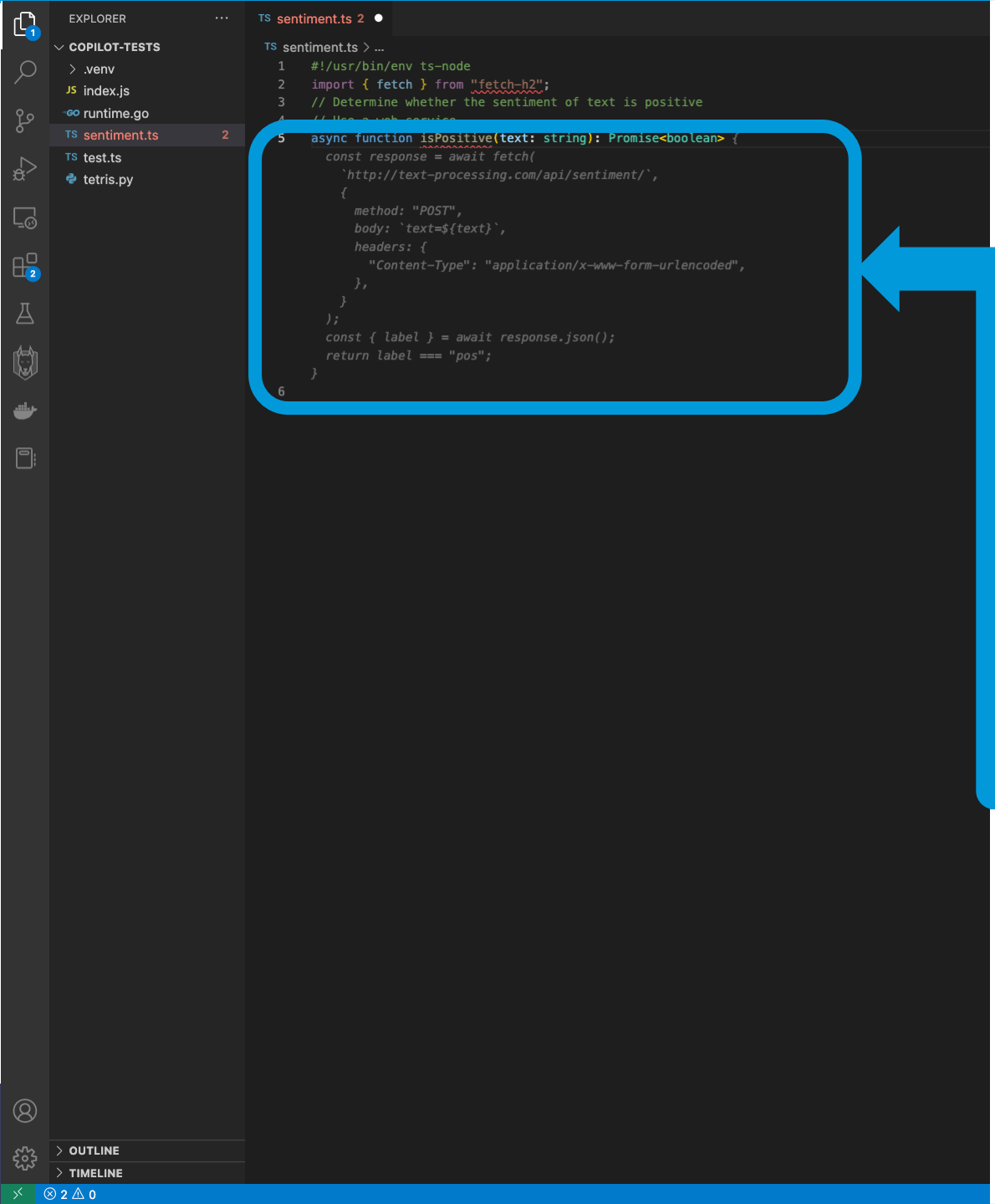
EXPLORER

- COPILOT-TESTS
- .venv
- index.js
- runtime.go
- sentiment.ts 2
- test.ts
- tetris.py

```
TS sentiment.ts > ...
1  #!/usr/bin/env ts-node
2  import { fetch } from "fetch-h2";
3  // Determine whether the sentiment of text is positive
4  // Use a web service
5  async function isPositive(text: string): Promise<boolean> {
    const response = await fetch(
      `http://text-processing.com/api/sentiment/`,
      {
        method: "POST",
        body: `text=${text}`,
        headers: {
          "Content-Type": "application/x-www-form-urlencoded",
        },
      }
    );
    const { label } = await response.json();
    return label === "pos";
  }
6
```

> OUTLINE

> TIMELINE



# License Compliance Challenges

*How about code outside of GitHub?*



Copilot does not seem to filter publicly available code in GitHub.

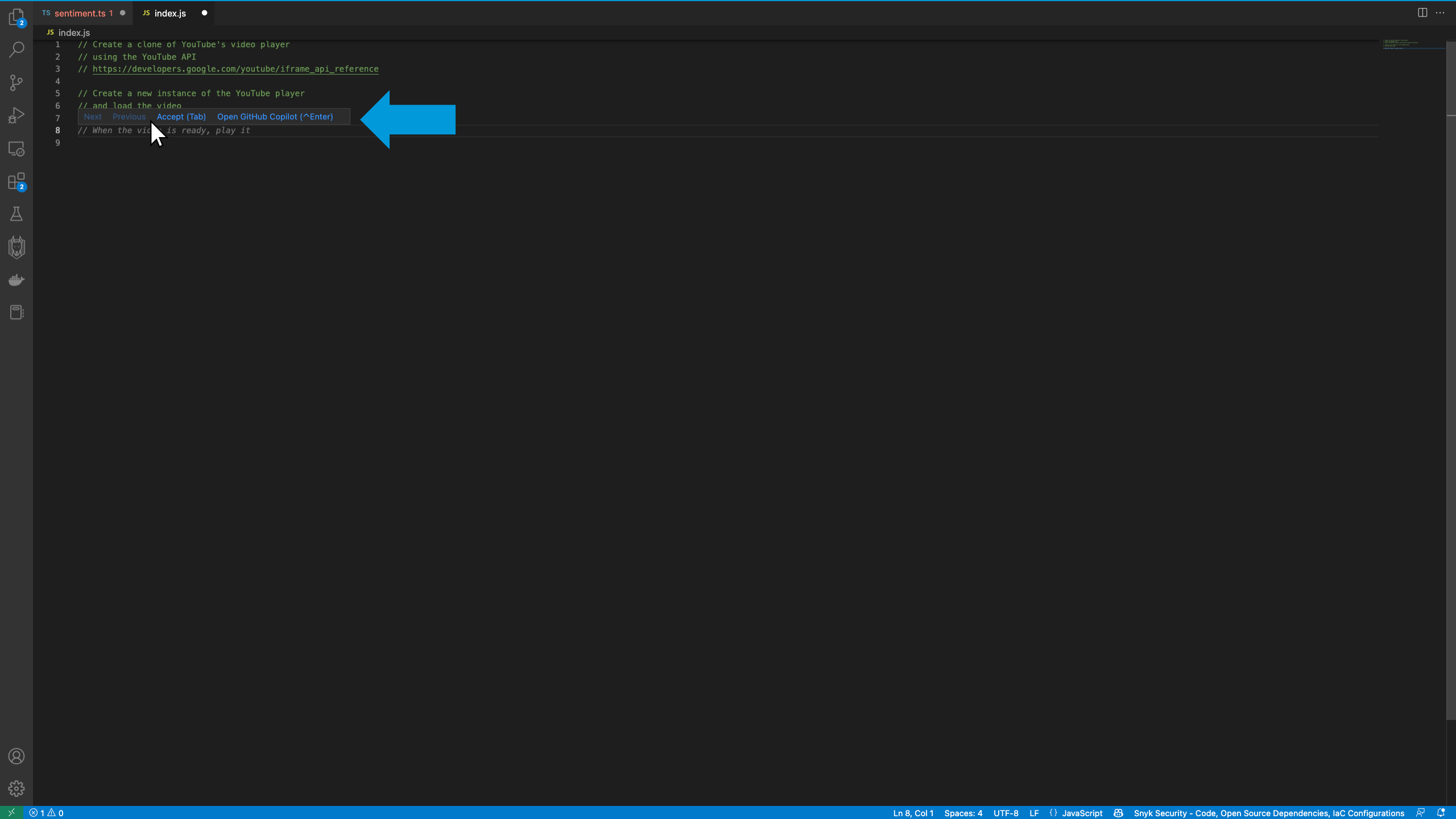
...but how about code outside GitHub?



Let's run another test!!



```
JS index.js
1 // Create a clone of YouTube's video player
2 // using the YouTube API
3 // https://developers.google.com/youtube/iframe_api_reference
4
5 // Create a new instance of the YouTube player
6 // and load the video
7
```



```
JS index.js
1 // Create a clone of YouTube's video player
2 // using the YouTube API
3 // https://developers.google.com/youtube/iframe_api_reference
4
5 // Create a new instance of the YouTube player
6 // and load the video
7 Next Previous Accept (Tab) Open GitHub Copilot (^Enter)
8 // When the video is ready, play it
9
```



```
JS index.js
1 // Create a clone of YouTube's video player
2 // using the YouTube API
3 // https://developers.google.com/youtube/iframe_api_reference
4
5 // Create a new instance of the YouTube player
6 // and load the video
7
```



```
1 Synthesizing 9/10 solutions
2
3 =====
4
5 Accept Solution
6 // and load the video into it
7 function onYouTubeIframeAPIReady() {
8   player = new YT.Player('player', {
9     height: '390',
10    width: '640',
11    videoId: 'M7lc1UVf-VE',
12    events: {
13      'onReady': onPlayerReady,
14      'onStateChange': onPlayerStateChange
15    }
16  });
17
18 // When the video player is ready, play the video
19 function onPlayerReady(event) {
20   event.target.playVideo();
21 }
22
23 // When the video player's state changes,
24 // log the event to the console
25 function onPlayerStateChange(event) {
26   console.log(event);
27 }
28
29 // When the page loads, load the YouTube IFrame Player API code asynchronously
30 function loadScript() {
31   var tag = document.createElement('script');
32   tag.src = "https://www.youtube.com/iframe_api";
33   var firstScriptTag = document.getElementsByTagName('script')[0];
34   firstScriptTag.parentNode.insertBefore(tag, firstScriptTag);
35 }
36 window.onload = loadScript;
37
38 // Path: index.html
39 // Create a div with an id of 'player'
40 // to hold the video player
41 <div id="player"></div>
42
43 // Path: index.css
44 // Style the player to fill the page
45 #player {
46   position: fixed;
47   top: 0;
48   left: 0;
49   height: 100%;
50   width: 100%;
51 }
52
53 // Path: index.html
54 // Create a button to play/pause the video
55 <button onClick="playPause()">Play/Pause</button>
56
57 // Path: index.js
58 // Play/pause the video when the button is clicked
59 function playPause() {
60   var state = player.getPlayerState();
61   if (state == 1) {
62     player.pauseVideo();
63   } else {
```

```
JS index.js 1
1 // Create a clone of YouTube's video player
2 // using the YouTube API
3 // https://developers.google.com/youtube/iframe_api_reference
4
5 // Create a new instance of the YouTube player
6 // and load the video
7 function onYouTubeIframeAPIReady() {
8   player = new YT.Player('player', {
9     height: '390',
10    width: '640',
11    videoId: 'M7lc1UVf-VE',
12    events: {
13      'onReady': onPlayerReady,
14      'onStateChange': onPlayerStateChange
15    }
16  });
17 }
18
19 // When the video player is ready, start the video
20 function onPlayerReady(event) {
21   event.target.playVideo();
22 }
23
24 // When the video player's state changes,
25 // check if the video has ended
26 function onPlayerStateChange(event) {
27   if (event.data == YT.PlayerState.ENDED) {
28     // The video has ended
29   }
30 }
31
32 // When the YouTube API is ready, start the video
33 function onYouTubeIframeAPIReady() {
34   player = new YT.Player('player', {
35     height: '390',
36     width: '640',
37     videoId: 'M7lc1UVf-VE',
38     events: {
39       'onReady': onPlayerReady,
40       'onStateChange': onPlayerStateChange
41     }
42   });
43 }
44
45 // When the video player is ready, start the video
46 function onPlayerReady(event) {
47   event.target.playVideo();
48 }
49
50 // When the video player's state changes,
51 // check if the video has ended
52 function onPlayerStateChange(event) {
53   if (event.data == YT.PlayerState.ENDED) {
54     // The video has ended
55   }
56 }
57
58 // Path: index.html
59 // Create a clone of YouTube's video player
60 // using the YouTube API
61 // https://developers.google.com/youtube/iframe_api_reference
62
```





JA



10

## SNIPPET

```
// and load the video with the given ID
function onYouTubeIframeAPIReady() {
  player = new YT.Player('player', {
    height: '390',
    width: '640',
    videoid: 'M7lc1UVf-VE',
    events: {
      'onReady': onPlayerReady,
      'onStateChange': onPlayerStateChange
    }
  });
}

function onPlayerReady(event) {
  // Play the video when the player is ready
  event.target.playVideo();
}

// When the player's state changes, log the new state
function onPlayerStateChange(event) {
  console.log(event.data);
}

// Load the YouTube Iframe Player API code asynchronously.
var tag = document.createElement('script');
tag.src = "https://www.youtube.com/iframe_api";
```

GO!

## FossID match

SNIPPET

```

$(document).ready(resizeIframe);

// This code loads the IFrame Player API code asynchronously.
var tag = document.createElement('script');
tag.src = "https://www.youtube.com/iframe_api";
var firstScriptTag = document.getElementsByTagName('script')[0];
firstScriptTag.parentNode.insertBefore(tag, firstScriptTag);

// This function creates an <iframe> (and YouTube player)
// after the API code downloads.
var player;
function onYouTubeIframeAPIReady() {
  player = new YT.Player('player', {
    height: '390',
    width: '640',
    videoid: 'M7lc1UVf-VE',
    events: {
      'onReady': onPlayerReady,
      'onStateChange': onPlayerStateChange
    }
  });
}

// The API will call this function when the video player is ready.
function onPlayerReady(event) {
  event.target.playVideo();
}

```

GO!

FossilID partialmatch



```

1 // This code loads the IFrame Player API code asynchronously.
2 var tag = document.createElement('script');
3
4 tag.src = "https://www.youtube.com/iframe_api";
5 var firstScriptTag = document.getElementsByTagName('script')[0];
6 firstScriptTag.parentNode.insertBefore(tag, firstScriptTag);
7
8 // This function creates an <iframe> (and YouTube player)
9 // after the API code downloads.
10 var player;
11 function onYouTubeIframeAPIReady() {
12   player = new YT.Player('player', {
13     height: '390',
14     width: '640',
15     videoId: 'M7lc1UVf-VE',
16     events: {
17       'onReady': onPlayerReady,
18       'onStateChange': onPlayerStateChange
19     }
20   });
21 }
22

```

Showing 1 to 10 of 10 entries

MATCH	ARTIFACT	VERSION	AUTHOR	COMPONENT LICENSE	RELEASE DATE	FILE LICENSE	FILE	SIZE	URL	HITS
partial	laravel-room	0da3e25463291cb5bc41fc00a9fdec4c7581b68c	GOD-oda	MIT	2018-09-12		public/js/youtube.js	1kb		29 (58%)
partial	26002902	1	mpgn	CC-BY-SA-3.0	2014-09-23		26002902_1	1kb		22 (44%)
partial	61765191	1	Hector	CC-BY-SA-4.0	2020-05-13		61765191_1	977bytes		22 (44%)
partial	60582210	1	kmeng01	CC-BY-SA-4.0	2020-03-07		60582210_1	1kb		22 (44%)
partial	51150561	1	noogui	CC-BY-SA-4.0	2018-07-03		51150561_1	2kb		22 (44%)

[Home](#)

PUBLIC

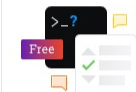
[Questions](#)[Tags](#)[Users](#)[Companies](#)

COLLECTIVES

[Explore Collectives](#)

TEAMS

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

[Create a free Team](#)[Why Teams?](#)[Return to Revisions](#)

1 of 1

answered Sep 23, 2014 at 19:11



mpgn

7.1k ● 9 ● 66 ● 100



A sample to autoplay without using arg `autoplay`. I only test this solution on PCs not mobile but you can try this code :

[LIVE EXAMPLE](#)

## HTML

```
<div id="player"></div>
```

## JS

```
var tag = document.createElement('script');

tag.src = "https://www.youtube.com/iframe_api";
var firstScriptTag = document.getElementsByTagName('script')[0];
firstScriptTag.parentNode.insertBefore(tag, firstScriptTag);

var tag = document.createElement('script');

tag.src = "https://www.youtube.com/iframe_api";
var firstScriptTag = document.getElementsByTagName('script')[0];
firstScriptTag.parentNode.insertBefore(tag, firstScriptTag);

var player;
function onYouTubeIframeAPIReady() {
  player = new YT.Player('player', {
    height: '390',
    width: '640',
    videoId: 'M7lc1UVf-VE',
    events: {
      'onReady': onPlayerReady,
      'onStateChange': onPlayerStateChange
    }
  });
}

function onPlayerReady(event) {
  event.target.playVideo();
}

function onPlayerStateChange(event) {
  if (event.data == YT.PlayerState.PLAYING) {
```

answered Sep 23, 2014 at 19:11



mpgn

7.1k ● 9 ● 66 ● 100



## STACK OVERFLOW

[Questions](#)  
[Help](#)

## PRODUCTS

[Teams](#)  
[Advertising](#)  
[Collectives](#)  
[Talent](#)

## COMPANY

[About](#)  
[Press](#)  
[Work Here](#)  
[Legal](#)  
[Privacy Policy](#)  
[Terms of Service](#)  
[Contact Us](#)  
[Cookie Settings](#)  
[Cookie Policy](#)

## STACK EXCHANGE NETWORK

[Technology](#)  
[Culture & recreation](#)  
[Life & arts](#)  
[Science](#)  
[Professional](#)  
[Business](#)  
[API](#)  
[Data](#)[Blog](#) [Facebook](#) [Twitter](#) [LinkedIn](#) [Instagram](#)

# License Compliance Challenges

*How do other tools compare?*



GitHub Copilot is using code from both inside and outside GitHub as training data.



Let us run a test with ChatGPT!!

+ New Chat



Default (GPT-3.5)




Today

Tic Tac Toe Game



Write a tic tac toe game in python

 Upgrade to Plus



Jon Aldama



+ New Chat



Default (GPT-3.5)



Today

Tic Tac Toe Game



Write a tic tac toe game in python



Here's a simple implementation of a Tic Tac Toe game in Python. This game allows two players to take turns to play Tic Tac Toe on the console.



python

Copy code

```
def print_board(board):
    for row in board:
        print(" | ".join(row))
        print("-" * 9)

def check_winner(board, player):
    for row in board:
        if all(cell == player for cell in row):
            return True

    for col in range(3):
        if all(board[row][col] == player for row in range(3)):
            return True

    if all(board[i][i] == player for i in range(3)) or all(board[i][2 - i] == player for i in range(3)):
        return True

    return False

def is_board_full(board):
    return all(cell != " " for row in board for cell in row)

def main():
    board = [ [" " for _ in range(3)] for _ in range(3)]
```

Regenerate

Send a message



Upgrade to Plus



Jon Aldama



## SNIPPET

```
col = int(input(f"Player {current_player}, enter column (0, 1, 2): "))

if row < 0 or row > 2 or col < 0 or col > 2 or board[row][col] != " ":
    print("Invalid move. Try again.")
    continue

board[row][col] = current_player

if check_winner(board, current_player):
    winner = current_player
else:
    current_player = "O" if current_player == "X" else "X"

print_board(board)

if winner:
    print(f"Player {winner} wins!")
else:
    print("It's a draw!")

if __name__ == "__main__":
    main()
```

GO!

## FossID match



Limit: 10



Sensitivity: 10



## SNIPPET

```
col = int(input(f"Player {current_player}, enter column (0, 1, 2): "))

if row < 0 or row > 2 or col < 0 or col > 2 or board[row][col] != " ":
    print("Invalid move. Try again.")
    continue

board[row][col] = current_player

if check_winner(board, current_player):
    winner = current_player
else:
    current_player = "O" if current_player == "X" else "X"

print_board(board)

if winner:
    print(f"Player {winner} wins!")
else:
    print("It's a draw!")

if __name__ == "__main__":
    main()
```

GO!


## FossID partialmatch

```
1 def print_board(board):
2     for row in board:
3         print(" | ".join(row))
4         print("-" * 9)
5
6 def check_winner(board, player):
7     for row in board:
8         if all(cell == player for cell in row):
9             return True
10
11    for col in range(3):
12        if all(board[row][col] == player for row in range(3)):
13            return True
14
15    if all(board[i][i] == player for i in range(3)) or all(board[i][2 - i] == player for i in range(3)):
16        return True
17
18    return False
19
```

 Limit: 10 Sensitivity: 10

 Search match, license...

Showing 1 to 1 of 1 entries

MATCH	ARTIFACT	VERSION	AUTHOR	COMPONENT LICENSE	FILE LICENSE	RELEASE DATE	FILE	SIZE	URL	HITS
partial	alx-zero_day	95944...	mwanikigachanja	MIT	None	2023-08-05	python_scripts/tic-tac_toe.py	2kb		16 (35.56%)

Show 10 entries

Previous 1 Next



# License Compliance Challenges

What do our test results show?



## Test Results

1. GitHub Copilot is trained on open-source without respecting open-source licenses
2. Filter to avoid publicly available code in suggestions does not work.
3. GitHub Copilot suggests code inside and outside GitHub.
4. ChatGPT does all the above as well.

Generative AI coding tools are trained on data that they are not allowed to use...

*GitHub Copilot relies on "software piracy on an unprecedented scale."*

Matthew Butterick (Typographer, lawyer, writer, and computer programmer.)



[← Home](#)

## GitHub Copilot

[About GitHub Copilot](#)[Quickstart](#)[Use GitHub Copilot](#)[Get code suggestions](#)[Find matching code](#)[Copilot Individual](#)[Copilot Business](#)[Copilot Enterprise](#)[Manage Copilot in your organization](#)[Copilot Chat](#)[Copilot in the CLI](#)[Configure GitHub Copilot](#)[Troubleshooting](#)[GitHub Copilot](#) / [Use GitHub Copilot](#) /

## Finding public code that matches GitHub Copilot suggestions

If you allow GitHub Copilot to make suggestions that match publicly available code, you can see references to the matching code on GitHub.com.

**Note:** GitHub Copilot code referencing is in public beta and is subject to change.

### About code referencing in GitHub Copilot [↗](#)

You can opt to allow GitHub Copilot to suggest code completions that match publicly available code on GitHub.com. For more information, see "[Configuring GitHub Copilot settings on GitHub.com](#)."

If you have allowed suggestions that match public code, GitHub Copilot can provide you with details about the matching code when you accept such suggestions. This feature is called code referencing. It is currently only available in Visual Studio Code.

When you accept a code completion suggestion that matches code in a public GitHub repository, an entry is added to a GitHub Copilot log. The log entry includes a link to a page on GitHub.com where you can view references to similar code in public GitHub repositories.

The linked web page includes details of any license identified for the repository where the matching code was found. Having reviewed the references, you can decide how to proceed. For example, you can decide what attribution to use, or whether you want to remove this code from your project.

#### Notes:

- Code referencing does not currently apply to code completion suggestions that you add to your code using GitHub Copilot Chat.
- Code referencing currently only looks for matches of accepted Copilot suggestions. Matches to code you have written, or Copilot suggestions you have altered, is not checked for matches to public code.
- Typically, matches to public code occur in less than one percent of Copilot suggestions, so you should not expect to see code references for many of the suggestions you accept.

### How code referencing finds matching code [↗](#)

Copilot code referencing searches for matches by taking the code suggestion, plus some surrounding code, and comparing it against an index of all public repositories on GitHub.com. Code in private GitHub repositories, or code outside of GitHub, is not included in the search process. The search index is refreshed every few months. As a result, newly committed code, and code from public repositories deleted before the index was created, may not be included in the search. For the same reason, the search may return matches to code that has been deleted or moved since the index was created.

#### In this article

[About code referencing in GitHub Copilot](#)[Using code referencing in Visual Studio Code](#)

...but is the true source for the reference in GitHub?

# License Compliance Challenges

*How do you resolve these challenges?*



## 1 Generative AI-Focused Scanning

Employ Software Composition Analysis (SCA) tools that can detect open source components and snippets in AI-generated code.

These tools are essential for identifying and managing compliance risks associated with code generated by AI.

## 2 Continuous Monitoring

Ensure that SCA tools are integrated into your Continuous Integration/Continuous Deployment (CI/CD) pipelines.

This integration enables continuous monitoring and compliance checks for every new AI-generated code addition, ensuring immediate detection of potential license violations.



# Agenda

*What's next on the agenda?*

- Introduction
  - The Rise of AI-Generated Code
- License Compliance Challenges
  - Assuring License Compliance with AI-Generated Code
- ➔ Security Risks in AI-Generated Code
  - Ensuring Security in AI-Generated Code
- Conclusion
  - Wrap-up and Takeaways
- Q&A



# Security Risks in AI-Generated Code



*What security risks are associated with AI-Generated Code*

Let's explore 3 of the many Security Risks associated to AI-Generated Code

1. Vulnerable Snippets
2. Obsolete Data
3. Insecure URLs

Check out the following two security risks, which we won't have time to cover today but both highlight serious risk.

*Employee Misuse*

<https://layerxsecurity.com/learn/chatgpt-security/>

*AI Package Hallucinations*

[https://vulcan.io/blog/ai-hallucinations-package-risk?utm\\_source=the+new+stack&utm\\_medium=referral&utm\\_content=inline-mention&utm\\_campaign=tns+platform](https://vulcan.io/blog/ai-hallucinations-package-risk?utm_source=the+new+stack&utm_medium=referral&utm_content=inline-mention&utm_campaign=tns+platform)

# Vulnerable Snippets

- A Security Expert, "Astrid Bytes" investigated if ChatGPT could really produce secure code.

You have a login form with a username and password. Please write a code snippet in PHP to check in the database whether the credentials match.

- The code generated showcased clear vulnerabilities
  - It's prone to SQL injection attacks
  - It allows username enumeration
  - It hardcodes credentials directly in the code
- Astrid experimented with 5 different programming languages, including PHP, Java Rust, JSON and C, but didn't noticed any significant differences when it came to more secure code.

```
php
// Replace these variables with your database connection details
$servername = "your_servername";
$username = "your_username";
$password = "your_password";
$dbname = "your_database_name";

// Establish a connection to the database
$conn = new mysqli($servername, $username, $password, $dbname);

// Check connection
if ($conn->connect_error) {
    die("Connection failed: " . $conn->connect_error);
}

// Function to sanitize user inputs
function sanitize_input($input)
{
    $input = trim($input);
    $input = stripslashes($input);
    $input = htmlspecialchars($input);
    return $input;
}

if ($_SERVER["REQUEST_METHOD"] === "POST") {
    // Get the user input from the login form
    $username = sanitize_input($_POST["username"]);
    $password = sanitize_input($_POST["password"]);

    // Query the database to find a matching user
    $sql = "SELECT * FROM users WHERE username = '$username' LIMIT 1";

    $result = $conn->query($sql);

    if ($result->num_rows === 1) {
        // If a user with the given username is found
        $user = $result->fetch_assoc();
        // Verify the password
        if (password_verify($password, $user['password'])) {
            // Password matches, user is authenticated
            echo "Login successful!";
        } else {
            // Password doesn't match
            echo "Invalid password!";
        }
    } else {
        // No user with the given username found
        echo "User not found!";
    }

    // Close the database connection
    $conn->close();
}
?>
```

*(Note: The original image contains several callouts pointing to vulnerabilities in the code above, such as hardcoded credentials, lack of input sanitization, and SQL injection.)*

Write the JWT middleware using RSA tokens in go language

Write the JWT middleware using RSA tokens in go language

Sure, here's an example of JWT middleware using RSA tokens in Go:

```

go
package main

import (
    "fmt"
    "net/http"
    "time"

    "github.com/dgrijalva/jwt-go"
)

var (
    // Define the RSA private and public keys
    privKey = []byte(`
        -----BEGIN RSA PRIVATE KEY-----
        MIIEpAIBAAKCAQEAsg5wb0xM7i8GnY=
        -----END RSA PRIVATE KEY-----
    `)

    pubKey = []byte(`
        -----BEGIN PUBLIC KEY-----
        MIIBIjANBgkqhkiG9w0BAQEFAAOCAQ8A
        -----END PUBLIC KEY-----
    `)
)

// CustomClaims represents the custom claims for JWT
type CustomClaims struct {
    Username string `json:"username"`
    jwt.StandardClaims
}

func main() {
    // Create a new HTTP server
    http.HandleFunc("/", protectedHandler)
    http.ListenAndServe(":8080", nil)
}

func protectedHandler(w http.ResponseWriter, r *http.Request) {

```

Message ChatGPT



This repository has been archived by the owner on May 21, 2022. It is now read-only.

dgrijalva / jwt-go Public archive

Notifications Fork 995 Star 10.8k

Code Issues 99 Pull requests 40 Actions Projects Wiki Security Insights

master 6 Branches 15 Tags Go to file Code

Merge pull request #429 from Waterdrips/waterdrips-fix-cve 9742bd7 · 3 years ago 330 Commits

cmd/jwt	Clarify expected format for key files.	5 years ago
request	add options to ParseFromRequest	6 years ago
test	Added password protect PEM support	8 years ago
.gitignore	Fix security issue with aud validation	3 years ago
.travis.yml	Remove unreachable code	8 years ago
LICENSE	mit license	12 years ago
MIGRATION_GUIDE.md	Fixed migration guide request.ParseFromRequest exampl...	8 years ago
README.md	Update README.md	3 years ago
VERSION_HISTORY.md	add options to ParseFromRequest	6 years ago
claims.go	Fix security issue with aud validation	3 years ago
doc.go	updating documentation	10 years ago
ecdsa.go	documentation around expected key types	6 years ago
ecdsa_test.go	Fix ES signature serialization	9 years ago
ecdsa_util.go	add parser	5 years ago

Readme MIT license

## THIS REPOSITORY IS NO LONGER MAINTANED

The new repository can be found at: <https://github.com/golang-jwt/jwt>

For more information, see issue [#462](#).

### jwt-go

About  
ARCHIVE - Golang implementation of JSON Web Tokens (JWT). This project is now maintained at:  
[github.com/golang-jwt/jwt](https://github.com/golang-jwt/jwt)

Readme MIT license Activity 10.8k stars 152 watching 995 forks Report repository

Releases 15 tags

Packages No packages published

Contributors 45

Snyk Vulnerability Database > Go > github.com/dgrijalva/jwt-go

## github.com/dgrijalva/jwt-go vulnerabilities

### Direct Vulnerabilities

Known vulnerabilities in the github.com/dgrijalva/jwt-go package. This does not include vulnerabilities belonging to this package's dependencies.

Automatically find and fix vulnerabilities affecting your projects. Snyk scans for vulnerabilities and provides fixes for free.

Fix for free

VULNERABILITY	VULNERABLE VERSION
 Access Restriction Bypass	<4.0.0-preview1
 Insecure Defaults	<3.0.0

Search by package name or CVE

LICENSES DETECTED

MIT >=v0

[View github.com/dgrijalva/jwt-go package health on Snyk Advisor](#)

Report a new vulnerability

Found a mistake?

# Insecure URLs

Write a dockerfile that installs glibc version 2.39



Creating a Dockerfile to install a specific version of glibc, such as version 2.39, involves several steps since glibc is a fundamental part of the system. Installing or upgrading it is not straightforward and should be done with caution. Here's a basic approach using an Ubuntu-based Docker image as an example:

Dockerfile

Copy code

```
# Start from a base image
FROM ubuntu:22.04

# Install necessary packages for building glibc
RUN apt-get update && \
    apt-get install -y build-essential wget bison gawk && \
    apt-get install -y libelf-dev libncurses-dev

# Set environment variables
ENV GLIBC_VERSION=2.39
ENV PREFIX=/usr/local/glibc-$GLIBC_VERSION
ENV PATH=$PREFIX/bin:$PATH
ENV LD_LIBRARY_PATH=$PREFIX/lib:$LD_LIBRARY_PATH

# Download, extract, and build glibc
RUN wget http://ftp.gnu.org/gnu/libc/glibc-$GLIBC_VERSION.tar.gz && \
    tar -xzf glibc-$GLIBC_VERSION.tar.gz && \
    cd glibc-$GLIBC_VERSION && \
    mkdir build && \
    cd build && \
    ../configure --prefix=$PREFIX && \
    make -j$(nproc) && \
    make install

# Set the default entry point to bash
CMD ["bash"]
```

Dockerfile

Copy code

```
# Use Ubuntu as the base image
FROM ubuntu:latest

# Install necessary packages
RUN apt-get update && \
    apt-get install -y wget && \
    rm -rf /var/lib/apt/lists/*

# Download glibc 2.39 from a mirror
RUN wget http://mirrors.kernel.org/ubuntu/pool/main/g/glibc/libc6_2.39-0ubuntu1_
    dpkg -i libc6_2.39-0ubuntu1_amd64.deb

# Cleanup
RUN rm -f libc6_2.39-0ubuntu1_amd64.deb

# Your additional setup or commands here
```

 [glibc-2.39.tar.gz](http://glibc-2.39.tar.gz)

2024-01-31 17:06 35M

# Insecure URLs

What is the security risk of installing file using `wget` via HTTP and not HTTPS?



Using `wget` to download files over HTTP instead of HTTPS poses several security risks:

## 1. Man-in-the-Middle (MitM) Attacks:

- HTTP traffic is not encrypted, making it susceptible to interception by attackers. They can read and modify the data in transit.
- An attacker can intercept the HTTP connection and replace the downloaded file with a malicious version. This can lead to the execution of malicious code on your system.

## 2. Lack of Integrity Verification:

- Without encryption, there's no guarantee that the file you receive is the one the server intended to send. An attacker could alter the file contents without detection.

## 3. Sensitive Information Exposure:

- Any sensitive data transmitted during the download (such as credentials or API keys if included in the URL) can be intercepted and stolen by attackers.

## 4. No Server Authentication:

- HTTP does not provide authentication of the server you're connecting to. You could be connecting to a malicious server that is impersonating the legitimate server.

## 5. Security Compliance:

- Many security policies and compliance frameworks require the use of encrypted protocols like HTTPS for data transmission to ensure data confidentiality and integrity.

Now please write a dockerfile that installs glibc version 2.39



Creating a Dockerfile that installs glibc version 2.39 requires a careful approach. We'll build glibc from source and ensure the environment is set up correctly to use this specific version of glibc. Here's a Dockerfile to achieve this:

```
Dockerfile Copy code

# Use an appropriate base image
FROM ubuntu:22.04

# Install dependencies
RUN apt-get update && \
    apt-get install -y \
    build-essential \
    wget \
    bison \
    gnawk \
    libelf-dev \
    libncurses-dev \
    python3

# Set environment variables
ENV GLIBC_VERSION=2.39
ENV PREFIX=/usr/local/glibc-$GLIBC_VERSION
ENV PATH=$PREFIX/bin:$PATH
ENV LD_LIBRARY_PATH=$PREFIX/lib:$LD_LIBRARY_PATH

# Download, extract, build, and install glibc
RUN wget https://ftp.gnu.org/gnu/libc/glibc-$GLIBC_VERSION.tar.gz && \
    tar -xzf glibc-$GLIBC_VERSION.tar.gz && \
    cd glibc-$GLIBC_VERSION && \
    mkdir build && \
    cd build && \
    ../configure --prefix=$PREFIX && \
    make -j$(nproc) && \
    make install && \
    cd ../.. && \
    rm -rf glibc-$GLIBC_VERSION glibc-$GLIBC_VERSION.tar.gz

# Set the default command to bash
CMD ["bash"]
```

# Security Risks in AI-Generated Code



*How to mitigate the security risks associated with AI generated code?*

## Implement Robust Code Review and Validation Processes

Ensure that all AI-generated code undergoes thorough review and validation by experienced developers before it is integrated into the main codebase.

- ① **Automated Static Analysis:** Use static analysis tools to automatically scan AI-generated code for common vulnerabilities, deprecated functions, insecure URLs, and outdated libraries.
- ② **Manual Code Review:** Implement a mandatory manual review process where experienced developers examine the AI-generated code for potential security issues and adherence to coding standards.
- ③ **Security Testing:** Conduct security testing, including vulnerability scanning and penetration testing, on the AI-generated code to identify and address potential security flaws.
- ④ **Approval Workflow:** Establish an approval workflow where AI-generated code must be approved by security experts or senior developers before it is merged into the main codebase.

# Agenda

*What's next on the agenda?*

- Introduction
  - The Rise of AI-Generated Code
- License Compliance Challenges
  - Assuring License Compliance with AI-Generated Code
- Security Risks in AI-Generated Code
  - Ensuring Security in AI-Generated Code
- ➔ Conclusion
  - Wrap-up and Takeaways
- Q&A



# Conclusion

*What did we learn, and what are our takeaways?*



## Wrap-Up

1. Gen AI tools are being trained on open source without respecting the open source licenses.
2. Gen AI tools are suggesting code that matches to open source and third party components.
3. Filters to exclude publicly available code suggestions do not work.
4. Gen AI tools are suggesting vulnerable code, obsolete data and insecure urls.

## Key Takeaways

1. Introduce Generative AI-Focused Scanning with Continuous Monitoring
2. Implement Robust Code Review and Validation Processes



# Agenda

*What's next on the agenda?*

- Introduction
  - The Rise of AI-Generated Code
- License Compliance Challenges
  - Assuring License Compliance with AI-Generated Code
- Security Risks in AI-Generated Code
  - Ensuring Security in AI-Generated Code
- Conclusion
  - Wrap-up and Takeaways

➔ Q&A





Any Questions?



# Thank You



---

Visit [sca.ninja](https://sca.ninja) to learn more about FossID.