





Automatic Code Generation by AI Illusion of Time Savings?



Philippe Charrière

CSM & CSE at GitLab

Developer

pcharriere@gitlab.com

@k33g on GitLab.com

@k33g_org on Twitter



Disclaimer

I'm not a Data Scientist
nor a Machine Learning Engineer
nor a AI Expert
I'm a AI user



Poll

How many Developers?

How many Data Scientists?

How many Machine Learning Engineer?

How many Ops?

How many Managers?

How many non Tech Persons?



What is GitLab? 

A collaborative tool to work on
source code, documentation, ...

And accelerate innovation



A Developer Oriented Platform? Not only

The screenshot shows the GitLab web interface for the repository 'www-gitlab-com' in the 'master' branch, specifically the 'doc' directory. The interface includes a sidebar with navigation options and a main content area displaying a table of files and folders.

Navigation options in the sidebar:

- Project overview
- Pinned
- Vulnerability report
- Pipelines
- Policies
- Environments
- Merge requests (2.2k)
- Releases
- Issues (18.9k)
- Manage
- Plan
- Code
- Build
- Secure
- Deploy
- Operate
- Monitor
- Analyze

Main content area table:

| Name | Last commit | Last update |
|----------------------------------|---|--------------|
| .. | | |
| legal | Fix reference link in individual_contributor_license_agreement.md | 6 years ago |
| templates | Fix typos in doc/ | 1 month ago |
| bamboohr.md | Update references to the team.yml file | 2 years ago |
| blog-posts.md | Made some changes | 1 month ago |
| development.md | Fix typos in doc/ | 1 month ago |
| docker.md | Update Directions for Docker to use Registry Image | 1 year ago |
| formatting_markdown.md | Rewrite markdown-formatter | 2 years ago |
| gitpod.md | Code/doc changes for renaming 'marketing' site to 'uncategorized' | 2 years ago |
| monorepo.md | Fix typos in doc/ | 1 month ago |
| netlifycms.md | Fix typos in doc/ | 1 month ago |
| press.md | Move UX handbook to Product | 9 months ago |
| projects.md | Update references to the team.yml file | 2 years ago |
| recurly.md | Fix typos in doc/ | 1 month ago |
| redirect_blogposts_org_to_com.md | Add the new site 🌟 | 7 years ago |

Issues

doc - master - GitLab.com / www | Issues - GitLab.com / www-gitlab-com | GitLab should maintain a calendar | Trainee Secure Analyzer Maint | Trainee BE maintainer (GitLab) | Merge requests - GitLab.com /

gitlab.com/gitlab-com/www-gitlab-com/-/issues/34020

Next | GitLab.com > www-gitlab-com > Issues > #34020

99+ | 53 | 99+

www-gitlab-com

- Project overview
- Pinned
- Vulnerability report
- Pipelines
- Policies
- Environments
- Merge requests 2.2k
- Releases
- Issues 18.9k
- Manage
- Plan
- Code
- Build
- Secure
- Deploy
- Operate
- Monitor
- Analyze

Open Trainee BE maintainer (GitLab) - Hitesh Raghuvanshi

Hitesh Raghuvanshi @hrahuvanshi · 2 months ago Author Developer

Fix caching to cover whole Service Ping payload: gitlab-org/gitlab!114992 (merged)

During review:

- Approved as-is, I did not add any comments.

Post-review:

- Merged as-is

@terrachu Please add feedback, and compare this review to the average maintainer review.

▼ Collapse replies

Terri Chu @terrachu · 2 months ago Developer

@hrahuvanshi Change was small, so I felt ok merging after testing it and looking at specs. Only advice I have is to manually test or ask for a domain expert review before merging if you're in an area you are unfamiliar with.

Hitesh Raghuvanshi @hrahuvanshi · 2 months ago Author Developer

Thanks for the feedback, I will keep that in mind.

Reply...

Hitesh Raghuvanshi @hrahuvanshi · 2 months ago Author Developer

Add user toggle for achievements: gitlab-org/gitlab!116258 (merged)

During review:

- Backend changes approved as is.

Add a to do

Assignee
Hitesh Raghuvanshi

Epic
None

Labels
None

Milestone
None

Iteration
None

Weight
None

Due date
None

Time tracking
No estimate or time spent

Health status
None

Confidentiality
Not confidential

14 Participants

Merge requests

doc - master · GitLab.com / www · Merge requests · GitLab.com / Update VSM category with whi · Add clarification for scenarios · Draft: introducing GitLab.com / Merge requests · GitLab.com /

gitlab.com/gitlab-com/www-gitlab-com/-/merge_requests/126188

GitLab.com > www-gitlab-com > Merge requests > 126188

99+ 53 99+

www-gitlab-com

Project overview
Pinned
Vulnerability report
Pipelines
Policies
Environments
Merge requests 2.2k
Releases
Issues 18.9k
Manage
Plan
Code
Build
Secure
Deploy
Operate
Monitor
Analyze

Open Add clarification for scenarios in which Workday can be leveraged for internal transfers gL-workday-lateral-transfe... into master

Overview 7 Commits 4 Pipelines 8 Changes 1 1 unresolved thread Add a to do

Alissa Meeks @ameeks started a thread on an old version of the diff 3 days ago Resolved 3 days ago by Giuliana Lucchesi Show thread

Giuliana Lucchesi resolved all threads 3 days ago

Giuliana Lucchesi added 1 commit 3 days ago

- 513b25e6 - Apply 1 suggestion(s) to 1 file(s)

Compare with previous version

Phil Calder @pcalder started a thread on the diff 3 days ago Hide thread

sites/handbook/source/handbook/people-group/promotions-transfers/index.html.md

| | | |
|-----|-----|---|
| 257 | 257 | |
| 258 | 258 | Regardless of the promotion level, it is critical that leaders work with their People Business Partner, Total Rewards, and FP&A [as outlined](/handbook/total-rewards/compensation/compensation-review-cycle/#compensation-program-budget) to identify tradeoffs we can review to fund the promotion. |
| 259 | 259 | |
| 260 | | - ### Greenhouse Promotions/Transfers Process |
| 260 | | + ### Transfers: Greenhouse or Workday |
| 261 | | + |
| 262 | | + #### Greenhouse |
| 261 | 263 | |
| 262 | 264 | The following types of promotions/transfers can be processed through Greenhouse: |
| 263 | 265 | - Interim Roles: Interim roles at all levels should go through the Greenhouse interview process (regardless of the number of applicants for the role). |
| 264 | 266 | - Example: Senior Backend Engineer applies to become a Manager, Engineering (interim) |
| 265 | 267 | - Interim roles that transition to the permanent role: All team members that started the Greenhouse process and were hired into the interim role will make the transition to the |

Assignee Edit
Alissa Meeks

0 Reviewers Edit
None - assign yourself

Labels Edit
handbook

Milestone Edit
None

Time tracking +
No estimate or time spent

15 Participants
+ 7 more

Automatic Code Generation by AI Illusion of Time Savings?



Some brief definitions reminders about AI for dummies (like me)



Some brief definitions reminders

- A **machine learning model** is a computer program that is used to recognize patterns in data or make predictions.
- **Machine learning for source code generation** is a field that uses machine learning models to write computer code from various inputs, such as natural language descriptions, incomplete code...
- **AI code generation** is a process of using machine learning models to write code from input that describes what the code should do.

Relevance of ML models for AI Code Generation



Do you know Chat GPT?





Please Help

Hello Bing; I need you to create me a cooking recipe based on the following elements: tomatoes, potatoes, sausages, basil, and chives. Can you **please** help me?



Hello, this is Bing. I can help you create a cooking recipe based on the following elements: tomatoes, potatoes, sausages, basil, and chives. Here is a possible recipe:

Potato and sausage bake with tomatoes

Ingredients:

- 2 lbs of potatoes
- 4 sausages of your choice
- 4 tomatoes
- 1 cup of heavy cream
- 1 cup of shredded cheese
- A few basil leaves
- A few chives
- Salt, pepper, olive oil

Directions:

- Preheat the oven to 350°F (177°C).
- Peel and slice the potatoes thinly. Cook them in a pot of boiling salted water for 15 minutes. Drain and set aside.
- Cut the sausages into pieces. Cook them in a skillet with some olive oil for 10 minutes. Set aside.
- Wash and slice the tomatoes. Chop the basil and the chives.
- In a baking dish, layer some potatoes, then some sausages, then some tomatoes. Season with salt, pepper and herbs. Repeat until all the ingredients are used up.
- Pour the heavy cream over the bake and sprinkle with cheese.

From a **Developer** perspective
Will AI code generation help me?



I did some tests (with Bing / Chat GPT4)

Can you please generate a program with a Golang function that returns a unique string Id, and call this function from the main function?

It works!



```
// Import the time, rand and fmt packages
import (
    "time"
    "math/rand"
    "fmt"
)

// GenerateID returns a unique string Id using the current time
and a random number
func GenerateID() string {
    // Get the current time in nanoseconds
    t := time.Now().UnixNano()
    // Generate a random number between 0 and 9999
    r := rand.Intn(10000)

    // Format the time and the random number as a string with
four digits
    return fmt.Sprintf("%d%04d", t, r)
}

// Main function
func main() {
    // Call the GenerateID function and print the result
    id := GenerateID()
    fmt.Println("The generated Id is:", id)
}
```

```
● 09:52:41 ubuntu@extism-web-ide demo ±|main x|→ go run main.go
16875067660978139916216
● 09:52:46 ubuntu@extism-web-ide demo ±|main x|→ go run main.go
16875067683180224426363
● 09:52:48 ubuntu@extism-web-ide demo ±|main x|→ go run main.go
The generated Id is: 16875069079794269330435
● 09:55:07 ubuntu@extism-web-ide demo ±|main x|→ go run main.go
The generated Id is: 16875069094650940489979
● 09:55:09 ubuntu@extism-web-ide demo ±|main x|→ go run main.go
The generated Id is: 16875069111085951584186
● 09:55:11 ubuntu@extism-web-ide demo ±|main x|→ go run main.go
The generated Id is: 16875069121807986866406
● 09:55:12 ubuntu@extism-web-ide demo ±|main x|→ go run main.go
The generated Id is: 16875072427962229887303
○ 10:00:42 ubuntu@extism-web-ide demo ±|main x|→ █
```

I did other tests (with Bing / Chat GPT4)

Can you **please**, generate in GoLang the source code of an HTTP server with a “/users” route. For this route if the HTTP method is POST it will create a user and if the HTTP method is GET it will get the list of the users. And I need a “/hello” route to get a greetings message.

It works again!



```
// HandleUsers is a handler function that handles both creating
// and getting users depending on the request method
func HandleUsers(w http.ResponseWriter, r *http.Request) {
    // Set the content type of the response to JSON
    w.Header().Set("Content-Type", "application/json")
    // Switch on the request method
    switch r.Method {
    case "POST":
        // Create a new User struct
        var user User
        // Decode the JSON data from the request body into the
        user struct
        err := json.NewDecoder(r.Body).Decode(&user)
        // If there is an error, write a bad request status and
        return
        if err != nil {
            w.WriteHeader(http.StatusBadRequest)
            return
        }
        // Append the user struct to the Users slice
        Users = append(Users, user)
        // Write a created status to the response
        w.WriteHeader(http.StatusCreated)
        // Encode the user struct as JSON and write it to the
        response body
        json.NewEncoder(w).Encode(user)
    case "GET":
        // Write an ok status to the response
        w.WriteHeader(http.StatusOK)
        // Encode the Users slice as JSON and write it to the
        response body
        json.NewEncoder(w).Encode(Users)
    default:
        // If the request method is not POST or GET, write a
        method not allowed status and return
        w.WriteHeader(http.StatusMethodNotAllowed)
    }
}
```

```
// SayHello is a handler function that returns a greetings
// message as plain text in the response body
func SayHello(w http.ResponseWriter, r *http.Request) {
    // Set the content type of the response to plain text
    w.Header().Set("Content-Type", "text/plain")
    // Check if the request method is GET
    if r.Method == "GET" {
        // Write an ok status to the response
        w.WriteHeader(http.StatusOK)
        // Write a greetings message to the response body
        w.Write([]byte("Hello, welcome to my HTTP server!"))
    } else {
        // If the request method is not GET, write a method not
        allowed status and return
        w.WriteHeader(http.StatusMethodNotAllowed)
    }
}
```

```
// Main function
func main() {
    // Initialize an empty Users slice
    Users = []User{}
    // Register the handler functions for each route with the
    default ServeMux
    http.HandleFunc("/users", HandleUsers)
    http.HandleFunc("/hello", SayHello)
    // Start listening for HTTP requests on port 8080 using the
    default ServeMux
    http.ListenAndServe(":8080", nil)
}
```

```
JSON_DATA:: command not found
● 10:25:35 ubuntu@extism-web-ide demo ±|main x| → JSON_DATA='{ "name": "Bob Morane", "email": "bob@gitlab.com" }'
● 10:25:44 ubuntu@extism-web-ide demo ±|main x| → curl -X POST http://localhost:8080/users \
-H 'Content-Type: application/json; charset=utf-8' \
-d "${JSON_DATA}"
{"name": "Bob Morane", "email": "bob@gitlab.com"}
● 10:25:49 ubuntu@extism-web-ide demo ±|main x| → JSON_DATA='{ "name": "Bill Balantine", "email": "bill@gitlab.com" }'
● 10:26:18 ubuntu@extism-web-ide demo ±|main x| → curl -X POST http://localhost:8080/users \
-H 'Content-Type: application/json; charset=utf-8' \
-d "${JSON_DATA}"
{"name": "Bill Balantine", "email": "bill@gitlab.com"}
● 10:26:23 ubuntu@extism-web-ide demo ±|main x| → curl http://localhost:8080/users
[{"name": "Bob Morane", "email": "bob@gitlab.com"}, {"name": "Bill Balantine", "email": "bill@gitlab.com"}]
● 10:26:49 ubuntu@extism-web-ide demo ±|main x| →
```

But, the life of a **Developer** is often more complicated...



Requirement: WebAssembly plugins

- WebAssembly ?
- Requirement: Run WASM functions from the Host
- Why: Efficiency, Speed, Safe >>> FaaS

>>> write an application that runs WASM programs

Problem

- The specification is half-baked (but in progress)
- Usually simple things are complicated to do
- Reinvent the wheel

I worked several days(weeks) to implement a solution

- Read the documentation
- Read the samples
- It works! 🎉
- I improved my expertise 💪 >>> 📄✍️ blog posts
- I can reuse it 📦

But let's imagine that I don't have the
solution

I'm lazy sometimes...



Ask ChatGPT

- The logic is **not bad**
- Some statements are **false**
- Some references are **not relevant**
- **The code does not compile**

I use the Wazero framework to make an application in Go that loads a wasm module developed in TinyGo, I need to call a "hello" function of the wasm module by passing it a character string (call the parameter variable "name "), and the result should be a string equal to the concatenation of "Greetings" and the name variable. Knowing that I have to use the WASI specification to do this, can you generate me a code for both the host application and the wasm module? According the WASI specification, a WASM function can only take numbers as parameters and can only return only one number. So, I need the helpers to copy the string values in memory from the host and the helpers to read memory from the WASM module.

So, ChatGPT 🐉 useful or not?

- Without expertise, it could be a wasted time
- You need to understand why it does not work
- As a Junior developer, you don't know if your code is optimized

Before using ChatGPT

For complex use cases

- Read the documentation
- First search tutorials on the internet
- Ask Senior Developers

Or ask ChatGPT for Tutorials



I found some resources that might help you with your task. Here are some summaries:

- **GitHub - tetratelabs/wazero: wazero: the zero dependency WebAssembly runtime for Go developers**¹: This is the official repository of wazero, a WebAssembly runtime written in Go that supports WASI specifications. You can find examples, documentation, and installation instructions here.
- **Wazero Cookbook - Part One: WASM function & Host application**²: This is a blog post that shows how to write a simple wasm function in Rust and call it from a Go host application using wazero. It explains how to pass a string value to the wasm function and get a string value back from the return value using the WASM memory buffer.
- **Wazero Cookbook - Part Two: Host functions**³: This is another blog post that shows how to write host functions in Go and register them with wazero. It also shows how to pass and return string values between the host functions and the wasm functions using the WASM memory buffer.

I hope these resources are useful for you. 😊

Learn more: [1. github.com](#) [2. k33g.hashnode.dev](#) [+1 more](#)

1 of 30 ●

ChatGPT 🤖, Risks

- **Productivity** of Junior developers **will not increase** (it may even be worse)
- Junior developers will produce **poor** quality code **faster**
- Senior developers will spend more time to fix the source code of the other developers



So, are you telling me AI is useless?



No! 🤗

AI is great, and it's only the beginning



The real opportunity goes beyond creating code



How developers spend their time

- 39%**
Understanding code,
writing new code or
improving existing code
- 61%**
Everything else

Source: Tidelfit



How to helps developers for real with AI features?

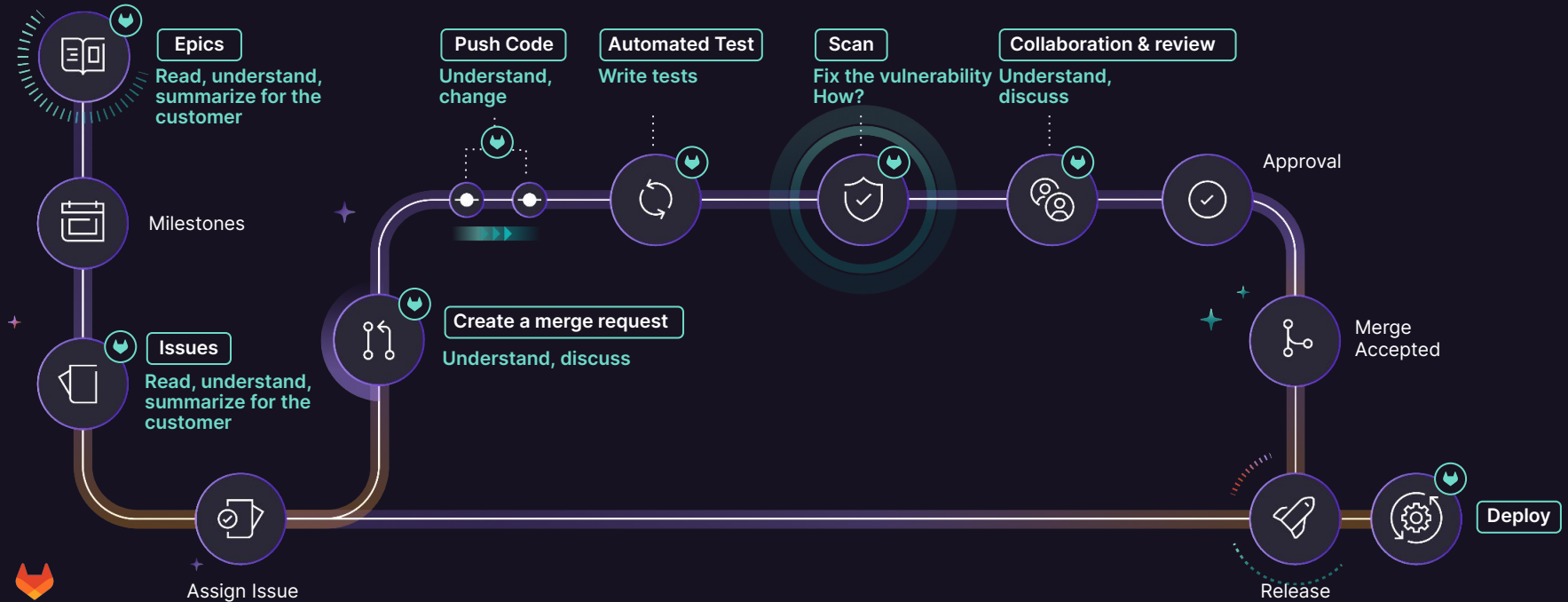


How to helps **GitLab users** for real
with **AI features**?

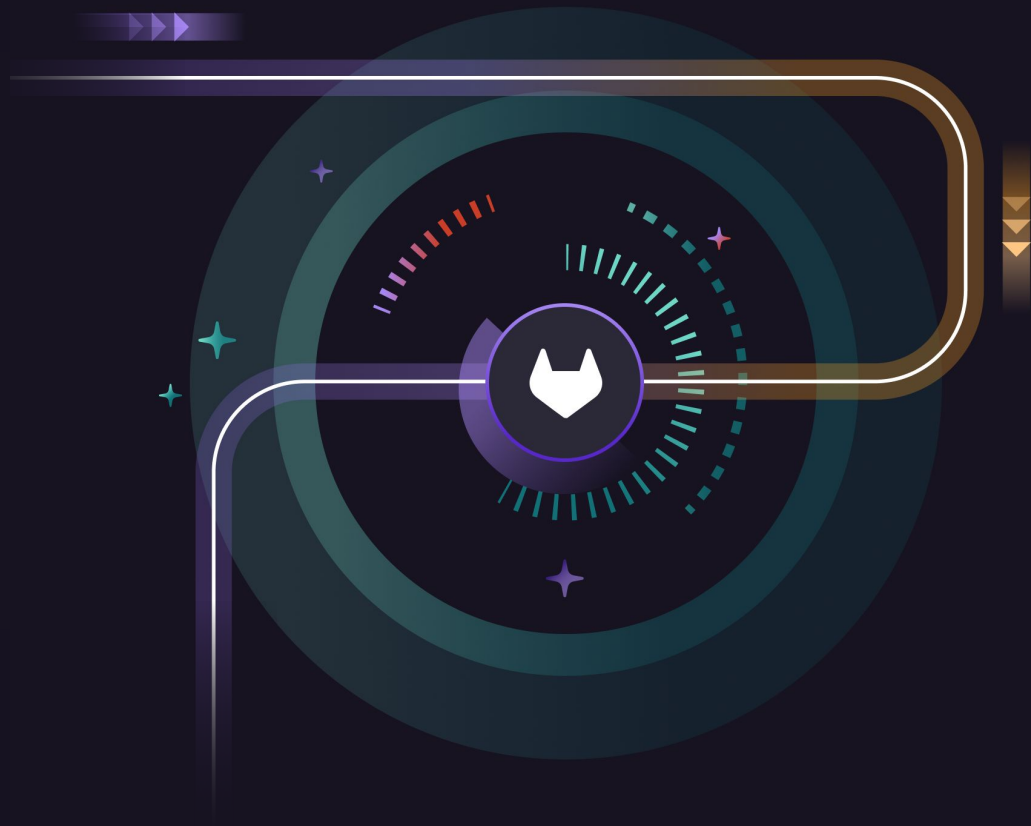


What is my day-to-day job at GitLab?

CSM, CSE, Developer... And the others



AI-assisted features

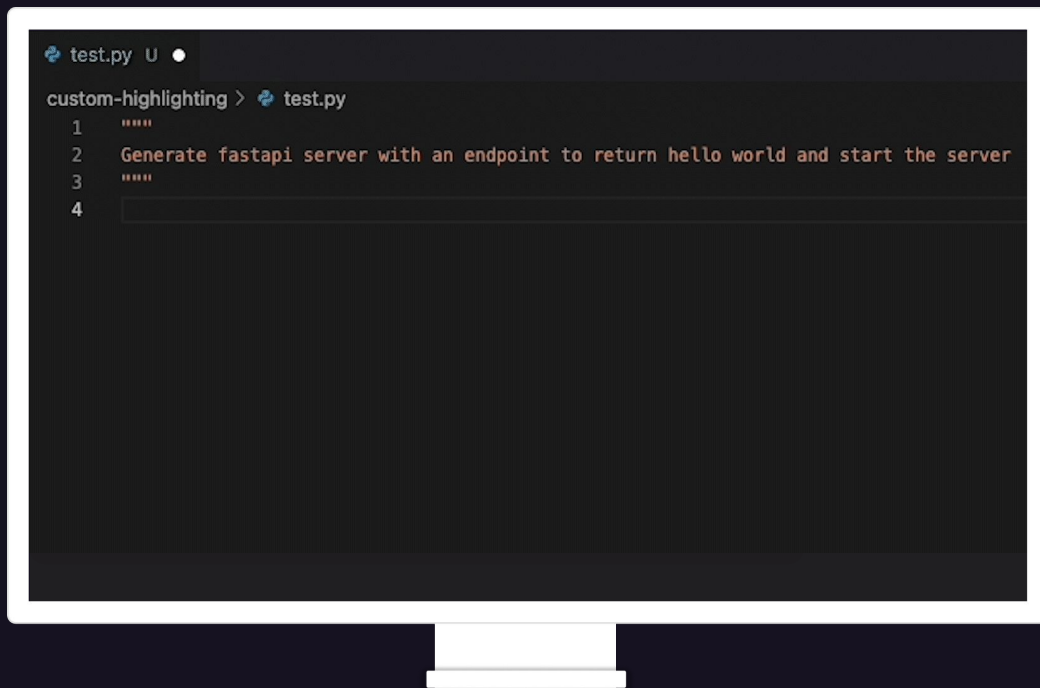


From a developer perspective

I want ...

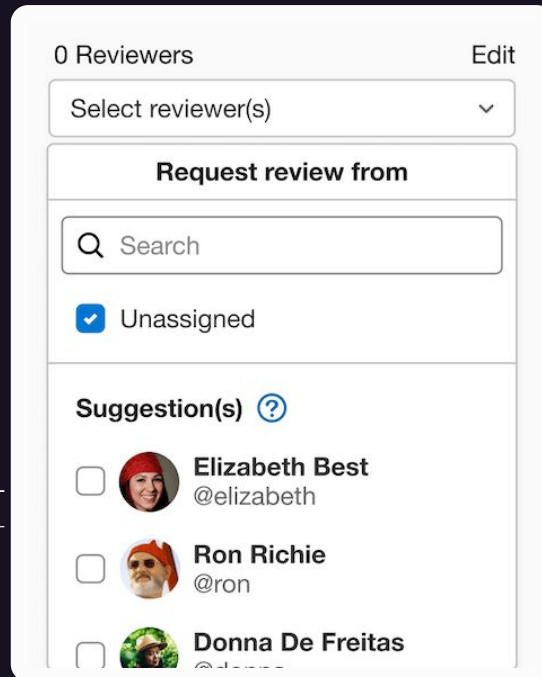


AI Assisted code suggestions



Suggested Reviewers

Find the best person to review your code changes






0 Reviewers Edit

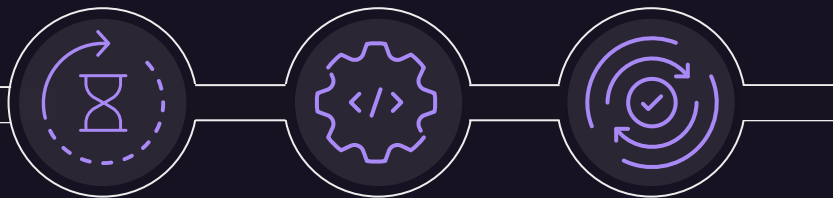
Select reviewer(s) ▼

Request review from

Unassigned

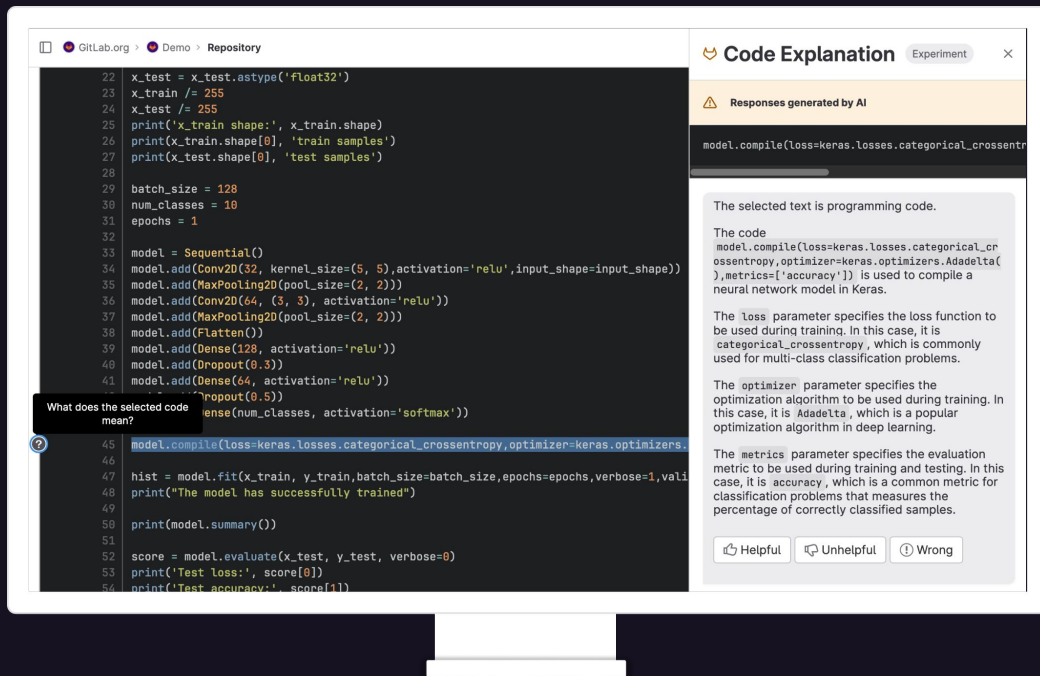
Suggestion(s) ?

-  **Elizabeth Best**
@elizabeth
-  **Ron Richie**
@ron
-  **Donna De Freitas**
@donna



Explain this code

Inline code explanations



The screenshot shows a GitLab IDE interface. The main editor displays Python code for training a neural network. A tooltip is visible over line 45, asking "What does the selected code mean?". The Code Explanation sidebar on the right provides detailed answers for the selected code.

```
22 x_test = x_test.astype('float32')
23 x_train /= 255
24 x_test /= 255
25 print('x_train shape:', x_train.shape)
26 print(x_train.shape[0], 'train samples')
27 print(x_test.shape[0], 'test samples')
28
29 batch_size = 128
30 num_classes = 10
31 epochs = 1
32
33 model = Sequential()
34 model.add(Conv2D(32, kernel_size=(5, 5), activation='relu', input_shape=input_shape))
35 model.add(MaxPooling2D(pool_size=(2, 2)))
36 model.add(Conv2D(64, (3, 3), activation='relu'))
37 model.add(MaxPooling2D(pool_size=(2, 2)))
38 model.add(Flatten())
39 model.add(Dense(128, activation='relu'))
40 model.add(Dropout(0.3))
41 model.add(Dense(64, activation='relu'))
42 model.add(Dropout(0.5))
43 model.add(Dense(num_classes, activation='softmax'))
44
45 model.compile(loss=keras.losses.categorical_crossentropy, optimizer=keras.optimizers.
46
47 hist = model.fit(x_train, y_train, batch_size=batch_size, epochs=epochs, verbose=1, vali
48 print("The model has successfully trained")
49
50 print(model.summary())
51
52 score = model.evaluate(x_test, y_test, verbose=0)
53 print('Test loss:', score[0])
54 print('Test accuracy:', score[1])
```

Code Explanation Experiment

Responses generated by AI

```
model.compile(loss=keras.losses.categorical_crossentropy,
```

The selected text is programming code.

The code `model.compile(loss=keras.losses.categorical_crossentropy, optimizer=keras.optimizers.Adadelta(), metrics=['accuracy'])` is used to compile a neural network model in Keras.

The `loss` parameter specifies the loss function to be used during training. In this case, it is `categorical_crossentropy`, which is commonly used for multi-class classification problems.

The `optimizer` parameter specifies the optimization algorithm to be used during training. In this case, it is `Adadelta`, which is a popular optimization algorithm in deep learning.

The `metrics` parameter specifies the evaluation metric to be used during training and testing. In this case, it is `accuracy`, which is a common metric for classification problems that measures the percentage of correctly classified samples.

Helpful Unhelpful Wrong

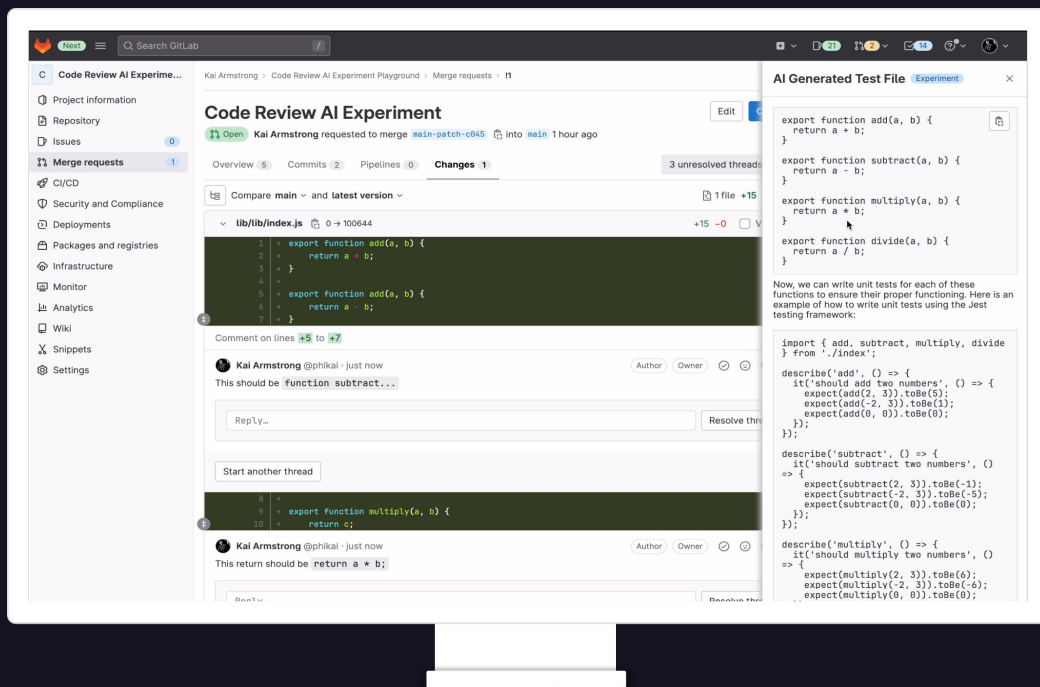


Generate Tests



in MR

Quickly add unit tests
to merge requests



The screenshot shows the GitLab Code Review AI Experiment interface. The main content area displays a merge request for the file `lib/lib/index.js`. The code is shown with line numbers 1 through 7. A comment from Kai Armstrong asks for a `function subtract`. The interface includes a sidebar with navigation options like Project Information, Repository, Issues, Merge requests, CI/CD, Security and Compliance, Deployments, Packages and registries, Infrastructure, Monitor, Analytics, Wiki, Snippets, and Settings. A right-hand panel titled "AI Generated Test File" shows the following code:

```
export function add(a, b) {
  return a + b;
}
export function subtract(a, b) {
  return a - b;
}
export function multiply(a, b) {
  return a * b;
}
export function divide(a, b) {
  return a / b;
}
```

Below the code, the text reads: "Now, we can write unit tests for each of these functions to ensure their proper functioning. Here is an example of how to write unit tests using the Jest testing framework:"

```
import { add, subtract, multiply, divide } from './index';

describe('add', () => {
  it('should add two numbers', () => {
    expect(add(2, 3)).toBe(5);
    expect(add(-2, 3)).toBe(1);
    expect(add(0, 0)).toBe(0);
  });
});

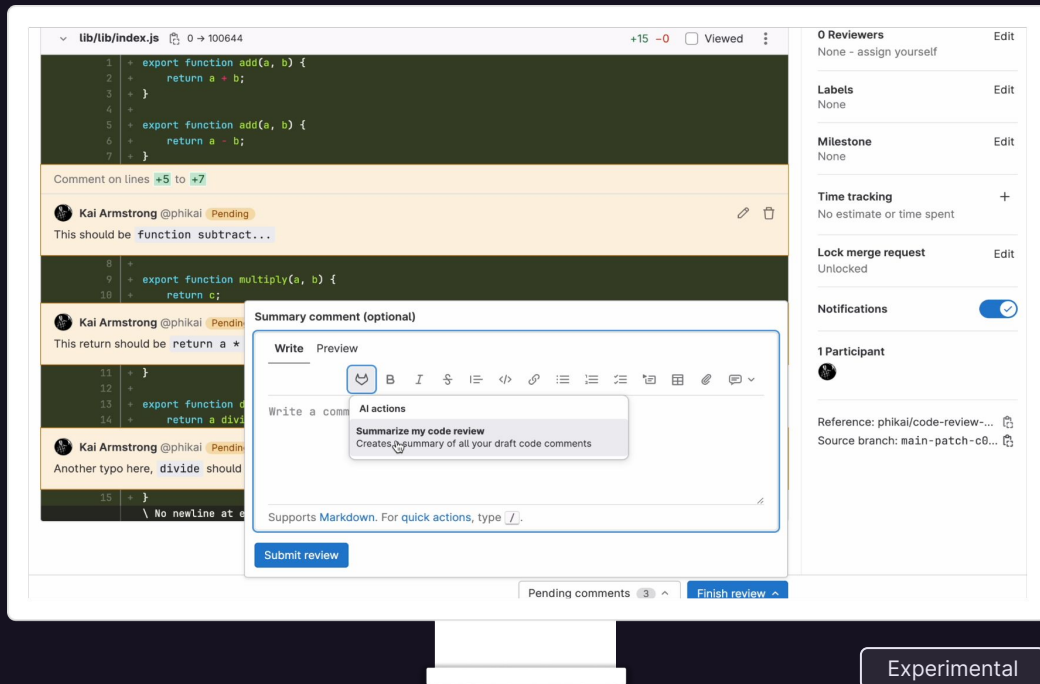
describe('subtract', () => {
  it('should subtract two numbers', () => {
    expect(subtract(2, 3)).toBe(-1);
    expect(subtract(-2, 3)).toBe(-5);
    expect(subtract(0, 0)).toBe(0);
  });
});

describe('multiply', () => {
  it('should multiply two numbers', () => {
    expect(multiply(2, 3)).toBe(6);
    expect(multiply(-2, 3)).toBe(-6);
    expect(multiply(0, 0)).toBe(0);
  });
});
```



Summarize my MR Review

Quickly summarize changes
you recommend making



The screenshot displays a GitLab Merge Request (MR) review interface for a file named `lib/lib/index.js`. The code diff shows several changes, including function definitions for `add` and `multiply`. Comments from user `Kai Armstrong` are visible, such as "This should be function subtract..." and "This return should be return a *".

An AI summary tool is open, titled "Summary comment (optional)". It features a "Write" tab and a "Preview" tab. The "Write" tab contains a text input field with the placeholder "Write a comment" and a rich text editor toolbar. Below the input field, there is a section for "AI actions" with a button labeled "Summarize my code review" and a description: "Creates a summary of all your draft code comments". At the bottom of the tool, it states "Supports Markdown. For quick actions, type `/`." and a "Submit review" button.

On the right side of the MR interface, there are sections for "0 Reviewers", "Labels", "Milestone", "Time tracking", "Lock merge request", "Notifications", and "1 Participant".

Experimental



java/src/main/TemplateInjection.java

+1 -1 Viewed

```
@@ -38,7 +38,7 @@ public class TemplateInjection {
38     VelocityContext context = new VelocityContext();
39     StringWriter swOut = new StringWriter();
40
41 -    Velocity.evaluate(context, swOut, "test", inputStream.toString());
```

```
42     Velocity.evaluate(context, swOut, "test", fileReader);
43     Velocity.evaluate(context, swOut, "test", template);
```

```
44     }
```

↓

```
@@ -38,7 +38,7 @@ public class TemplateInjection {
38     VelocityContext context = new VelocityContext();
39     StringWriter swOut = new StringWriter();
40
41 +    //Velocity.evaluate(context, swOut, "test", inputStream.toString());
```

```
42     Velocity.evaluate(context, swOut, "test", fileReader);
43     Velocity.evaluate(context, swOut, "test", template);
```

```
44     }
```

↓



Philippe Charrière @k33g Pending



Why only remove the code? It's useless



Philippe Charrière @k33g Pending



I'm not sure that using a template is a good idea

Activity

All activity ▾



- **Philippe Charrière** assigned to [@k33g](#) 1 hour ago
 - **Philippe Charrière** added 1 commit 1 hour ago
 - [78701697](#) - Update file TemplateInjection.java
- [Compare with previous version](#)



Philippe Charrière @k33g Pending



Here's a quick summary of the code review:

- `java/src/main/TemplateInjection.java`: Consider alternative solutions to using a template.
- `java/src/main/TemplateInjection.java`: Remove unnecessary code instead of commenting it out.

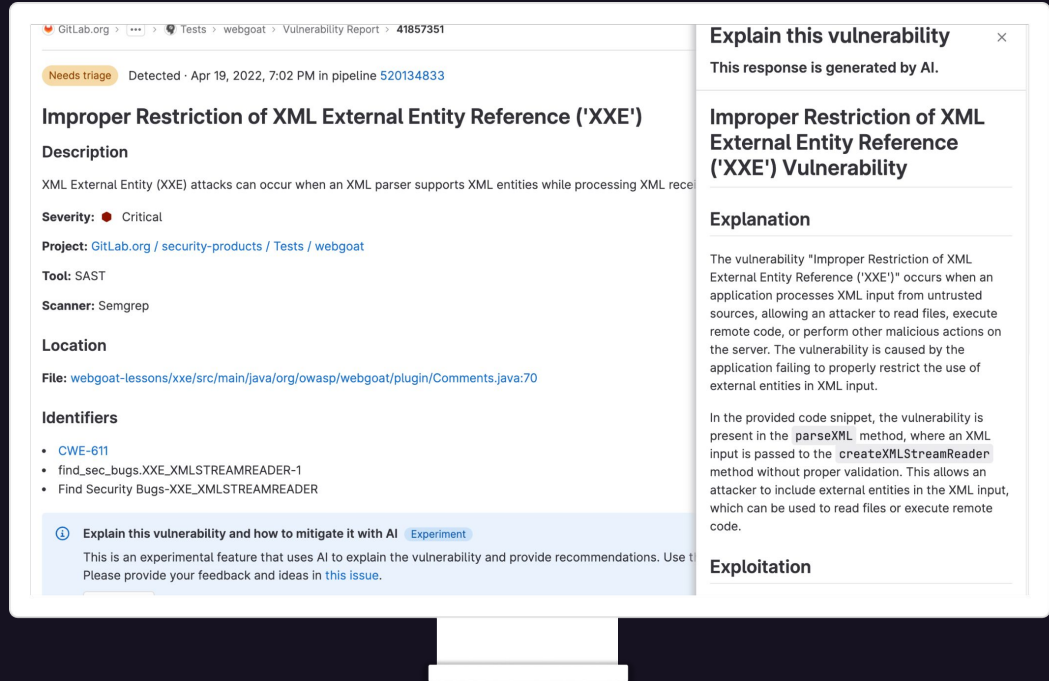
This comment was generated by AI

From a Security perspective I want ...



Explain this vulnerability

Continuous security improvements



GitLab.org > Tests > webgoat > Vulnerability Report > 41857351

Needs triage Detected · Apr 19, 2022, 7:02 PM in pipeline 520134833

Improper Restriction of XML External Entity Reference ('XXE')

Description

XML External Entity (XXE) attacks can occur when an XML parser supports XML entities while processing XML rece

Severity: ● Critical

Project: [GitLab.org / security-products / Tests / webgoat](#)

Tool: SAST

Scanner: Semgrep

Location

File: [webgoat-lessons/xxe/src/main/java/org/owasp/webgoat/plugin/Comments.java:70](#)

Identifiers

- CWE-611
- find_sec_bugs.XXE_XMLSTREAMREADER-1
- Find Security Bugs-XXE_XMLSTREAMREADER

Explain this vulnerability and how to mitigate it with AI Experiment

This is an experimental feature that uses AI to explain the vulnerability and provide recommendations. Use t
Please provide your feedback and ideas in [this issue](#).

Explain this vulnerability

This response is generated by AI.

Improper Restriction of XML External Entity Reference ('XXE') Vulnerability

Explanation

The vulnerability "Improper Restriction of XML External Entity Reference ('XXE')" occurs when an application processes XML input from untrusted sources, allowing an attacker to read files, execute remote code, or perform other malicious actions on the server. The vulnerability is caused by the application failing to properly restrict the use of external entities in XML input.

In the provided code snippet, the vulnerability is present in the `parseXML` method, where an XML input is passed to the `createXMLStreamReader` method without proper validation. This allows an attacker to include external entities in the XML input, which can be used to read files or execute remote code.

Exploitation





What is the vulnerability?

The vulnerability is that the code allows an attacker to inject arbitrary code into the template engine. This can be done by providing malicious input to the `inputStream` parameter. The attacker can then execute arbitrary code on the server.

How can an attacker take advantage of the vulnerability?

The attacker can take advantage of the vulnerability by providing malicious input to the `inputStream` parameter. For example, the attacker could provide a string that contains a malicious script. When the code is executed, the malicious script will be executed on the server.

How can it be fixed?

The vulnerability can be fixed by sanitizing the input to the `inputStream` parameter. This can be done by using a whitelist of allowed characters. For example, the code could only allow characters that are allowed in HTML.

Example of vulnerable code

The following is an example of vulnerable code:

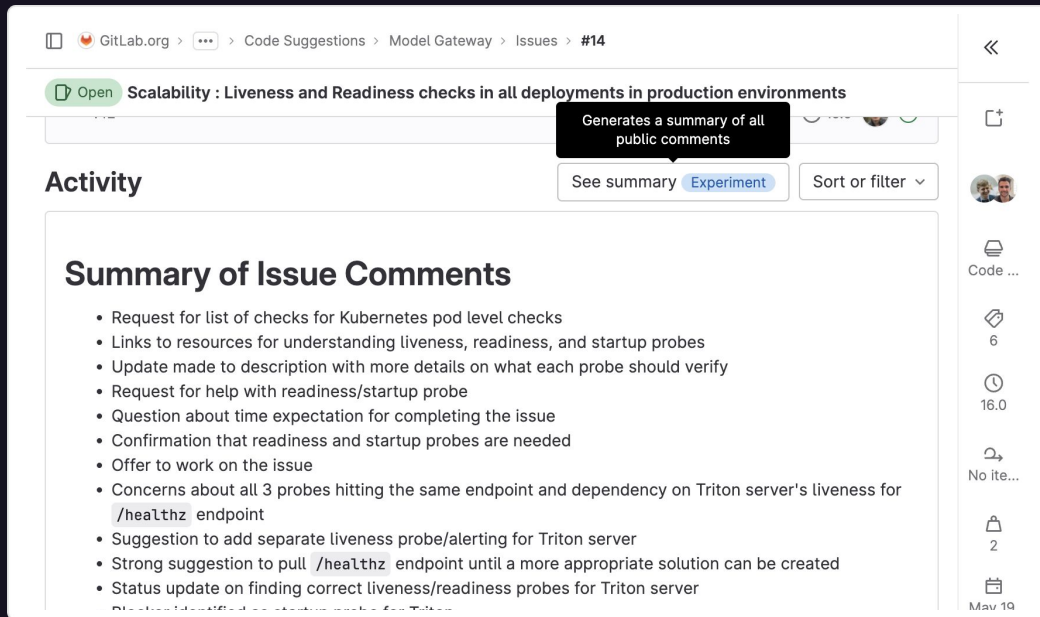
From a CSM & CSE perspective (and more)

I want ...



Issue summaries

Summarize long issue and merge
request comment threads

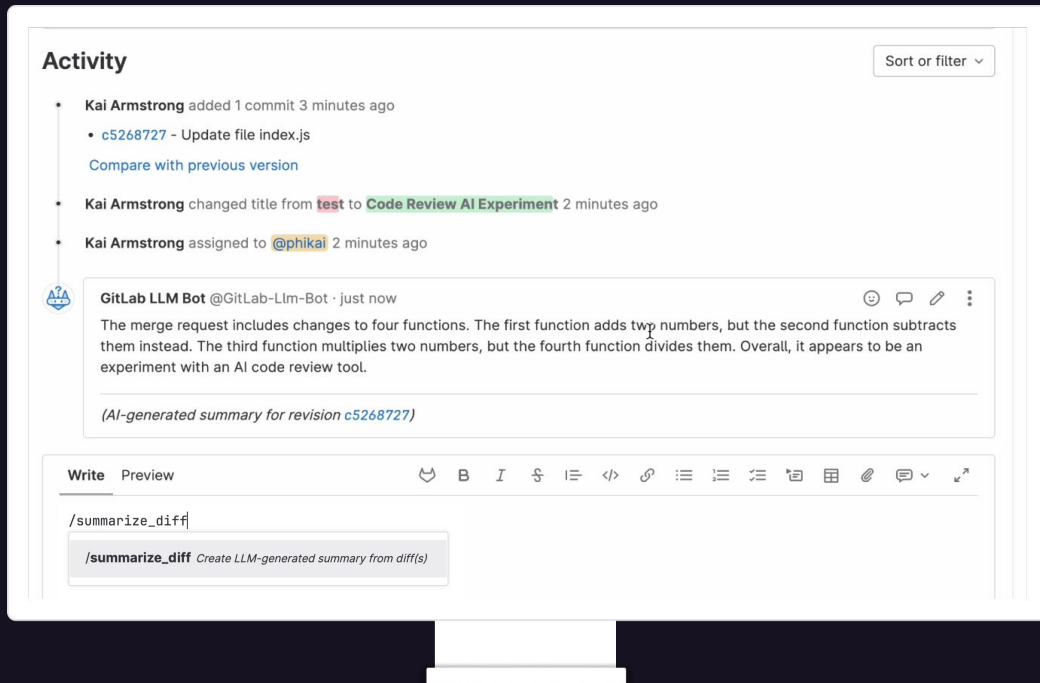


Note: UX may not represent final design



Summarize MR Changes

Drive alignment and
understanding of changes



Activity Sort or filter ▾

- Kai Armstrong** added 1 commit 3 minutes ago
 - [c5268727](#) - Update file index.js
 - [Compare with previous version](#)
- Kai Armstrong** changed title from `test` to `Code Review AI Experiment` 2 minutes ago
- Kai Armstrong** assigned to [@phika](#) 2 minutes ago

GitLab LLM Bot @GitLab-Llm-Bot · just now 🗨️ 💬 ✎️ ⋮

The merge request includes changes to four functions. The first function adds two numbers, but the second function subtracts them instead. The third function multiplies two numbers, but the fourth function divides them. Overall, it appears to be an experiment with an AI code review tool.

(AI-generated summary for revision [c5268727](#))

Write Preview 🔄 B I ↺ ↻ </> 🔗 ☰ ☷ ☸ 📄 📁 📧 📧 ⌵ ↗

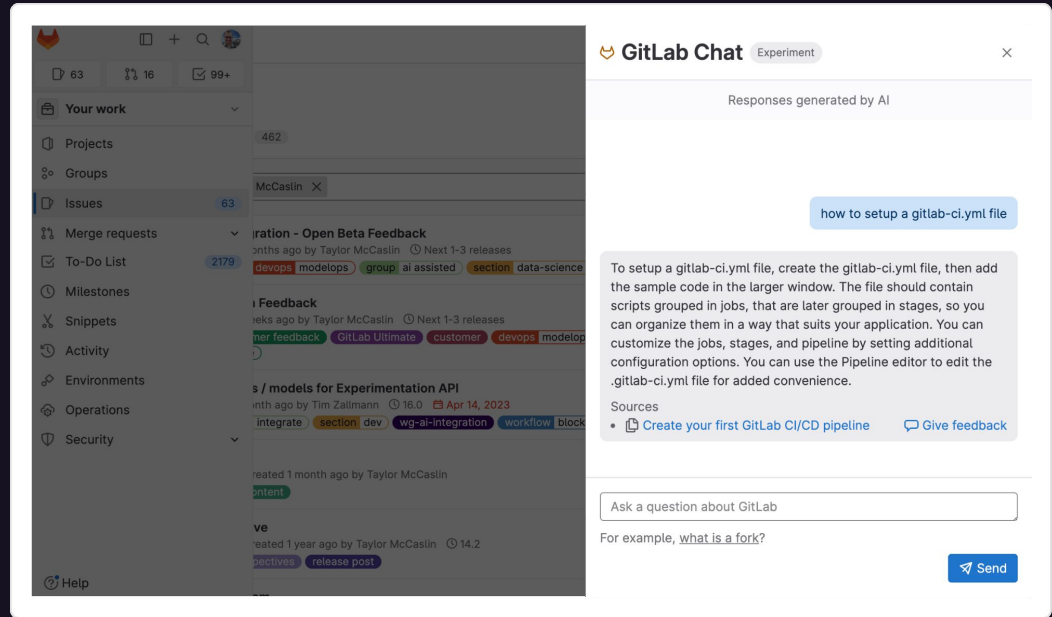
```
/summarize_diff
```

`/summarize_diff` Create LLM-generated summary from diff(s)



GitLab Chat

Ask questions rather than search for answers



Experimental



Closed :mega: Feedback issue for: [Experiment] Explain Vulnerability with AI

Adjust AI mutation rate limits !118202

Provide a single line of vulnerable code !118270

Activity

[View summary](#) [Sort or filter](#)

Phil Calder added to epic &10284 (closed) 2 months ago

Phil Calder changed title from Feedback issue for: [AI Proposal] Explain this Vulnerability to Feedback issue for: [AI Proposal] Explain this Vulnerability :mega 2 months ago

Phil Calder changed the description 2 months ago · [Compare with previous version](#)

Phil Calder added `group` `threat insights` `devops` `govern` `section` `sec` scoped labels 2 months ago

Phil Calder added `wg-ai-integration` `wg-ai-integration-prioritized-prototype` labels 2 months ago

Phil Calder changed title from Feedback issue for: [AI Proposal] Explain this Vulnerability :mega to Feedback issue for: [AI Proposal] Explain this Vulnerability :mega: 2 months ago

Phil Calder mentioned in design management/design #407124 (closed)[design_1681421934437.png] 2 months ago

Becka Lippert mentioned in design management/design #407124 (closed)[design_1681485526732.png] 2 months ago

Alana Bellucci changed title from Feedback issue for: [AI Proposal] Explain this Vulnerability {->-} to {+>+}Feedback issue for: [AI Proposal] Mitigate Vulnerability with AI 1 month ago

Alana Bellucci changed the description 1 month ago · [Compare with previous version](#)

Becka Lippert mentioned in design management/design #407124 (closed)[design_1681420879588.png] 1 month ago

Becka Lippert mentioned in design management/design #407124 (closed)[design_1681846032408.png] 1 month ago

Becka Lippert changed title from Feedback issue for: [AI Proposal] Mitigate Vulnerability with AI to Feedback issue for: [AI Proposal] Explain Vulnerability with AI 1 month ago

Becka Lippert mentioned in design management/design #407124 (closed)[design_1681921959398.png] 1 month ago

Alana Bellucci made the issue visible to everyone 1 month ago

Responses generated by AI

How to setup and use Flux CD with GitLab CI

To setup and use Flux CD with GitLab CI, use the Flux command `bootstrap gitlab` to configure a Kubernetes cluster to manage itself from a Git repository. Authenticate your installation with either a project access token, a group access token, or a personal access token. To trigger immediate Git repository reconciliation, use the agent for Kubernetes with Flux. Some Flux features require write access to the source repositories.

Source:

[Flux \(Beta\)](#)

[Helpful](#)

[Unhelpful](#)

[Wrong](#)

Give me a link to learn how to use the kubernetes agent

Link:

<https://docs.gitlab.com/ee/user/project/clusters/agents/>

Source:

[Working with the agent for Kubernetes](#)

[Helpful](#)

[Unhelpful](#)

[Wrong](#)

You can ask AI for more information.



I'm a happy 🥳
CSM/CSE/Developer



GitLab Duo

The suite of AI capabilities powering your workflows



GitLab Duo: AI-assisted capabilities for everyone involved in the software development lifecycle



Developer Teams

Available features

Code Suggestions

AI paired programming

Suggested Reviewers

Better code reviews

Summarize MR Changes

Drive alignment and action

Summarize My MR Review

Get your point across

Help with Git Commands

Natural language CLI assistant



Security & Operations

Explain This Vulnerability

Remediate security issues

Generate Tests in MRs

Automate repetitive tasks

Explain This Code

Uplevel and understand



For Everyone

Issue Comment Summaries

Understand and take action

GitLab Chat

Get help fast

Value Stream Forecasting

Predict the future



Experimental

Beta

General Availability

GitLab AI Vision

AI-assisted features



GitLab Duo

Capabilities coming Soon

- Generate Planning Descriptions
- Natural Language Chart Generation
- Suggest A Fix For Failed Tests
- CI/CD Pipeline Failure Resolution
- Suspicious Activity Alerting
- Vulnerability Report Summary
- Merge Request Assistance
- Natural Language CI Configuration



GitLab AI Roadmap:

Next step: ModelOps

Stay tuned

Data Science

Empower Business transformation by allowing customers to build and integrate data science workloads within GitLab





Thank you for your attention





Question time

Only simple questions :)



