# CWE Detail – CWE-1007

## Description

The product displays information or identifiers to a user, but the display mechanism does not make it easy for the user to distinguish between visually similar or identical glyphs (homoglyphs), which may cause the user to misinterpret a glyph and perform an unintended, insecure action.

## Extended Description

Some glyphs, pictures, or icons can be semantically distinct to a program, while appearing very similar or identical to a human user. These are referred to as homoglyphs. For example, the lowercase "l" (ell) and uppercase "I" (eye) have different character codes, but these characters can be displayed in exactly the same way to a user, depending on the font. This can also occur between different character sets. For example, the Latin capital letter "A" and the Greek capital letter "Α" (Alpha) are treated as distinct by programs, but may be displayed in exactly the same way to a user. Accent marks may also cause letters to appear very similar, such as the Latin capital letter grave mark "À" and its equivalent "Á" with the acute accent. Adversaries can exploit this visual similarity for attacks such as phishing, e.g. by providing a link to an attacker-controlled hostname that looks like a hostname that the victim trusts. In a different use of homoglyphs, an adversary may create a back door username that is visually similar to the username of a regular user, which then makes it more difficult for a system administrator to detect the malicious username while reviewing logs.

## Threat-Mapped Scoring

Score: 0.0

Priority: Unclassified

## Observed Examples (CVEs)

**•** CVE-2013-7236: web forum allows impersonation of users with homoglyphs in account names

**•** CVE-2012-0584: Improper character restriction in URLs in web browser

**•** CVE-2009-0652: Incomplete denylist does not include homoglyphs of "/" and "?" characters in URLs

**•** CVE-2017-5015: web browser does not convert hyphens to punycode, allowing IDN spoofing in URLs

**•** CVE-2005-0233: homoglyph spoofing using punycode in URLs and certificates

**•** CVE-2005-0234: homoglyph spoofing using punycode in URLs and certificates

**•** CVE-2005-0235: homoglyph spoofing using punycode in URLs and certificates

## Related Attack Patterns (CAPEC)

* CAPEC-632

## Modes of Introduction

**•** Architecture and Design: This weakness may occur when characters from various character sets are allowed to be interchanged within a URL, username, email address, etc. without any notification to the user or underlying system being used.

**•** Implementation: N/A

## Common Consequences

**•** Impact: Other — Notes: An attacker may ultimately redirect a user to a malicious website, by deceiving the user into believing the URL they are accessing is a trusted domain. However, the attack can also be used to forge log entries by using homoglyphs in usernames. Homoglyph manipulations are often the first step towards executing advanced attacks such as stealing a user's credentials, Cross-Site Scripting (XSS), or log forgery. If an attacker redirects a user to a malicious site, the attacker can mimic a trusted domain to steal account credentials and perform actions on behalf of the user, without the user's knowledge. Similarly, an attacker could create a username for a website that contains homoglyph characters, making it difficult for an admin to review logs and determine which users performed which actions.

## Potential Mitigations

**•** Implementation: Use a browser that displays Punycode for IDNs in the URL and status bars, or which color code various scripts in URLs. Due to the prominence of homoglyph attacks, several browsers now help safeguard against this attack via the use of Punycode. For example, Mozilla Firefox and Google Chrome will display IDNs as Punycode if top-level domains do not restrict which characters can be used in domain names or if labels mix scripts for different languages. (Effectiveness: N/A)

**•** Implementation: Use an email client that has strict filters and prevents messages that mix character sets to end up in a user's inbox. Certain email clients such as Google's GMail prevent the use of non-Latin characters in email addresses or in links contained within emails. This helps prevent homoglyph attacks by flagging these emails and redirecting them to a user's spam folder. (Effectiveness: N/A)

## Applicable Platforms

**•** None (Class: Not Language-Specific, Prevalence: Undetermined)

## Demonstrative Examples

**•** However, the URL above is comprised of Cyrillic characters that look identical to the expected ASCII characters. This results in most users not being able to distinguish between the two and assuming that the above URL is trusted and safe. The "e" is actually the "CYRILLIC SMALL LETTER IE" which is represented in HTML as the character &#x435, while the "a" is actually the "CYRILLIC SMALL LETTER A" which is represented in HTML as the character &#x430. The "p", "c", and "o" are also Cyrillic characters in this example. Viewing the source reveals a URL of "http://www.&#x435;x&#x430;m&#x440;l&#x435;.&#x441;&#x43e;m". An adversary can utilize this approach to perform an attack such as a phishing attack in order to drive traffic to a malicious website.

**•** Assume an adversary visits a legitimate, trusted domain and creates an account named "admin", except the 'a' and 'i' characters are Cyrillic characters instead of the expected ASCII. Any actions the adversary performs will be saved to the log file and look like they came from a legitimate administrator account.