**Homoglyph Link checker**

**Tool Name:** Homoglyph Link Checker

**By:** Gautam Poojari (Intern ID: 251)

**Organisation:** DigiSuraksha — Parhari Foundation

**What’s This Tool All About?**

The **Homoglyph Domain Detector** is a lightweight web tool that identifies deceptive URLs by catching Unicode-based domain spoofing attacks. These homoglyph attacks rely on lookalike characters from other scripts (like Cyrillic or Greek) to trick users. This tool scans any pasted or uploaded text for such suspicious links using a Python-Flask backend.

**Why Does It Matter?**

Phishing emails, scam messages, and fake login pages often use **visually identical but malicious URLs**. A single Cyrillic or Greek character swap can create an evil twin of your trusted site.

This tool:

* Finds and flags homoglyph (visually confusing) domains
* Helps in threat triage, email audits, and awareness training
* Supports English, Greek, Russian — all the sneaky Unicode boys

# Core Features

* Detects homoglyphs (script mixing like Cyrillic “о” vs Latin “o”)
* Supports both **text input** and **file upload**
* Shows both **all links** and **suspicious links**
* Simple, clean Flask web interface (no command line needed)
* Uses homoglyphs Python package internally
* Easy to extend into full anti-phishing automation

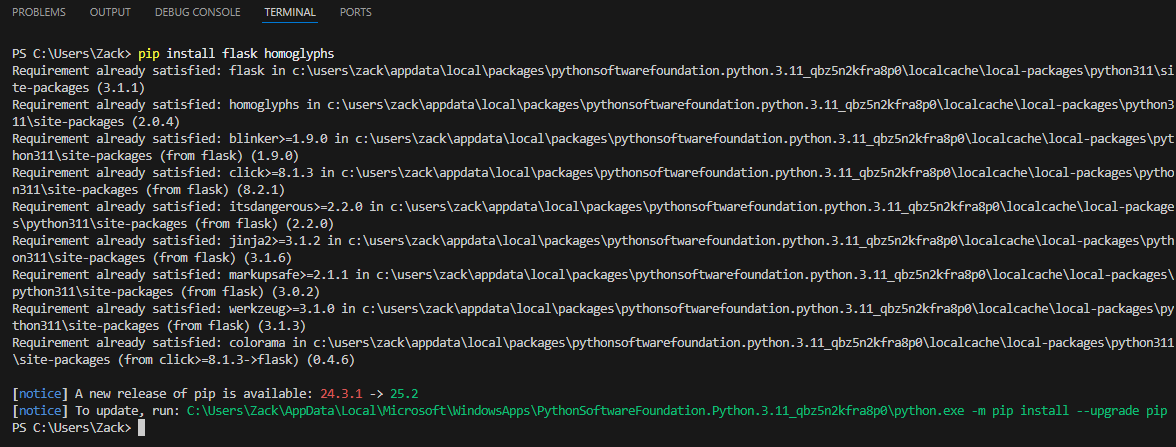
# How It Works

When you paste or upload text:

1. The app extracts all URLs using regex
2. It checks each link using the **homoglyphs** module
3. Suspicious ones (i.e., not ASCII-normalized) are flagged

# Dependencies & Setup

To get started, install the required dependencies:



pip

install

flask

homoglyphs

# How To Run It?

Run the following command to launch the tool locally:

python

homoglyphs251.py

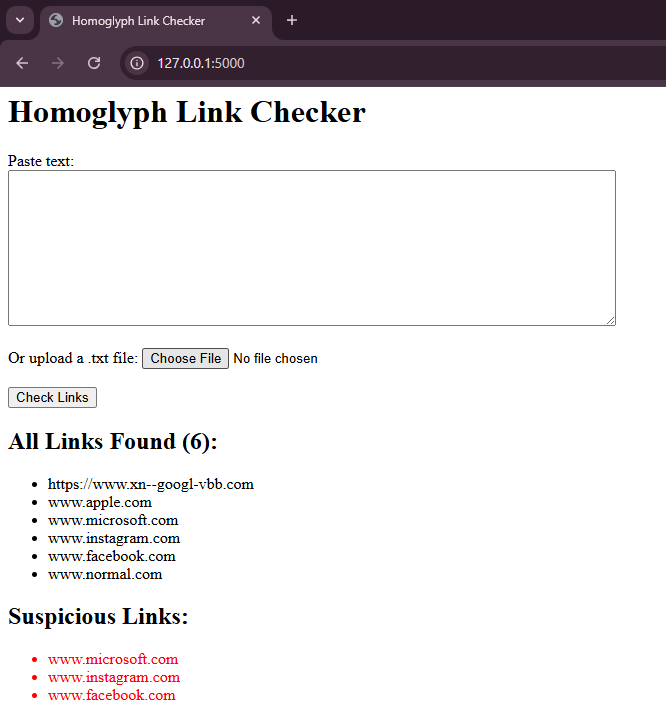
This will start the Flask development server at

http://127.0.0.1:5000

.

# 

# Web UI Output:

****

Once the app is running, open your browser and paste some suspicious-looking text or upload a

.txt

file

to scan.

# Sample Input & Output

You already got this thanks to the UI:

**Input Sample:**

<https://www.apple.com>

[www.facebоok.com](http://www.facebоok.com)

[www.normal.com](http://www.normal.com)

<https://www.xn--googl-vbb.com>

1. It checks each link using the homoglyphs module

Suspicious ones (i.e., not ASCII-normalized) are flagged

**Output:**

**All Links Found:**

* [www.apple.com](http://www.apple.com/)
* www.facebоok.com
* [www.normal.com](http://www.normal.com/)
* [www.microsoft.com](http://www.microsoft.com)
* www.instagram.com
* [https://www.xn--googl-vbb.com](https://www.xn--googl-vbb.com/)

**Suspicious Links:**

* www.facebоok.com
* [www.microsoft.com](https://www.xn--googl-vbb.com/)
* www.instagram.com

**Who Should Use This?**

* SOC Analysts
* Cybersecurity Investigators
* Threat Intel Teams
* Trainers for cyber awareness

**When to Use?**

* Triage during phishing incident reports
* Email/message audits before sending/receiving
* Red teaming or blue teaming exercises
* As part of CI/CD pipeline security checks

# Limitations

* Doesn’t support HTML or rich email parsing yet
* Might miss new Unicode tricks (needs regular script DB updates)
* No “threat score” or integration with real-time threat feeds

**What Makes It Good?**

* Simple to run, even for non-coders
* Portable across OSes
* Great for awareness + automation
* Lightweight and clean interface

# Summary

The **Homoglyph Link Checker** helps you **catch phishy links** before they catch you. It’s a fast, lightweight, and powerful addition to any digital investigation toolkit or awareness training module. Whether you're a pro analyst or just checking sketchy links in an email, this tool’s got your back.