

# Should I Trust ChatGPT To Review My Program

**OCTOBER 4, 2023**

Mark Sherman  
Technical Director, Cybersecurity Foundations, CERT  
Telephone: +1 412.268.9223  
mssherman@cert.org



Copyright 2023 Carnegie Mellon University.

This material is based upon work funded and supported by the Department of Defense under Contract No. FA8702-15-D-0002 with Carnegie Mellon University for the operation of the Software Engineering Institute, a federally funded research and development center.

The view, opinions, and/or findings contained in this material are those of the author(s) and should not be construed as an official Government position, policy, or decision, unless designated by other documentation.

References herein to any specific commercial product, process, or service by trade name, trade mark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by Carnegie Mellon University or its Software Engineering Institute.

NO WARRANTY. THIS CARNEGIE MELLON UNIVERSITY AND SOFTWARE ENGINEERING INSTITUTE MATERIAL IS FURNISHED ON AN "AS-IS" BASIS. CARNEGIE MELLON UNIVERSITY MAKES NO WARRANTIES OF ANY KIND, EITHER EXPRESSED OR IMPLIED, AS TO ANY MATTER INCLUDING, BUT NOT LIMITED TO, WARRANTY OF FITNESS FOR PURPOSE OR MERCHANTABILITY, EXCLUSIVITY, OR RESULTS OBTAINED FROM USE OF THE MATERIAL. CARNEGIE MELLON UNIVERSITY DOES NOT MAKE ANY WARRANTY OF ANY KIND WITH RESPECT TO FREEDOM FROM PATENT, TRADEMARK, OR COPYRIGHT INFRINGEMENT.

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution. Please see Copyright notice for non-US Government use and distribution.

This material may be reproduced in its entirety, without modification, and freely distributed in written or electronic form without requesting formal permission. Permission is required for any other use. Requests for permission should be directed to the Software Engineering Institute at [permission@sei.cmu.edu](mailto:permission@sei.cmu.edu).

Carnegie Mellon® and CERT® are registered in the U.S. Patent and Trademark Office by Carnegie Mellon University.

DM23-1070

# CMU Software Engineering Institute (SEI)

Bringing innovation to the U.S. Government



- CMU 2022-2023 *U.S. News and World Report* rankings:
  - #1 in artificial intelligence, computer engineering, cybersecurity, management information systems, mobile/web applications, programming languages, software engineering, and quantitative analysis
  - #1 in overall computer science
- Federally Funded Research and Development Center (FFRDC) chartered in 1984 and sponsored by the DoD
- Leader in researching complex software engineering, cyber security, and artificial intelligence (AI) engineering solutions
- Critical to the U.S. Government's ability to acquire, develop, operate, and sustain software systems that are innovative, affordable, trustworthy, and enduring
- Growing depth in data science, statistics, machine learning (LLM) and artificial intelligence (see booth)

# Lots of Hype about Using ChatGPT for Coding

## Okay, so ChatGPT just debugged my code. For real.

Not only can ChatGPT write code, it can read code. On one hand, that's very helpful. On the other hand, that's truly terrifying.



Written by David Geelitz, Senior Contributing Editor on April 3, 2023

<https://www.zdnet.com/article/okay-so-chatgpt-just-debugged-my-code-for-real/>

## How to use ChatGPT to write code

By Allen Helton · March 22, 2023

**Greetings, developers!** Have you ever had a case of "coder's block" where you need a little help or inspiration to write code? Or need some help understanding code that you inherited from someone else? Well, you're in luck! From generating boilerplate code to debugging or explaining existing code, ChatGPT is a no-brainer way to be a faster, more efficient software engineer.

OpenAI's ChatGPT falls into the category of **generative AI**, which is a type of artificial intelligence that generates new or novel content based on data it's been trained on. That means ChatGPT can not only write essays, jokes and poems for you, it can also be your pair programmer.

So let's dive in and see how you can use it to work smarter, not harder.

<https://www.pluralsight.com/blog/software-development/how-use-chatgpt-programming-coding>

## How good is ChatGPT at writing code?

Learn more about ChatGPT and discover how it can be used to generate programming solutions.

Are you looking for a new way to write code? ChatGPT is a revolutionary AI-based system that can help you generate programming solutions quickly and easily. In this article, we'll explain how effective ChatGPT is at writing code and whether or not it really lives up to its promises.

<https://botpress.com/blog/how-good-is-chatgpt-at-writing-code>

AI

## ChatGPT Changed How I Write Software

By Allen Helton · 31 May 2023

AI is buzzing right now. All my social media feeds are about some new thing you can do with ChatGPT plugins or showing off a new photo created by generative AI. To be honest, it's pretty cool to see and I'm not mad about it.

<https://www.readyssetcloud.io/blog/allen.helton/chatgpt-changed-how-i-write-software/>

# CERT Secure Coding Standards



Collected wisdom from thousands of contributors on community wiki since Spring 2006

<http://securecoding.cert.org>

- SEI CERT C Coding Standard
- SEI CERT C++ Coding Standard
- CERT Oracle Secure Coding Standard for Java



Secure Coding Training and Professional Certificates

- CERT Secure Coding in C and C++
- CERT Secure Coding in Java

International Standards Participation

- ISO/IEC C Programming Language
- ISO/IEC C++ Programming Language

# Rule Organization – Noncompliant Code Example

## Noncompliant Code Example

In this noncompliant code example, `memcmp()` is used to compare the contents of two structures, including any padding bytes:

```
#include <string.h>

struct s {
    char c;
    int i;
    char buffer[13];
};

void compare(const struct s *left, const struct s *right) {
    if ((left && right) &&
        (0 == memcmp(left, right, sizeof(struct s)))) {
        /* ... */
    }
}
```

Noncompliant Code  
*Don't try this at home!*

Noncompliant code examples or antipatterns in a pink frame—do not copy and paste into your code. Most examples were found in the wild.

# Rule Organization – Compliant Solution

## Compliant Solution

In this compliant solution, all of the fields are compared manually to avoid comparing any padding bytes:

```
#include <string.h>

struct s {
    char c;
    int i;
    char buffer[13];
};

void compare(const struct s *left, const struct s *right) {
    if ((left && right) &&
        (left->c == right->c) &&
        (left->i == right->i) &&
        (0 == memcmp(left->buffer, right->buffer, 13))) {
        /* ... */
    }
}
```

Compliant solutions in a blue frame that conform with all rules and can be reused in your code

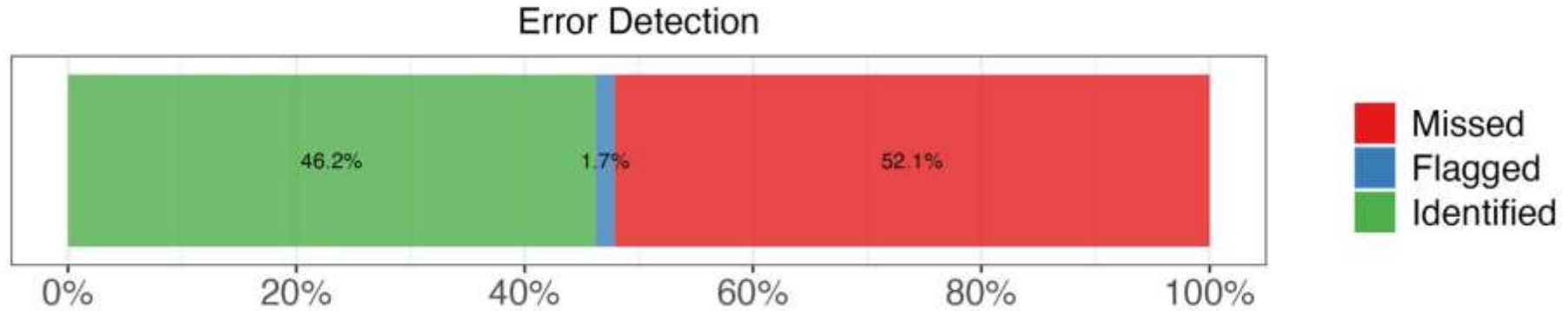
Compliant Code  
*Fixes noncompliant code.*

# Experiment

Ask ChatGPT to identify errors in 238 examples of noncompliant code from CERT C Secure Coding Standard

- ChatGPT 3.5 as of March 27, 2023
- Each trial run as a new conversation
- No trial repeated
- All examples (with solutions) available on the web during ChatGPT's training data capture
- Some examples have comments suggesting error
- ChatGPT's performance evaluated by SMEs

# Overall Results – Finding Errors

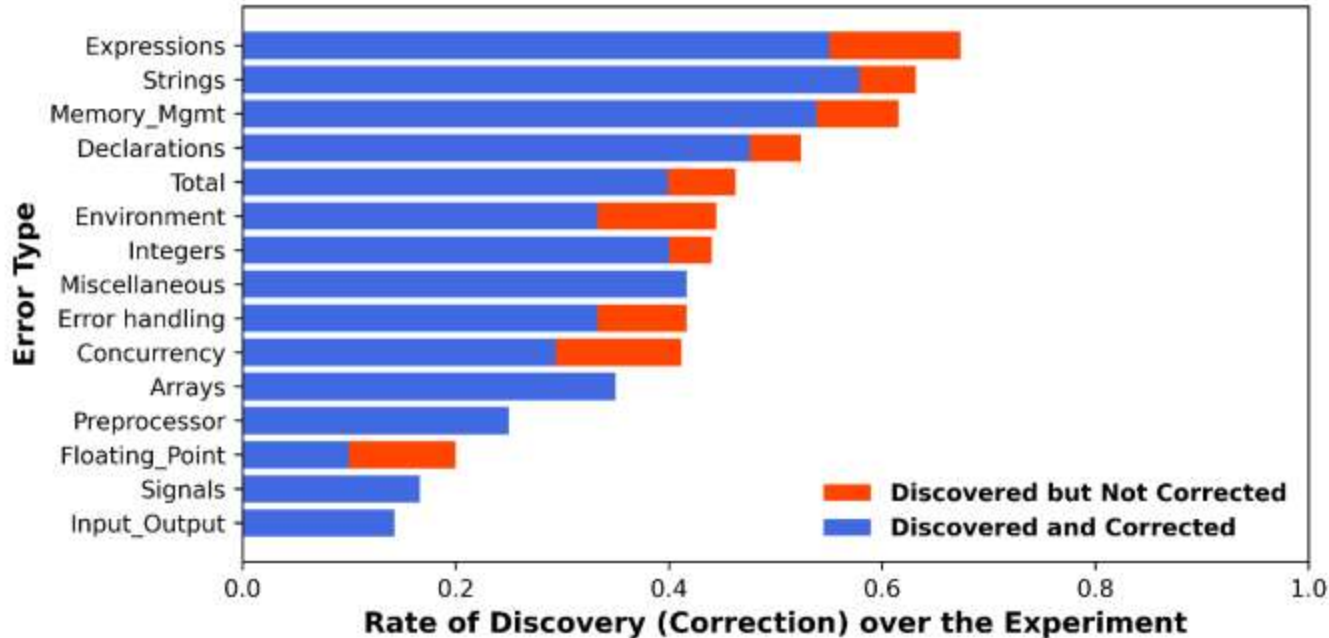


Note: Underlying data in backup

# Overall Results – Fixing Errors



# Overall Results – By Feature Tested



## Seems particularly good at:

- Fixing Integers
- Finding/Fixing Expressions
- Finding/Fixing Memory Management
- Finding/Fixing Strings

## Seems particularly challenged by:

- Finding Floating point
- Finding Input/Output
- Finding Signals

It's not that ChatGPT did not find all the errors .....



"He's not all that smart. I beat him most of the time."

CartoonStock.com

It's that it found as many as it did.

# Summary

- Experiment illustrates promise but also limitations
- Like many applications of LLMs, knowledgeable users must review output
- Unfortunately, programmers are not very good at reading and evaluating code
- ChatGPT certainly should not be used to replace programmers, and any other application should be monitored with great care

# Acknowledgements



**Robert Schiela**  
Secure Coding Technical Manager

Telephone: +1 412.268.3965  
[rschiela@cert.org](mailto:rschiela@cert.org)



**David Svoboda**  
Software Security Engineer

Telephone: +1 412.268.3965  
[Svoboda@cert.org](mailto:Svoboda@cert.org)



**Matthew Walsh**  
Senior Data Scientist

Telephone: +1 412.268.9121  
[mwalsh@cert.org](mailto:mwalsh@cert.org)



**Clarence Worrell**  
Senior Data Scientist

Telephone: +1 412.268.9059  
[cworrell@cert.org](mailto:cworrell@cert.org)



**Jamie Glenn**  
Operations Coordinator

Telephone: +1 412.268.3346  
[jglenn@cert.org](mailto:jglenn@cert.org)

I want to thank the following members of the Cybersecurity Foundations group at CERT for their expert assistance with evaluating and preparing these materials:

Jamie Glenn	Experiment management
Robert Schiela	C SME
David Svoboda	C SME (Representative, ISO/IEC JTC1/SC22/WG14 (C Programming Language) and lead of Undefined Behaviors Study Group (UBSG))
Matt Walsh	Data analysis and visualization
Clarence Worrell	Data analysis and visualization