
Web Defacement Scanner

Program Document

Version 1.0 approved

29 MAY 2024

Prepared By ITSC CSOC Internship 2024

Setsiri Matewin 630610766

Phanuwat Ngoenthok 640610659

CPE.ENG.CMU

สารบัญ

Web Defacement Scanner.....	1
29 MAY 2024.....	1
Setsiri Matewin 630610766.....	1
Phanuwat Ngoenthok 640610659.....	1
สารบัญ.....	2
Introduction.....	3
ชื่อโปรแกรม : Web Defacement Scanner.....	3
เวอร์ชันของโปรแกรม : 1.0.....	3
วันที่เผยแพร่ : 29 พฤษภาคม พ.ศ. 2567.....	3
ระบบปฏิบัติการ : Window.....	3
วัตถุประสงค์ของโปรแกรม.....	3
ภาพรวมของการทำงานของโปรแกรม.....	3
ผู้ใช้เป้าหมาย.....	3
Scope.....	4
Installation and Configuration.....	4
User Guide.....	5
หน้าตาของโปรแกรมเมื่อเริ่มต้น.....	5
ส่วนประกอบและฟังก์ชันของโปรแกรมในหน้าเริ่มต้น.....	5
หน้าตาของโปรแกรมเมื่อเริ่มสแกน.....	6
ส่วนประกอบและฟังก์ชันของโปรแกรมในหน้าเริ่มสแกน.....	6
หน้าตาของโปรแกรมเมื่อสิ้นสุดการสแกน.....	7
ส่วนประกอบและฟังก์ชันของโปรแกรมในหน้าสิ้นสุดการสแกน.....	7
Technical Details.....	8
Libraries and Imports.....	8
1. Standard Libraries:.....	8
2. Third-party Libraries:.....	8
Global Variables.....	8
Functions.....	8
Core Functionalities.....	8
GUI Components.....	9
Scanning and Defacement Detection.....	9
File Handling.....	9
Thread Management.....	9
Progress and Status Updates.....	9
Execution Flow.....	9
Troubleshooting.....	10
License and Terms.....	11
Appendices.....	12
Appendix A: References.....	12
Appendix B: Sample Code.....	12

Introduction

ชื่อโปรแกรม : Web Defacement Scanner

เวอร์ชันของโปรแกรม : 1.0

วันที่เผยแพร่ : 29 พฤษภาคม พ.ศ. 2567

ระบบปฏิบัติการ : Window

วัตถุประสงค์ของโปรแกรม

จุดประสงค์ของโปรแกรมนี้คือการพัฒนาและใช้งานเครื่องมือสแกนเนอร์สำหรับการตรวจจับการแฮกและการแก้ไขหน้าเว็บ (web defacement) โดยผู้ใช้สามารถป้อน URL และคีย์เวิร์ดที่ต้องการตรวจสอบ เมื่อโปรแกรมได้รับข้อมูลดังกล่าวแล้ว โปรแกรมจะทำการรวบรวมและวิเคราะห์เนื้อหาของเว็บไซต์ที่ระบุ โดยใช้กระบวนการการรวบรวมข้อมูลแบบต่อเนื่อง (crawling) ไปยังลิงก์ต่าง ๆ ภายในเว็บไซต์นั้น โปรแกรมจะทำการตรวจจับและรายงานผลการตรวจพบการเปลี่ยนแปลงหรือการแทรกแซงหน้าเว็บที่อาจเป็นอันตราย เช่น การฝังลิงก์ไปยังเว็บไซต์พนันหรือเนื้อหาที่ไม่เหมาะสม เป้าหมายหลักของโปรแกรมนี้คือการช่วยผู้ดูแลระบบและผู้ใช้ทั่วไปในการรักษาความปลอดภัยของเว็บไซต์จากการถูกโจมตีและการเปลี่ยนแปลงที่ไม่พึงประสงค์

ภาพรวมของการทำงานของโปรแกรม

โปรแกรมนี้ถูกพัฒนาโดยใช้ภาษา Python โดยมีการนำไลบรารี BeautifulSoup มาใช้สำหรับการรวบรวมและวิเคราะห์ข้อมูลจากเว็บไซต์ และใช้ Tkinter สำหรับการสร้างอินเทอร์เฟซกราฟิกที่ใช้งานง่ายให้กับผู้ใช้ เมื่อผู้ใช้ป้อน URL และคีย์เวิร์ดที่ต้องการตรวจสอบลงในโปรแกรม โปรแกรมจะเริ่มกระบวนการรวบรวมข้อมูลโดยใช้ BeautifulSoup เพื่อดึงข้อมูลจากหน้าเว็บและลิงก์ภายในเว็บไซต์ จากนั้นโปรแกรมจะทำการสแกนหาเนื้อหาที่เกี่ยวข้องกับคีย์เวิร์ดที่ระบุ หากพบเนื้อหาที่เข้าข่ายการแฮกหรือการแทรกแซง เช่น ลิงก์ไปยังเว็บไซต์พนันตามคีย์เวิร์ดที่กำหนดไว้ โปรแกรมจะรายงานผลการตรวจพบให้กับผู้ใช้ผ่านอินเทอร์เฟซที่จัดทำด้วย Tkinter นอกจากนี้ โปรแกรมยังถูกพัฒนาเป็นไฟล์ .exe เพื่อให้ผู้ใช้สามารถรันโปรแกรมได้โดยง่าย เพียงแค่ดับเบิลคลิกไฟล์ ก็สามารถเริ่มต้นการสแกนและตรวจสอบเว็บไซต์ได้ทันที

ผู้ใช้เป้าหมาย

- ผู้ดูแลระบบเว็บไซต์ (Web Administrators): ผู้ที่มีหน้าที่ดูแลและจัดการเว็บไซต์ ต้องการตรวจสอบความปลอดภัยและป้องกันการแฮกหรือการแทรกแซงที่ไม่พึงประสงค์
- นักพัฒนาเว็บ (Web Developers): ผู้ที่พัฒนาเว็บไซต์และต้องการเครื่องมือในการตรวจสอบความปลอดภัยของเว็บแอปพลิเคชันที่พัฒนาขึ้น
- ผู้เชี่ยวชาญด้านความปลอดภัยไซเบอร์ (Cybersecurity Experts): ผู้ที่ทำงานด้านความปลอดภัยทางไซเบอร์ ต้องการเครื่องมือในการตรวจสอบและวิเคราะห์การโจมตีเว็บไซต์

4.เจ้าของธุรกิจออนไลน์ (Online Business Owners): ผู้ที่ดำเนินธุรกิจผ่านเว็บไซต์และต้องการรักษาความปลอดภัยของข้อมูลและความน่าเชื่อถือของเว็บไซต์

5.นักวิจัยด้านความปลอดภัย (Security Researchers): ผู้ที่ศึกษาด้านความปลอดภัยทางไซเบอร์และต้องการเครื่องมือในการตรวจสอบการแสกและการแก้ไขหน้าเว็บเพื่อการวิจัยและพัฒนา

6.ผู้ให้บริการโฮสติ้ง (Hosting Providers): ผู้ให้บริการโฮสติ้งที่ต้องการตรวจสอบความปลอดภัยของเว็บไซต์ที่อยู่นบนเซิร์ฟเวอร์ของตน

Scope

Web Defacement Scanner จัดทำเพื่อให้ผู้ดูแลระบบเว็บไซต์ นักพัฒนาเว็บ หรือเจ้าของธุรกิจออนไลน์ โดย Web Defacement Scanner จะนำ url ที่ user เขียนเข้ามาไป crawling ตาม HTML content ที่สามารถ fetch ได้ และจะ crawling เพียงครั้งที่ขึ้นต้นด้วย url ตามที่ user ได้เขียนเข้ามา

Installation and Configuration

การติดตั้งใช้งาน ผู้ใช้สามารถดาวน์โหลดตัวโปรแกรมในรูปแบบของ Application (.exe) ได้จาก Github ของโปรแกรม หรือทำการโคลนตัว Git Repository ในกรณีที่ต้องการ Source code ทั้งหมดได้จาก ลิ้งนี้ https://github.com/nunu111/web_defacement_scanner

โดยในกรณีที่รันโปรแกรมจากโค้ด ผู้ใช้ต้องลงซอฟต์แวร์ ไลบรารี ที่เกี่ยวข้องเสียก่อน ซึ่งได้แก่

1. IDE software เช่น VS code
2. ภาษาเขียนโปรแกรม python (แนะนำ python ver 3.12.3)
3. BeautifulSoup (python library) use command : pip install beautifulsoup4
4. Requests (python library) use command : pip install requests
5. urllib (python library) use command : pip install urllib

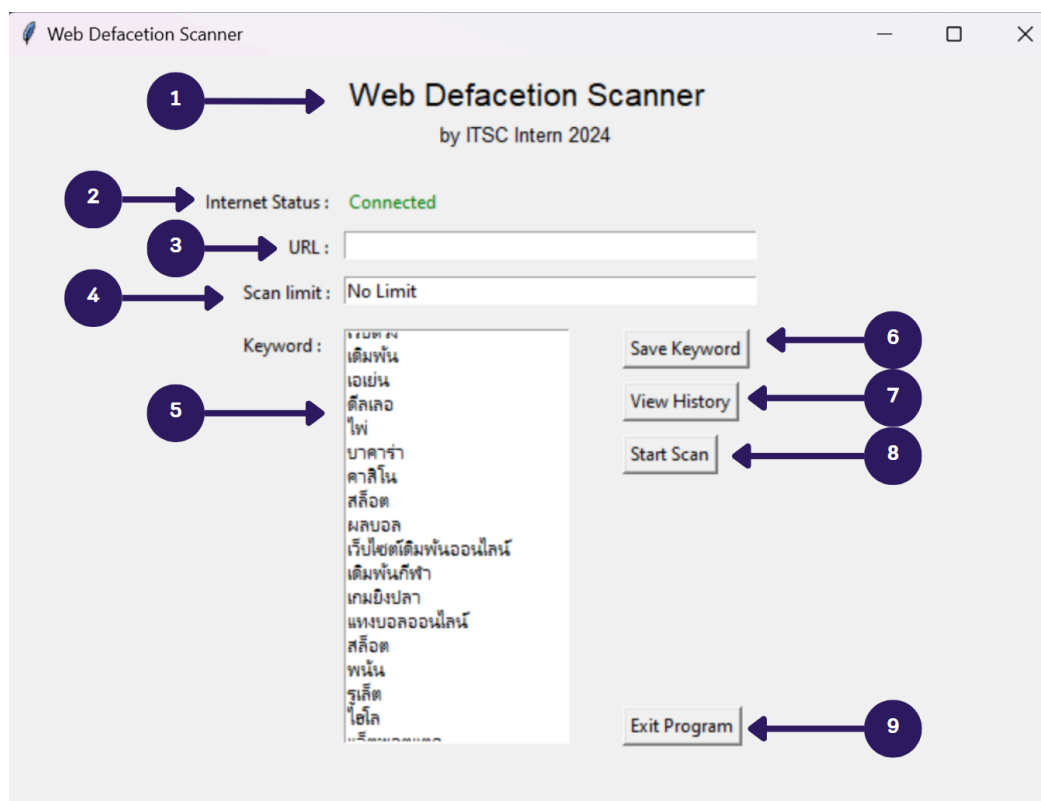
ในดั่งต้นจะมีเพียงตัวโปรแกรมอย่างเดียว และเพื่อการใช้โปรแกรมจะทำการสร้าง ไฟล์ keyword.txt และ โฟลเดอร์สำหรับเก็บ result ชื่อ History เมื่อโปรแกรมตรวจพบว่ายังไม่มีไฟล์ดังกล่าว โดยจะสร้างอยู่ใน path เดียวกันที่โปรแกรมนี้อยู่

ผลลัพธ์การรันโปรแกรมจะถูกสร้างเก็บไว้เป็น ไฟล์ .txt โดยอยู่ในรูปแบบการสร้างดังนี้ URL-DATE.txt โดยจะถูกรวมอยู่ในโฟลเดอร์ History

โดยสรุปตัวโปรแกรมมี 3 ส่วนหลักนั่นคือ ตัวโปรแกรม (.pyหรือ.exe) ตัวไฟล์ keyword.txt และ โฟลเดอร์ History นั่นคือเมื่อต้องการลบโปรแกรม ผู้ใช้สามารถ Delete 3 ส่วนนี้ ได้ตามความต้องการ

User Guide

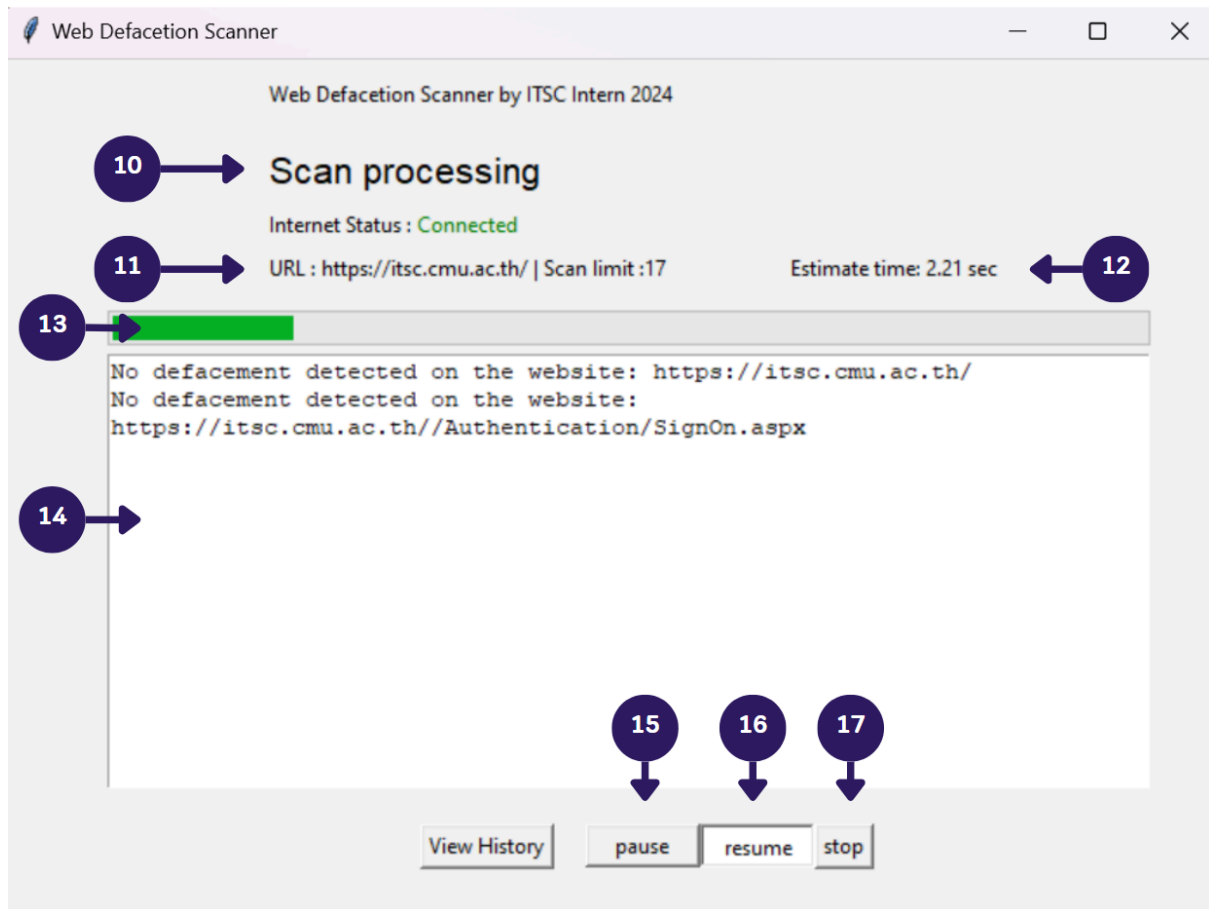
หน้าตาของโปรแกรมเมื่อเริ่มต้น



ส่วนประกอบและฟังก์ชันของโปรแกรมในหน้าเริ่มต้น

1. ชื่อโปรแกรม
2. ฟังก์ชันแสดงสถานะการเชื่อมต่ออินเทอร์เน็ตของเครื่องคอมพิวเตอร์ของผู้ใช้
3. ส่วนรับ Input จากผู้ใช้เป็น URL เดิมรูปแบบ เช่น <https://itsc.cmu.ac.th>
4. Scan limit เมื่อต้องการการแสกนที่จำกัดจำนวนการ crawling ให้ใส่เป็นเลขจำนวนเต็มบวก ถ้าปล่อยว่างจะขึ้นว่า No limit ตัวโปรแกรมจะทำการแสกนหมดจนหมดเท่าที่สามารถไปได้
5. หน้าต่าง text editor ของไฟล์ keyword.txt ผู้ใช้สามารถแก้ไขไฟล์ได้โดยจะอยู่ในรูปแบบของหนึ่งคีย์เวิร์ดต่อหนึ่งบรรทัด
6. ปุ่มฟังก์ชัน ชื่อ Save Keyword เมื่อกดแล้วจะบันทึกไฟล์ keyword.txt ตามข้อมูลที่อยูบนหน้าต่าง text editor
7. ปุ่มฟังก์ชัน ชื่อ View history เมื่อกดจะเปิดตัว History โฟลเดอร์ขึ้นมา
8. ปุ่มฟังก์ชัน ชื่อ Start Scan เมื่อกดแล้วจะเริ่มทำการแสกนเข้าสู่หน้าต่างของการ processing
9. ปุ่มฟังก์ชัน ชื่อ Exit Program เมื่อกดจะทำการปิดโปรแกรม

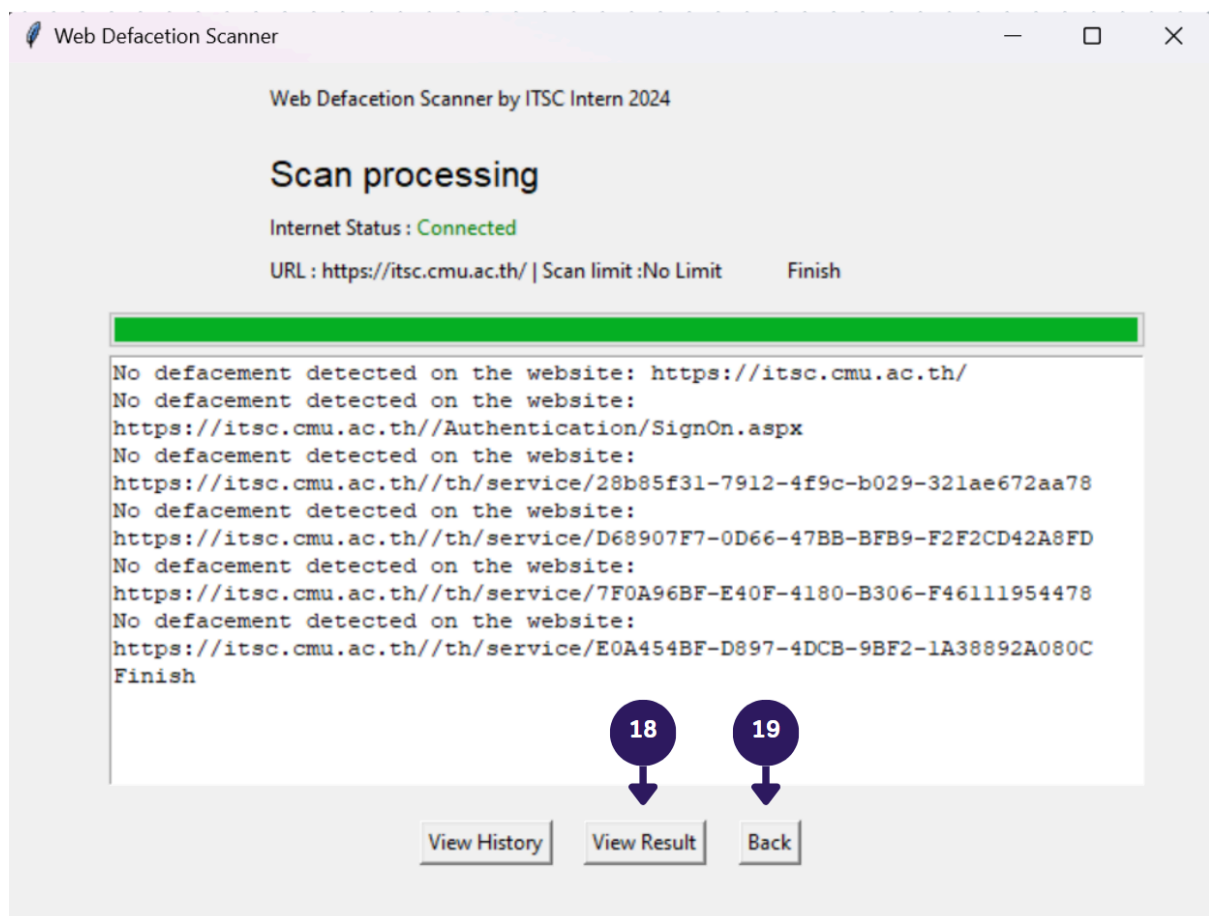
หน้าต่างของโปรแกรมเมื่อเริ่มสแกน



ส่วนประกอบและฟังก์ชันของโปรแกรมในหน้าเริ่มสแกน

10. ข้อความแสดงหน้าการสแกนของโปรแกรม
11. ฟังก์ชันข้อความแสดงสถานะรายละเอียดของการสแกนได้แก่ URL และ Scan limit
12. ฟังก์ชันข้อความแสดงเวลาที่คาดว่าจะประมวลผลเสร็จ
13. ฟังก์ชัน Progress bar แสดงความคืบหน้าของการสแกนยึดตาม scan limit ในกรณีที่ใส่เป็น No Scan limit ตัว progress bar จะไม่แสดงความคืบหน้าระหว่างการสแกน แต่ยังคงแสดงได้อย่างถูกต้องเมื่อโปรแกรมทำการสแกนเสร็จสิ้น
14. ฟังก์ชันหน้าต่างแสดง output ของการสแกนแบบ real time
15. ปุ่มฟังก์ชันสวิตช์ ชื่อ pause เมื่อกดแล้วจะหยุดการสแกนชั่วคราว
16. ปุ่มฟังก์ชันสวิตช์ ชื่อ resume เมื่อกดแล้วจะเริ่มการสแกนต่อ
17. ปุ่มฟังก์ชัน ชื่อ stop เมื่อกดแล้วจะหยุดการสแกนเพื่อจบการทำงานโดยมีการสอบถามความแน่ใจ และตัวเลือกกว่าผู้ใช้จะเก็บ result ที่สแกนมาหรือไม่

หน้าต่างของโปรแกรมเมื่อสิ้นสุดการสแกน



ส่วนประกอบและฟังก์ชันของโปรแกรมในหน้าสิ้นสุดการสแกน

18. ปุ่มฟังก์ชัน ชื่อ View Result เมื่อกดแล้วจะเปิดไฟล์ผลการสแกนที่เป็น .txt ขึ้นมา
19. ปุ่มฟังก์ชัน ชื่อ Back เมื่อกดแล้วจะนำพาไปยังหน้าแรกของโปรแกรมพร้อมที่จะเริ่มทำงานใหม่

Technical Details

ส่วนนี้เป็นการอธิบายเกี่ยวกับรายละเอียดทางเทคนิคของโปรแกรมสแกนการโจมตีเว็บไซต์ (web defacement scanning) โดยโปรแกรมถูกพัฒนาด้วย Python และไลบรารีต่างๆ เพื่อสร้างอินเทอร์เฟซกราฟิก (GUI) ดำเนินการสแกนเว็บ จัดการไฟล์ และจัดการการจับคู่ระบบเครือข่าย ด้านล่างนี้เป็นองค์ประกอบทางเทคนิคหลัก และคำอธิบายของแต่ละส่วน:

Libraries and Imports

1. Standard Libraries:

os: ใช้สำหรับการดำเนินการเกี่ยวกับระบบปฏิบัติการ เช่น การจัดการเส้นทางและการเปิดไดเรกทอรี

sys: ใช้เพื่อเข้าถึงพารามิเตอร์และฟังก์ชันที่เฉพาะเจาะจงของระบบ

datetime: ใช้สำหรับการจัดการกับวันที่และเวลา

socket: สำหรับการดำเนินการที่เกี่ยวข้องกับเครือข่าย เช่น การตรวจสอบชื่อโดเมน

threading: สำหรับการจัดการการดำเนินการที่เกิดขึ้นพร้อมกันโดยใช้เธรด

2. Third-party Libraries:

tkinter: สำหรับสร้างอินเทอร์เฟซกราฟิก (GUI)

BeautifulSoup จาก bs4: สำหรับการวิเคราะห์เอกสาร HTML และ XML

requests: สำหรับการทำคำร้องขอ HTTP เพื่อดึงเนื้อหาเว็บ

urllib.request และ urllib.parse: สำหรับการจัดการด้าน URL

csv: สำหรับอ่านและเขียนไฟล์ CSV

Global Variables

'running', 'stop', 'thread', 'keepHistory', 'Domain_URL': ควบคุมสถานะของโปรแกรมและการดำเนินการเธรด

Functions

Core Functionalities

run: เริ่มกระบวนการสแกนโดยการตรวจสอบ URL, อัปเดตอินเทอร์เฟซ, และเริ่มเธรดใหม่เพื่อเรียกใช้ฟังก์ชัน find_defacement

back: รีเซ็ตอินเทอร์เฟซเพื่อกลับสู่หน้าหลักและหยุดกระบวนการสแกนที่กำลังดำเนินการ

is_valid_domain: ตรวจสอบชื่อโดเมนว่ามีจริงไหม

open_file: เปิดไฟล์ keyword.txt หรือสร้างไฟล์ใหม่หากไม่มี และแสดงเนื้อหาในอินเทอร์เฟซ

save_file: บันทึกคีย์เวิร์ดที่พบบนอินเทอร์เฟซไปยังไฟล์ keyword.txt

check_internet_connection: ตรวจสอบการเชื่อมต่ออินเทอร์เน็ตโดยพยายามเชื่อมต่อกับ Google

update_status: อัปเดตสถานะการเชื่อมต่ออินเทอร์เน็ตที่แสดงในอินเตอร์เฟซเป็นระยะเวลา
get_program_directory: ค้นหาไดเรกทอรีที่โปรแกรมตั้งอยู่
open_History_folder: เปิดโฟลเดอร์ประวัติเพื่อดูผลลัพธ์การสแกนที่ผ่านมา
open_Result_file: เปิดไฟล์ผลลัพธ์ที่สร้างโดยการสแกนล่าสุด
validate_entry: ตรวจสอบข้อมูลที่ป้อนในช่องจำกัดการสแกนเพื่อให้แน่ใจว่าเป็นตัวเลขหรือไม่จำกัด

GUI Components

Labels, Entry fields, Buttons, Frames, and Text widgets: ใช้ในการสร้างอินเตอร์เฟซผู้ใช้ เช่น ฟิลด์ป้อนข้อมูล การแสดงสถานะ และปุ่มดำเนินการ

Scanning and Defacement Detection

fetch_website_content: ดึงเนื้อหา HTML ของ URL ที่กำหนด
find_defacement: ฟังก์ชันหลักที่สแกนหน้าเว็บเพื่อค้นหาการบุกรุกโดยการค้นหาคำหลักที่ระบุ มันจัดการดึงเนื้อหาเว็บ แยกวิเคราะห์ด้วย BeautifulSoup และอัปเดตอินเตอร์เฟซด้วยผลลัพธ์

File Handling

write_result: เขียนผลลัพธ์การสแกนลงในไฟล์ในโฟลเดอร์ประวัติ โดยสร้างโฟลเดอร์หากไม่มีการเก็บข้อมูล

Thread Management

pausing, stopping, resuming: ฟังก์ชันในการควบคุมสถานะของกระบวนการสแกน (หยุดชั่วคราว หยุด และดำเนินการต่อ) โดยใช้ตัวแปรโลกและการจัดการเธรด

Progress and Status Updates

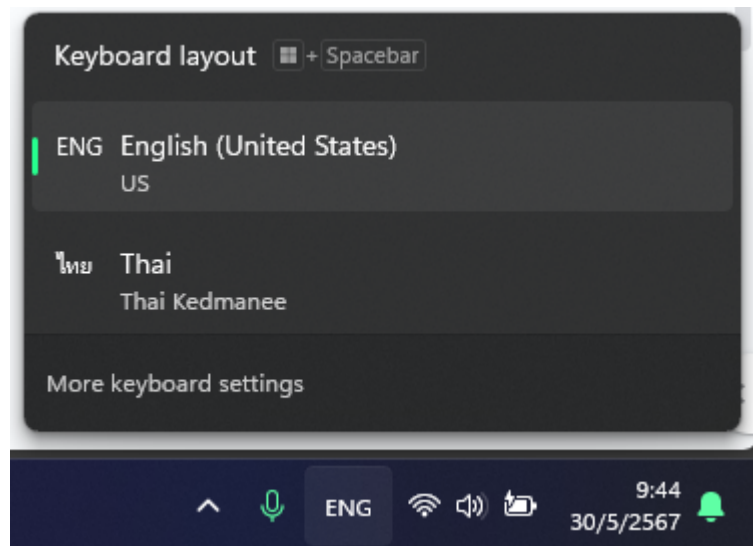
update_status: อัปเดตสถานะการเชื่อมต่ออินเทอร์เน็ตเป็นระยะๆ
progressbar: อัปเดตแถบความคืบหน้าในอินเตอร์เฟซเพื่อแสดงความคืบหน้าของการสแกน

Execution Flow

Initialization: โปรแกรมเริ่มต้นโดยการกำหนดค่าคอมโพเนนต์อินเตอร์เฟซและการตั้งค่าเหตุการณ์
User Input: ผู้ใช้ป้อน URL และขีดจำกัดการสแกนและสามารถเริ่มการสแกน บันทึกคำหลัก หรือดูประวัติ
Scanning Process: ฟังก์ชัน run ตรวจสอบ URL เริ่มเธรดใหม่ และเรียกใช้ find_defacement เพื่อสแกนหน้าเว็บสำหรับคำหลักที่ระบุ
Results Handling: ผลลัพธ์ถูกแสดงในอินเตอร์เฟซและบันทึกไว้ในไฟล์ตามต้องการ
User Actions: ผู้ใช้สามารถหยุดชั่วคราว ดำเนินการต่อ หยุดการสแกน หรือดูผลลัพธ์ได้

Troubleshooting

1. หากไม่สามารถเขียน url ในช่องที่ 3 ได้ ให้เปลี่ยนภาษาคีย์บอร์ดเป็นภาษาอังกฤษ หากเปลี่ยนแล้วยังไม่สามารถเขียนได้ ให้ทำการปิดโปรแกรมเปิดใหม่



2. หากโปรแกรมค้างระหว่างการ scan ซึ่งไม่สามารถ pause หรือ stop ได้ ให้รอสักพักประมาณ 5-10 นาที ซึ่งโปรแกรมอยู่ระหว่างการทำงาน หากไม่สามารถรอได้ ให้ทำการกดปิดโปรแกรม (โปรแกรมมีโอกาสค้างระหว่างหา keyword ภายใน HTML)

3. หากพบปัญหาใช้งานกับ PHP ไม่ได้ตอนนี้ version ปัจจุบันยังไม่รองรับ

4. หากใส่ url แล้วโปรแกรมไม่สามารถ crawling ได้ ลองให้ลองเช็ค url ซึ่งเราสามารถใส่ path ด้านหลัง url ได้ หากจำเป็นเช่น `www.example.com` เป็น `www.example.com/web2017` ซึ่งกรณีนี้ขึ้นอยู่กับโครงสร้างเว็บไซต์ของ User เองด้วย

License and Terms

1. เงื่อนไขการใช้งาน

1.1 การใช้งานที่ได้รับอนุญาต

ผู้ใช้ได้รับอนุญาตให้ใช้โปรแกรมนี้เพื่อวัตถุประสงค์ส่วนตัวและการศึกษา ซึ่งรวมถึงสิทธิ์ในการคัดลอก แก้ไข ผสมผสาน เผยแพร่ และแจกจ่ายสำเนาของซอฟต์แวร์ภายใต้เงื่อนไขของใบอนุญาตที่ได้รับ การแจกจ่าย

1.2 การใช้งานที่ต้องห้าม

ผู้ใช้ไม่ได้รับอนุญาตให้ใช้โปรแกรมนี้เพื่อกิจกรรมที่ผิดกฎหมายใด ๆ ซึ่งรวมถึงแต่ไม่จำกัดเพียง การเข้าถึงเว็บไซต์อื่นโดยไม่ได้รับอนุญาต การละเมิดข้อมูล หรือการโจมตีทางไซเบอร์ใด ๆ โปรแกรมนี้ต้องไม่ถูกใช้ในการสแกนเว็บไซต์โดยไม่ได้รับอนุญาตจากเจ้าของเว็บไซต์

2. ข้อจำกัดความรับผิดชอบ

โปรแกรมนี้ถูกจัดทำขึ้น "ตามสภาพ" โดยไม่มีการรับประกันใด ๆ ไม่ว่าจะเป็นการแสดงออกหรือโดยนัย รวมถึงแต่ไม่จำกัดเพียงการรับประกันเกี่ยวกับความสามารถในการค้า ความเหมาะสมสำหรับวัตถุประสงค์เฉพาะ และการไม่ละเมิดลิขสิทธิ์ ไม่ว่าในกรณีใด ๆ ผู้เขียนหรือผู้ถือลิขสิทธิ์จะไม่รับผิดชอบต่อ การเรียกร้อง ความเสียหาย หรือความรับผิดอื่นใด ไม่ว่าจะเป็นในรูปแบบของการกระทำทางสัญญา การละเมิด หรือในทางอื่น ๆ ที่เกิดขึ้นจากหรือเกี่ยวข้องกับซอฟต์แวร์หรือการใช้หรือการจัดการอื่นใดในซอฟต์แวร์นี้

3. ความรับผิดชอบของผู้ใช้

ผู้ใช้โปรแกรมนี้ต้องรับผิดชอบในการตรวจสอบให้แน่ใจว่าการใช้โปรแกรมนี้เป็นไปตามกฎหมาย และระเบียบข้อบังคับที่บังคับใช้ทั้งหมด ผู้เขียนโปรแกรมไม่รับผิดชอบต่อการใช้งานซอฟต์แวร์ที่ไม่ถูกต้องใด ๆ

4. การเปลี่ยนแปลงเงื่อนไข

ผู้เขียนขอสงวนสิทธิ์ในการเปลี่ยนแปลงเงื่อนไขเหล่านี้ได้ทุกเมื่อ ผู้ใช้ควรตรวจสอบเงื่อนไขเป็นระยะ ๆ เพื่อการเปลี่ยนแปลงใด ๆ การใช้โปรแกรมต่อไปหลังจากการเปลี่ยนแปลงใด ๆ แสดงว่าคุณยอมรับเงื่อนไขใหม่

Guidelines for further development

1. โปรแกรมเวอร์ชันล่าสุดไม่สามารถใช้งานได้กับ HTML ที่มี a tag href นำหน้าด้วย “.” ยกตัวอย่างเช่น `` ซึ่งได้ทำการ comment #!unfinish ในส่วนนั้นไว้แล้วใน source code สามารถไปแก้ไขทำได้ โดยส่วนใหญ่เว็บไซต์ที่มี tag แบบนี้จะเป็น PHP ซึ่งจะมีผลต่อ path ด้านหลัง ยกตัวอย่างเช่นเว็บไซต์ `www.example.com/pathA/pathB/pathC` เมื่อโปรแกรมเจอ `` ณ page นั้นผลลัพธ์ของ url ควรจะได้ `www.example.com/pathA/pathB/content` หาก `` ควรจะเป็น www.example.com/content

2. จากหัวข้อ Troubleshooting ในข้อ 2 ปัญหาเกิดจากการติดลูปในฟังก์ชัน `find_defacement` ณ จุดใดจุดหนึ่งสามารถแก้ไขด้วยนำ code ส่วนที่นำไปใช้เพื่อหยุดการทำงานไปใส่ในลูปได้ หรือจะให้ดีเลยก็ทำฟังก์ชันใหม่ให้เริ่ม thread โดยสามารถบันทึกข้อมูลจากฟังก์ชัน `find_defacement` เมื่อ user หยุดให้บังคับ thread หยุดทำงานได้เลย

Appendices

Appendix A: References

1. GeeksforGeeks - GeeksforGeeks เป็นเว็บไซต์ที่ให้แหล่งข้อมูลและบทความเกี่ยวกับการเขียนโปรแกรมและเทคโนโลยีต่าง ๆ ซึ่งมีบทความที่ครอบคลุมถึงการใช้งาน Tkinter ใน Python (Website: <https://www.geeksforgeeks.org/python-gui-tkinter/>)
2. Tkinter Documentation - เอกสารอ้างอิงสำหรับ Tkinter ซึ่งเป็นไลบรารีสำหรับสร้าง GUI (Graphical User Interface) ใน Python โดยเอกสารนี้ครอบคลุมการใช้งานทุกฟังก์ชันและวิดเจ็ตต่าง ๆ (Website: <https://docs.python.org/3/library/tkinter.html>)
3. Stack Overflow - Stack Overflow เป็นชุมชนออนไลน์สำหรับนักพัฒนาโปรแกรมที่สามารถถามและตอบคำถามเกี่ยวกับปัญหาต่าง ๆ ในการเขียนโปรแกรม รวมถึงการใช้งาน Tkinter และการทำเว็บครอว์ลิ่ง (Website: <https://stackoverflow.com/>)
4. Chiang Mai University - มหาวิทยาลัยเชียงใหม่เป็นสถาบันการศึกษาที่มีการจัดทำและเผยแพร่ข้อมูลวิชาการต่าง ๆ รวมถึงบทความและวิจัยที่เกี่ยวข้องกับการพัฒนาโปรแกรมและเทคโนโลยี (Website: <https://www.cmu.ac.th/>)
5. ITSC - สำนักบริการเทคโนโลยีสารสนเทศ ผู้ให้บริการโซลูชันดิจิทัลแบบมืออาชีพสำหรับมหาวิทยาลัย | A Trusted-Professional Digital Solution Provider for CMU (Website: <https://itsc.cmu.ac.th/>)

Appendix B: Sample Code

B.1 Tkinter Sample Code

```
import tkinter as tk

def on_click():
    print("Button clicked!")

root = tk.Tk()
root.title("Sample Tkinter App")

button = tk.Button(root, text="Click Me", command=on_click)
```

```
button.pack()
```

```
root.mainloop()
```

B.2 Web Crawling Sample Code

```
import requests
```

```
from bs4 import BeautifulSoup
```

```
def crawl(url):
```

```
    response = requests.get(url)
```

```
    soup = BeautifulSoup(response.text, 'html.parser')
```

```
    for link in soup.find_all('a'):
```

```
        print(link.get('href'))
```

```
crawl('http://example.com')
```