AI-powered reverse-engineering

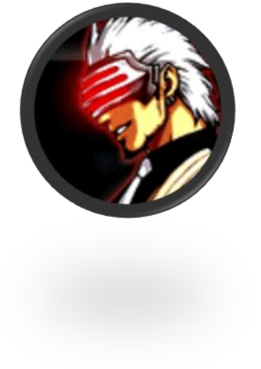
Gepetto

kaspersky

# whoami

Ivan Kwiatkowski

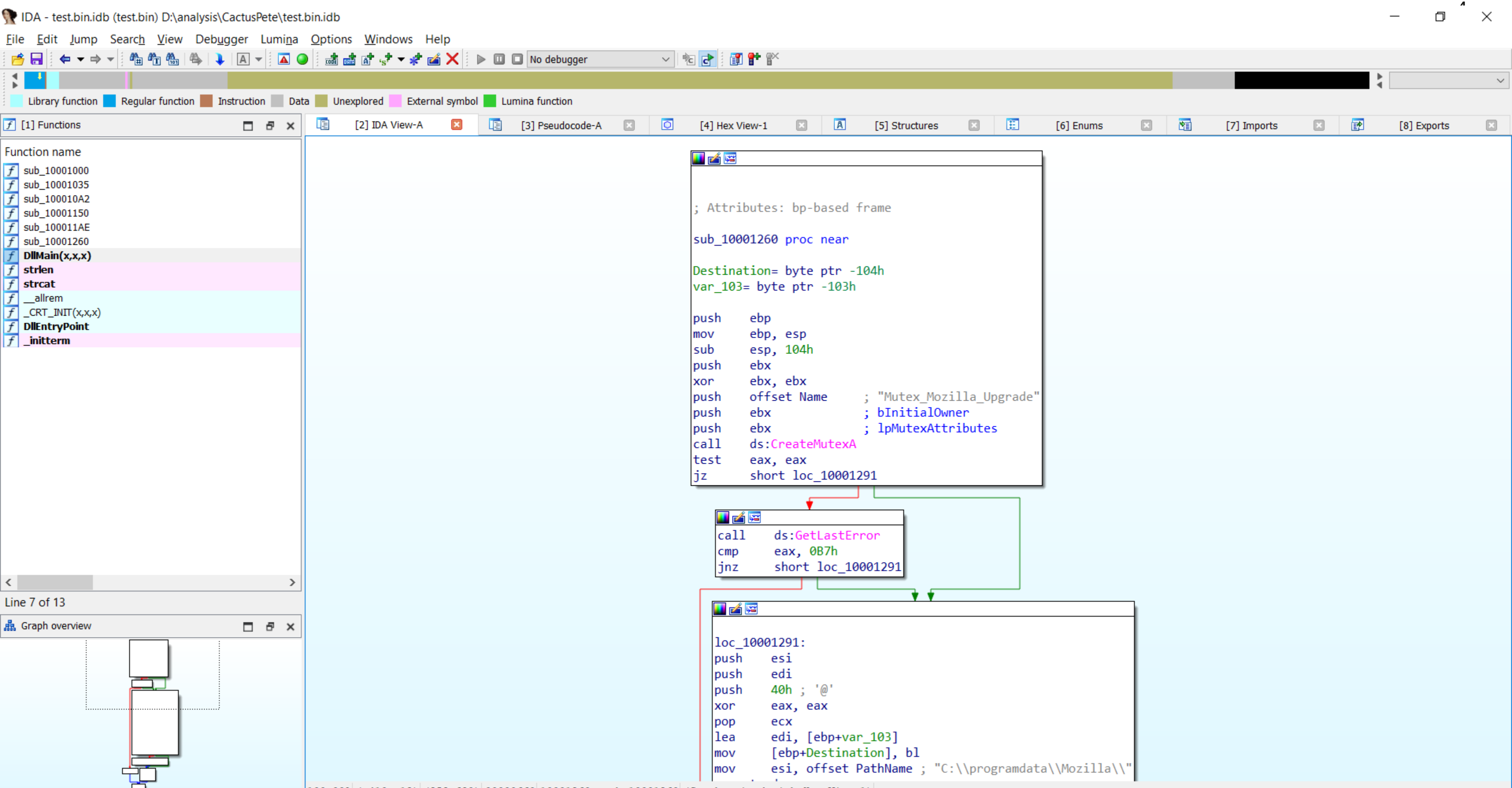
- Senior Security Researcher @ Kaspersky
- Threat Intelligence
- Reverse engineering
- @JusticeRage



~~December 2022~~

The month of “can ChatGPT do my job?”

What is my job?



```
Function name
sub_10001000
sub_10001035
sub_100010A2
sub_10001150
sub_100011AE
sub_10001260
DllMain(x,x,x)
strlen
strcat
__allrem
_CRT_INIT(x,x,x)
DllEntryPoint
_initterm
```

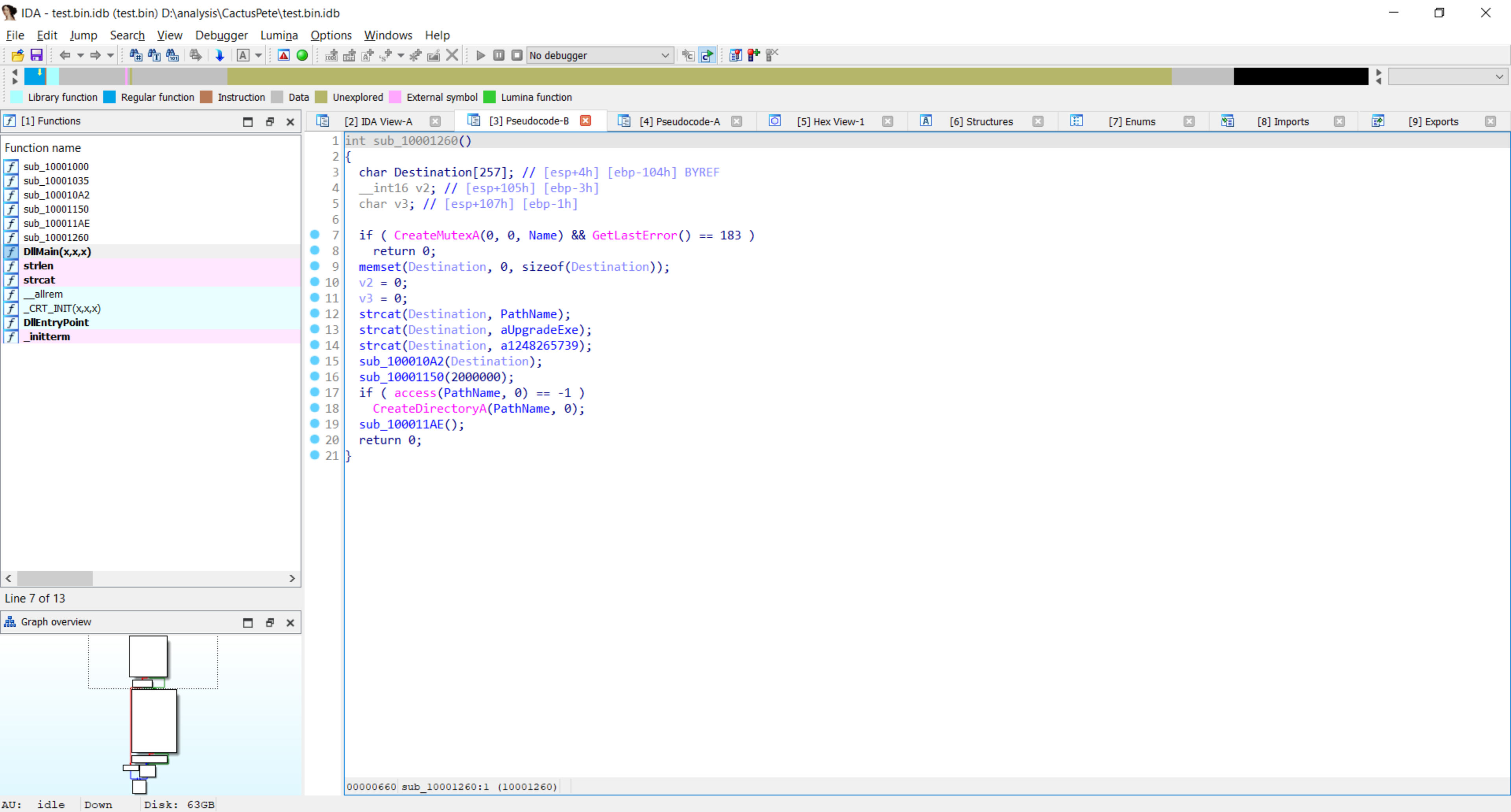
Line 7 of 13

```
; Attributes: bp-based frame

sub_10001260 proc near

Destination= byte ptr -104h
var_103= byte ptr -103h

push    ebp
mov     ebp, esp
sub     esp, 104h
push    ebx
xor     ebx, ebx
push    offset Name     ; "Mutex_Mozilla_Upgrade"
push    ebx             ; bInitialOwner
push    ebx             ; lpMutexAttributes
call    ds:CreateMutexA
test    eax, eax
jz      short loc_10001291
```

```
call    ds:GetLastError
cmp     eax, 0B7h
jnz     short loc_10001291
```

```
loc_10001291:
push    esi
push    edi
push    40h ; '@'
xor     eax, eax
pop     ecx
lea     edi, [ebp+var_103]
mov     [ebp+Destination], bl
mov     esi, offset PathName ; "C:\\programdata\\Mozilla\\"
```

:: OFFICIAL 4CHAN /b/ KEYBOARD ::
F5
C
V
Enter
Page Up
Page Down



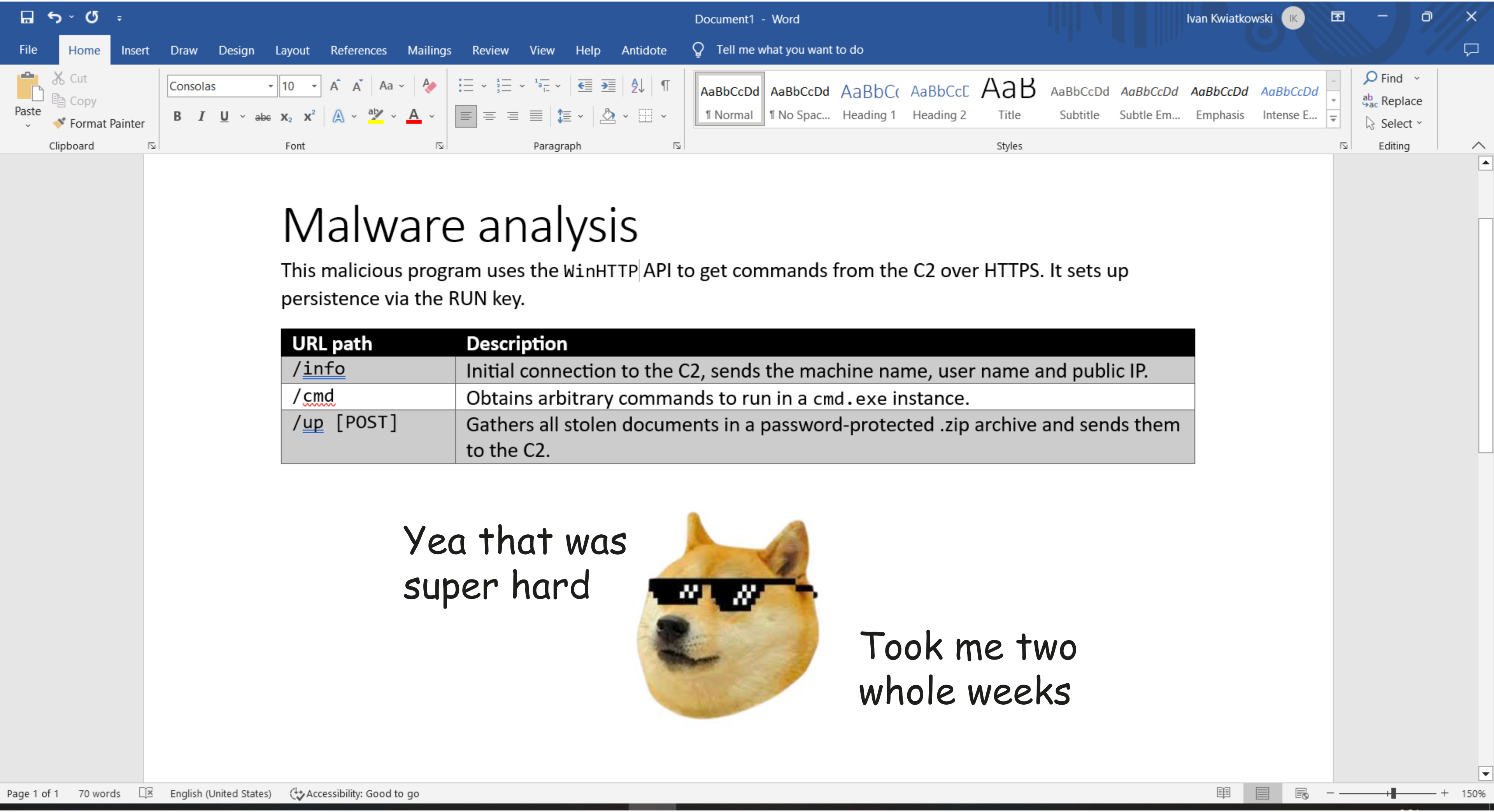
```
Function name
sub_10001000
sub_10001035
sub_100010A2
sub_10001150
sub_100011AE
sub_10001260
DllMain(x,x,x)
strlen
strcat
__allrem
_CRT_INIT(x,x,x)
DllEntryPoint
_initterm
```

Line 7 of 13

```
int sub_10001260()
{
  char Destination[257]; // [esp+4h] [ebp-104h] BYREF
  __int16 v2; // [esp+105h] [ebp-3h]
  char v3; // [esp+107h] [ebp-1h]

  if ( CreateMutexA(0, 0, Name) && GetLastError() == 183 )
    return 0;
  memset(Destination, 0, sizeof(Destination));
  v2 = 0;
  v3 = 0;
  strcat(Destination, PathName);
  strcat(Destination, aUpgradeExe);
  strcat(Destination, a1248265739);
  sub_100010A2(Destination);
  sub_10001150(2000000);
  if ( access(PathName, 0) == -1 )
    CreateDirectoryA(PathName, 0);
  sub_100011AE();
  return 0;
}
```

```
00000660 sub_10001260:1 (10001260)
```

```
AU:  idle   Down    Disk: 63GB
```



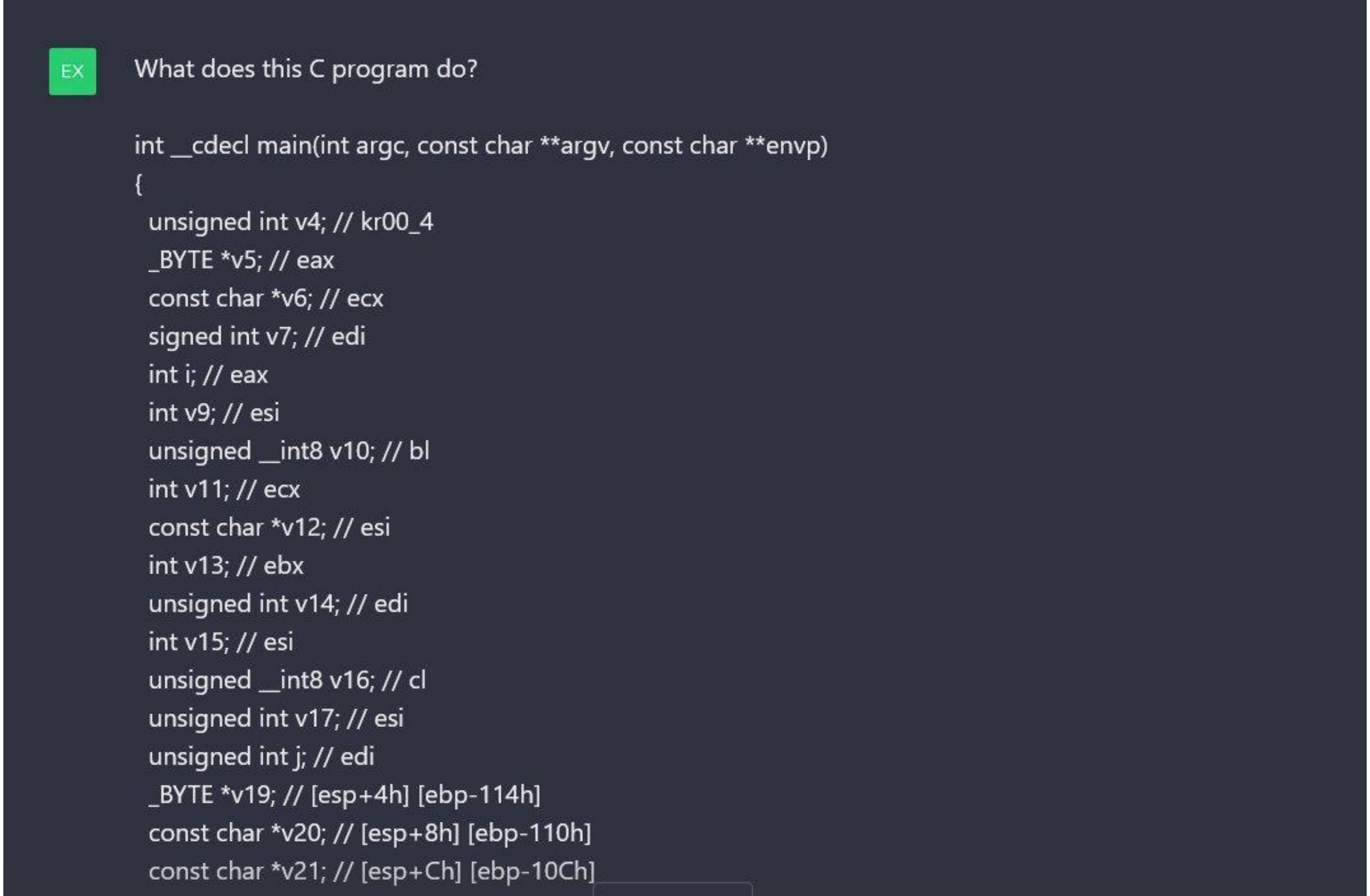
Malware analysis

This malicious program uses the `WinHTTP` API to get commands from the C2 over HTTPS. It sets up persistence via the RUN key.

URL path	Description
`/info`	Initial connection to the C2, sends the machine name, user name and public IP.
`/cmd`	Obtains arbitrary commands to run in a `cmd.exe` instance.
`/up [POST]`	Gathers all stolen documents in a password-protected .zip archive and sends them to the C2.

Yea that was super hard

Took me two whole weeks



What does this C program do?

```
int __cdecl main(int argc, const char **argv, const char **envp)
{
  unsigned int v4; // kr00_4
  _BYTE *v5; // eax
  const char *v6; // ecx
  signed int v7; // edi
  int i; // eax
  int v9; // esi
  unsigned __int8 v10; // bl
  int v11; // ecx
  const char *v12; // esi
  int v13; // ebx
  unsigned int v14; // edi
  int v15; // esi
  unsigned __int8 v16; // cl
  unsigned int v17; // esi
  unsigned int j; // edi
  _BYTE *v19; // [esp+4h] [ebp-114h]
  const char *v20; // [esp+8h] [ebp-110h]
  const char *v21; // [esp+Ch] [ebp-10Ch]
```

```
v24[v15] = v16;
...2[v20 - v19] ^ v24[(unsigned __int8)(v
while ( v14
```

...ndom generation algorithm (...

...encrypt the given string. Finally, it prints the encrypted strin...

hexadecimal format.

Ali Hadi | B!n@ry @binaryz0ne · Dec 3, 2022

This made me signup!

downbtn @downbtn_ · Dec 2, 2022

Does this mean all the hours I spent as a kid reading assembly code is wasted? lol

@mikko @mikko · Dec 2, 2022

WHAT

Why does it work???

What LLMs are good at	What we have
Processing language (i.e., summarizing, explaining…)	• Pseudo-C • Programming **language** • Strict rules
Processing code	• Trained on all of GitHub • GitHub copilot *works*
Translation	• Convert C to English • DeepL *works*

Obvious next step

Write an IDA plugin that automatically queries ChatGPT

Demo time

kaspersky

Features

- Analyze functions / rename variables

  But is it useful?

  Look for security vulnerabilities?

- Multi-model support (GPT-3.5, GPT-4)

  More models waiting in a PR (ChatSonic, Bing…)

- Multi-language support (EN, FR, ES, IT, KO, CN…)

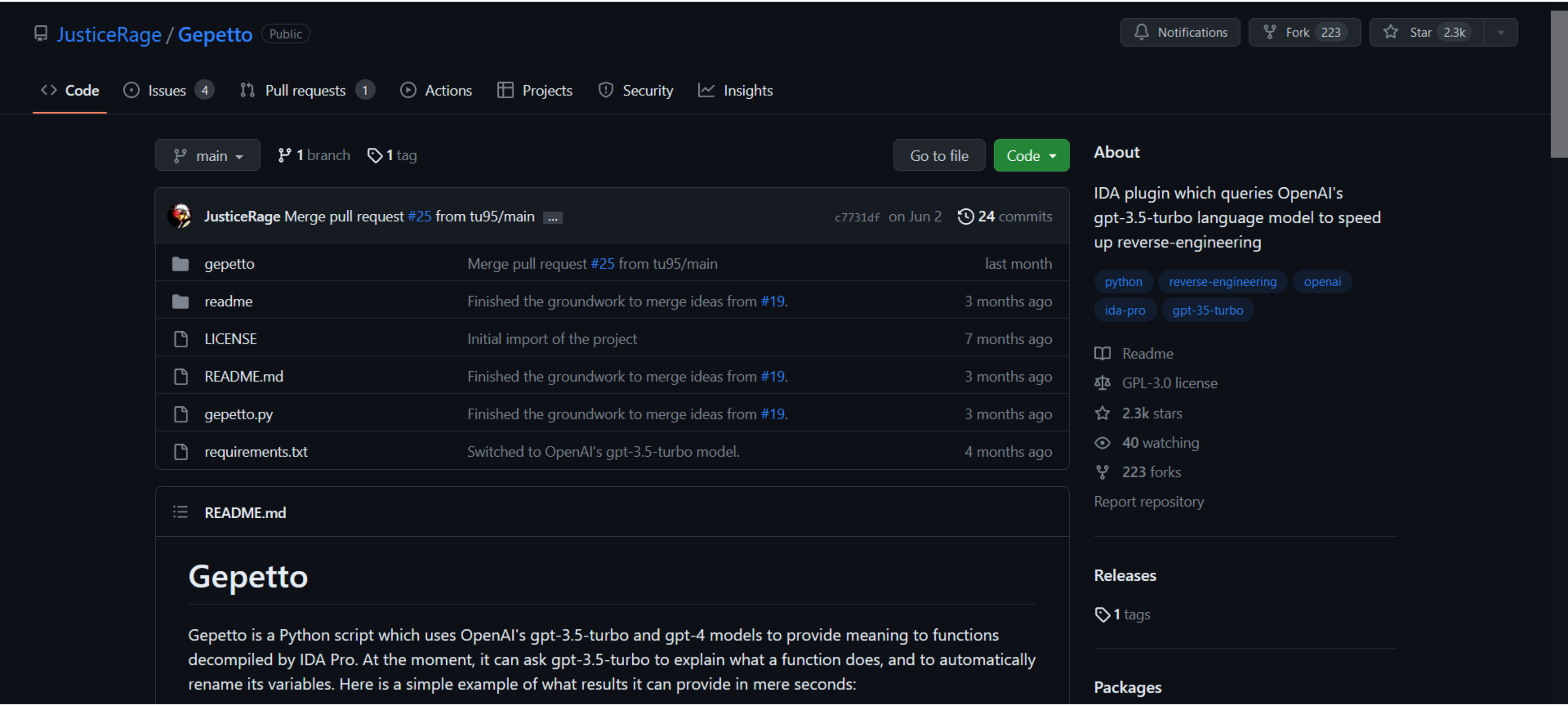
  https://www.transifex.com/gepetto/

  (If you don’t help I’ll make ChatGPT or DeepL do it)

Limitations

- API costs: ~$3 / reversing session
- Token limit

FOSS (GPL 3.0)



JusticeRage Merge pull request #25 from tu95/main ... c7731df on Jun 2 24 commits

gepetto	Merge pull request #25 from tu95/main	last month
readme	Finished the groundwork to merge ideas from #19.	3 months ago
LICENSE	Initial import of the project	7 months ago
README.md	Finished the groundwork to merge ideas from #19.	3 months ago
gepetto.py	Finished the groundwork to merge ideas from #19.	3 months ago
requirements.txt	Switched to OpenAI's gpt-3.5-turbo model.	4 months ago

README.md

Gepetto

Gepetto is a Python script which uses OpenAI's gpt-3.5-turbo and gpt-4 models to provide meaning to functions decompiled by IDA Pro. At the moment, it can ask gpt-3.5-turbo to explain what a function does, and to automatically rename its variables. Here is a simple example of what results it can provide in mere seconds:

About

IDA plugin which queries OpenAI's gpt-3.5-turbo language model to speed up reverse-engineering

python reverse-engineering openai ida-pro gpt-35-turbo

Readme
GPL-3.0 license
2.3k stars
40 watching
223 forks
Report repository

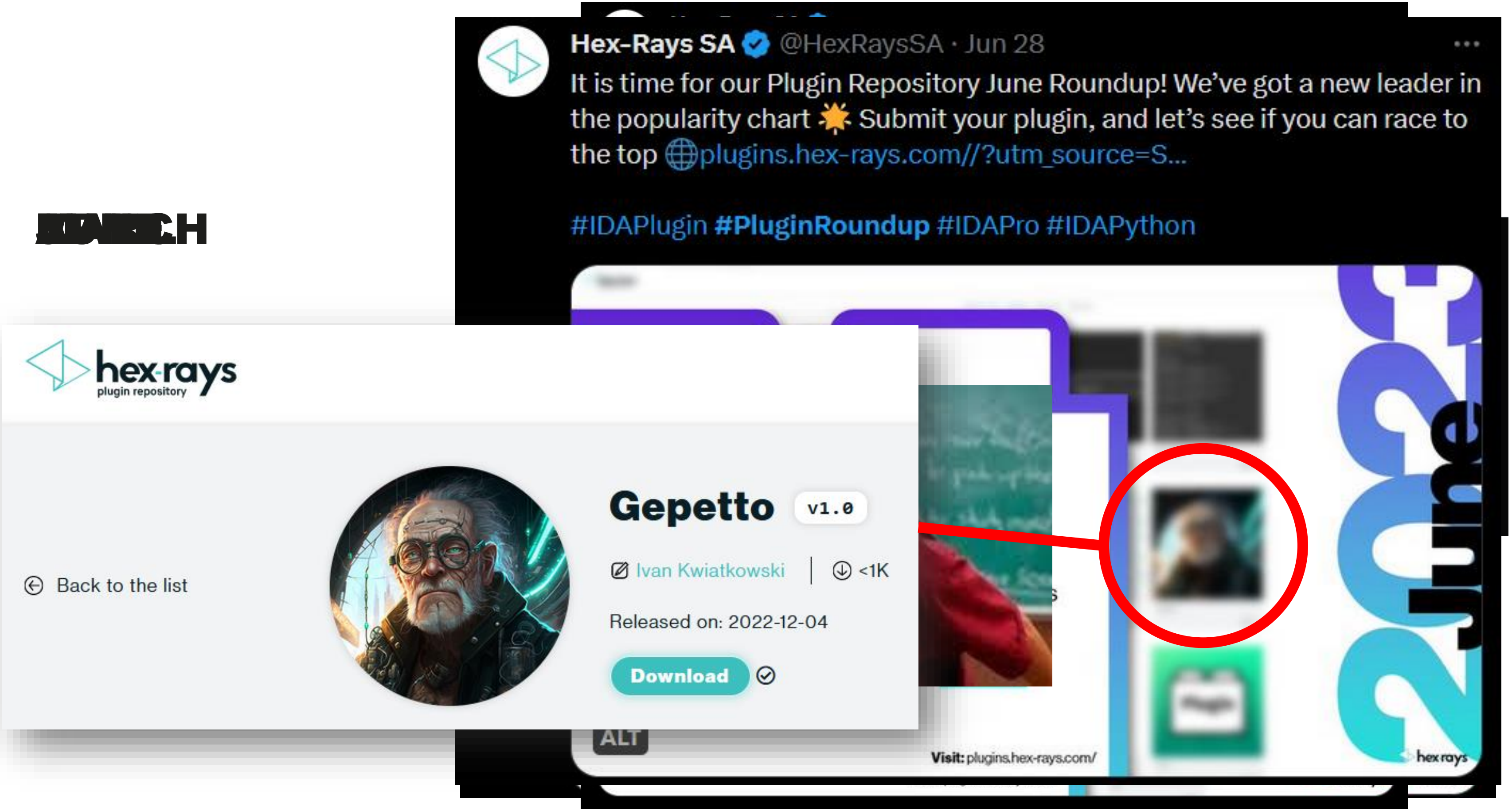
Releases

1 tags

Packages

Adoption

[illegible]



Hex-Rays SA @HexRaysSA · Jun 28

It is time for our Plugin Repository June Roundup! We've got a new leader in the popularity chart 🌟 Submit your plugin, and let's see if you can race to the top 🌐plugins.hex-rays.com//?utm_source=S...

#IDAPlugin **#PluginRoundup** #IDAPro #IDAPython

hex-rays plugin repository

Back to the list

Gepetto v1.0

Ivan Kwiatkowski | <1K

Released on: 2022-12-04

The future

- GPT-4 is a general purpose model.

    Can we fine-tune LLaMA for RE?
    How to evaluate the output of the LLM?

- Recursive program analysis

    https://github.com/moyix/gpt-wpre
    Quite expensive to test (~$30 / attempt)
    Risk of compounding errors

- Train dedicated models

    ~~Image~~ function recognition
    Right-click > What C++ STL function is this? > kthxbye

The end!

kaspersky