CSC

Threatening Domains Targeting the Top 10 Most Valuable Brands

Key findings and highlights

Between August 2021 and August 2022, CSC identified 8,480 unique third-party domain names comprising a very close match to the brand names of the global top ten most valuable companies, with more than

99%

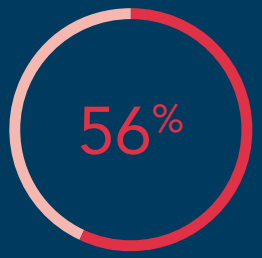
of those (where information is available) registered by third parties, leaving those brands vulnerable to targeted attacks by bad actors.

Domains definitively determined to be official were excluded from the remainder of the analysis.

Where registration information is available,

two thirds

of the domains used domain privacy services—indicating an intention by the owner to mask their identity—or have redacted information.



56% of the domains still registered at the time of analysis resolved to a live webpage. Among the live sites, we observed a range of high-concern content types, including fraud issues like potential phishing sites, and other brand infringements.

35% of the registered domains were configured with active mail exchange (MX) records, indicating their ability to send and receive emails, making them capable of launching phishing attacks.

Many of the domain names are chosen to appear deceptively similar to official brand domain names or feature common typo variants to catch misdirected web traffic. Domains using non-Latin characters raise the greatest potential for confusion (and thereby for fraudulent use) as they can appear almost identical to their Latin equivalents, for example:

amazon.com **apple.com** **facebook.com** **google.com** **microsoft.com**

Methodology of the analysis

In this analysis, we dive into potentially the most egregious and threatening set of domains targeting the top 10 most valuable company brands in 2022[1]. These domains are those where the second-level domain name (SLD)—the part of the domain name to the left of the dot—consists solely of an exact or very close match to the targeted brand name.

We used **CSC's 3D Domain Security & Enforcement** technology—powered by the DomainSec℠ platform, which uses proprietary Machine Learning Deep Search (MLDS) technology and combines machine learning, artificial intelligence, and clustering technology to identify leading indicators of compromise—to conduct the analysis. We considered new registrations (N), re-registrations (R) or drops (D)—collectively referred to as domain registration activity events—over the period August 2021 to August 2022.

The study focuses on domains containing any of the following brand variants as the SLD:

Exact matches—where the SLD is identical to the brand term under consideration, but with a different extension (TLD) to the official brand website (i.e., a 'cousin' domain).

Homoglyph matches—where one or more characters in the brand term are replaced by a visually similar, non-Latin character.

Fuzzy matches—featuring typos (misspellings) affecting a single character, covering missing characters, additional characters, transposed characters, and other character replacements.

Each of these domain names therefore appears extremely similar to that of the brand's official site and raises significant potential for targeted attacks. These variants may have been deliberately selected by those registering the domains to be confusing or to circumvent detection efforts by brand owners, who may be monitoring only for exact matches to the brand string. Such activity presents significant threat vectors to the targeted brands, as those domains can be used for a variety of purposes, including active fraudulent use (e.g., creating phishing sites) or potential brand confusion and traffic misdirection, by taking advantage of the status of and consumer trust in the brand being infringed, or attracting traffic from mis-typed browser requests or search engine queries.

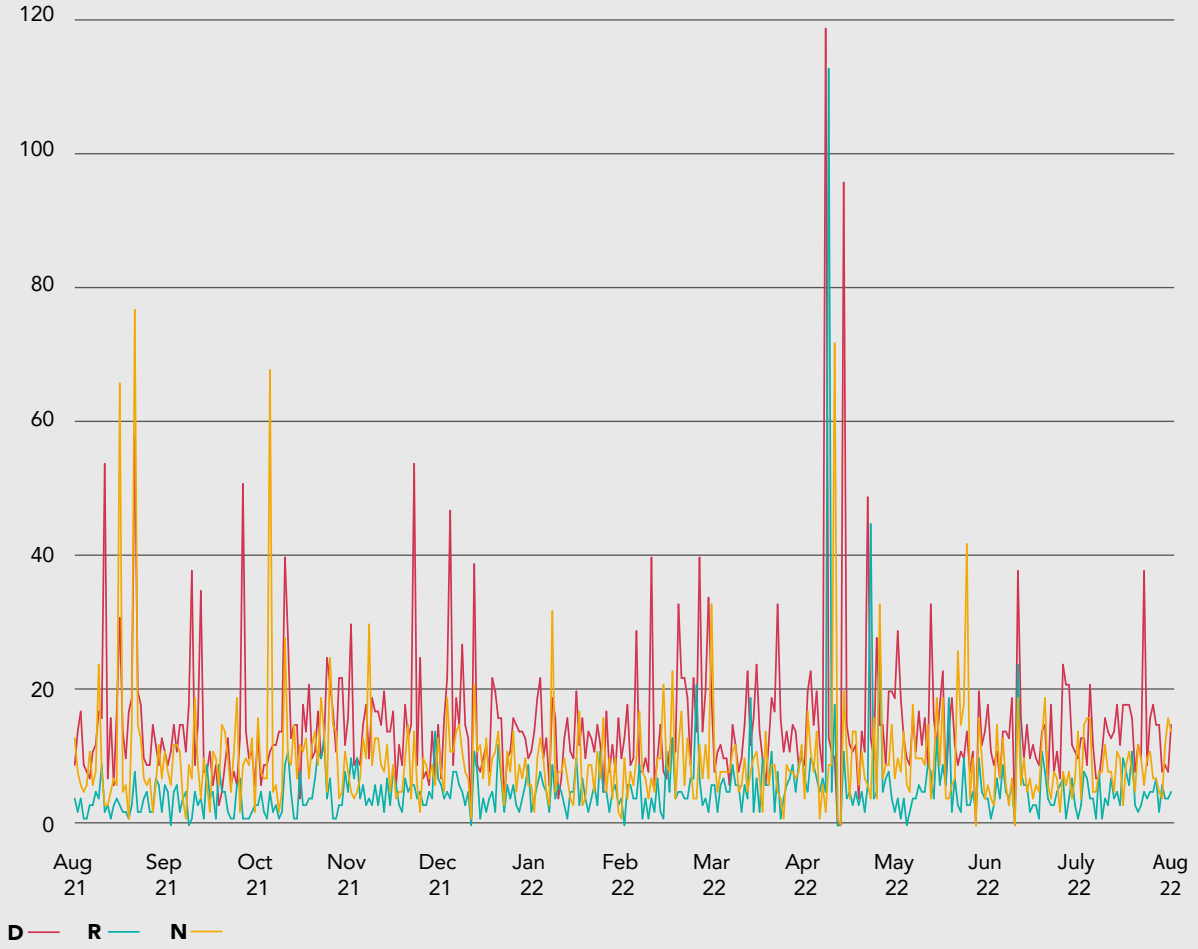
1. kantar.com/inspiration/brands/what-are-the-most-valuable-global-brands-in-2022; the brand terms used in our analysis are apple; google; amazon; microsoft; tencent; mcdonalds; visa; facebook; alibaba; and vuitton.

Findings

1. Domain activity

Over the analysis period, CSC identified more than 11,000 domain registration activity events across the ten brands under consideration, focusing only on the very close matches as described in the methodology. As with a previous CSC study considering deceptive domains with names beginning 'www' or 'http'[2], we saw continuous activity across the period (Figure 1).

Figure 1: Daily numbers of new registrations (N), re-registrations (R) and dropped (D) domains with names with a close match to any of the ten brand names under consideration.



2. cscdbs.com/blog/registration-patterns-of-deceptive-domains/

Findings

In some cases, individual spikes in activity can be tied to specific events or coordinated registration campaigns.

For example, a batch of 56 domains featuring misspellings of "mcdonalds.co.uk" was registered on August 27, 2021, followed by 40 domains on September 1, 2021 that comprised "microsoft" typos across the .BIZ and .ONE extensions, and then 46 "google.fr" and five "amazon.fr" typo variants on October 16, 2021. Additionally, a set of domains featuring "amazon" (or typos) and "ten-cent" as the SLD was registered on April 22, 2022 across a range of different new gTLD extensions. Previous research by CSC[3] established that a peak in registrations of domain names featuring a key pharmaceutical brand with energy-related keywords took place immediately coinciding with the launch of the Energize program[4].

3. 'Domain registration patterns analysis' (unpublished)
4. se.com/ww/en/about-us/newsroom/news/press-releases/10-global-pharmaceutical-companies-launch-first-of-its-kind-supplier-program-to-advance-climate-action-6182848cf01af478b619ddd4

In total, we identified 8,552 unique domain names in the dataset. For the active domains where WHOIS information was available, only 72 domains (<1%) were explicitly registered by the official brand owners—presumably as defensive registrations or acquired domains to prevent third-party use. The remaining domains were registered by third parties. The analysis presented here focuses on these 8,480 third-party domains.

8,480

unique third-party domain names in the dataset

Several types of brand variants (listed below) were present in the set of third-party SLDs, with their relative proportions within the dataset shown in Figure 2.

- Exact match (i.e., 'cousin' domains)
- Missing character
- Extra character
- Transposed characters (i.e., a swapped pair of adjacent characters)
- Replaced characters – either non-Latin homoglyphs or other character replacements, with the number of replaced characters shown in Figure 2

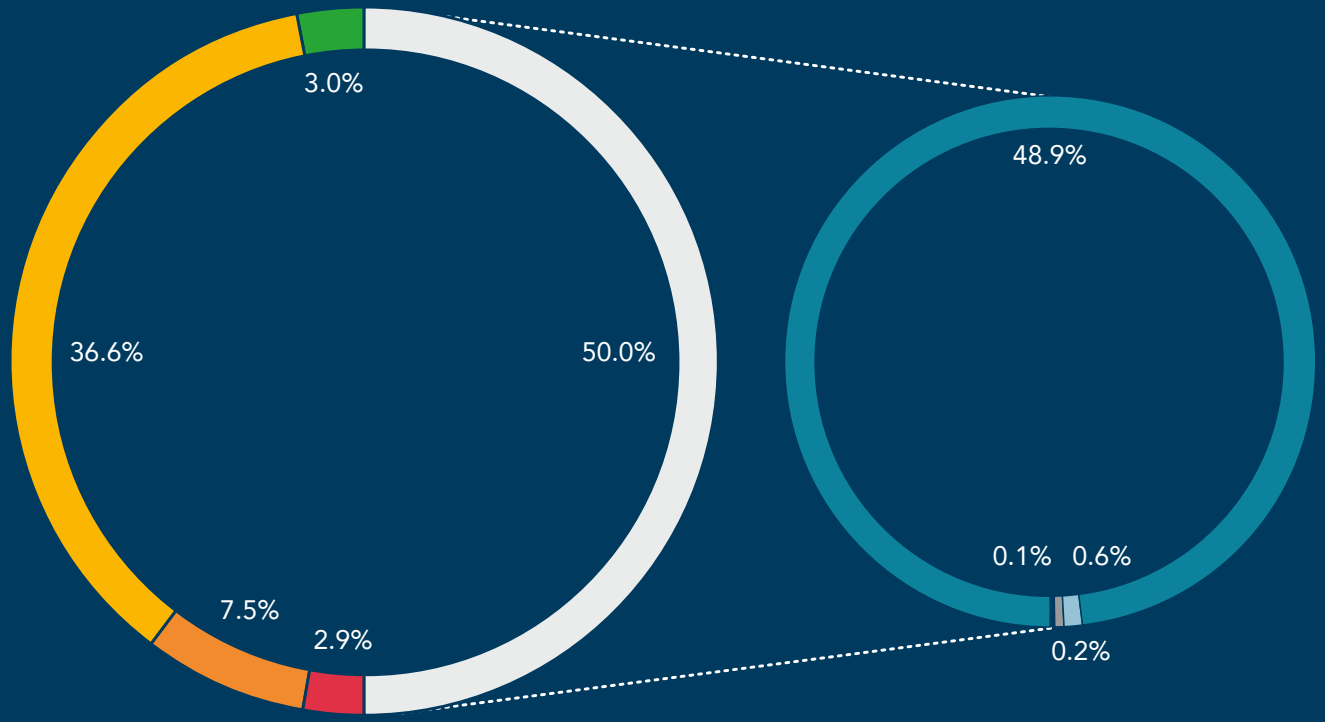
  N.B. Those domains with multiple character substitutions are generally homoglyph (internationalized-character) domain names, since fuzzy-match searches (covering other character substitutions) explicitly focus on close matches with a single character differing from the search term.

Findings

2. Brand variant types

The high potential for confusion between these domain names and those of the official brands' websites poses a significant threat to their security postures, as well as the risk of infringing use by bad actors for phishing activity, or other traffic misdirection.

Figure 2: Frequency of brand variant types used in the dataset of unique domains.



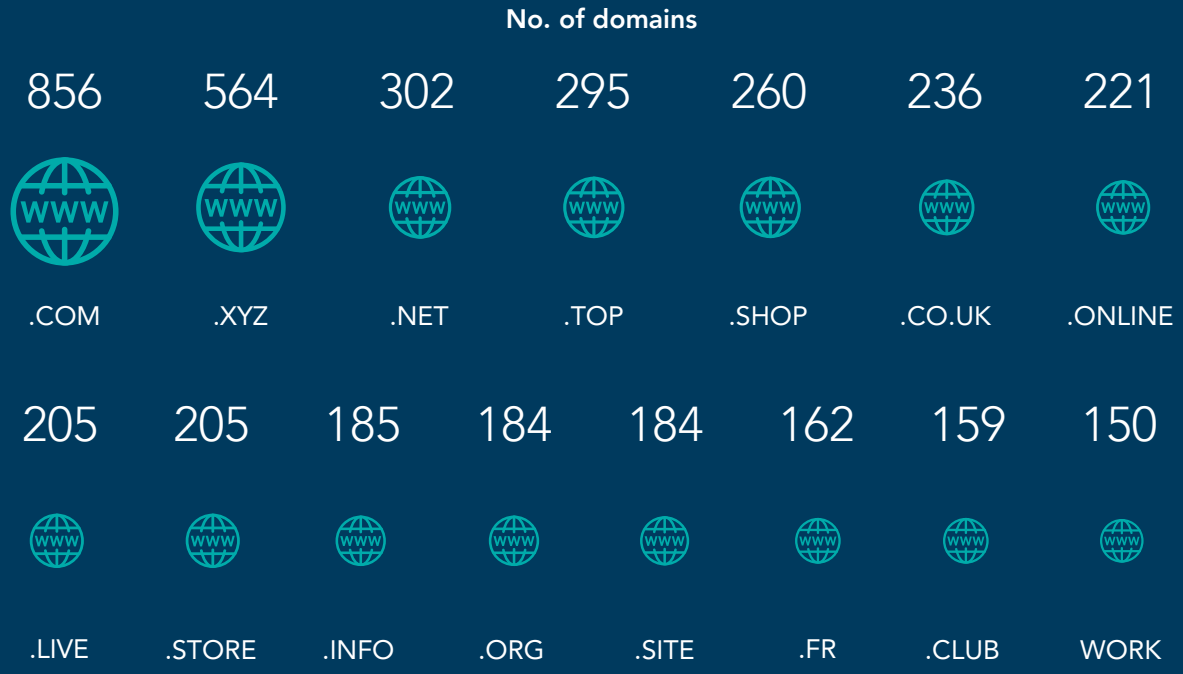
Within the dataset, 3% of the domains featured an exact match to the targeted brand name (i.e., cousin domains), 44% featured a missing or extra character, 3% featured transposed characters, and 50% featured one or more replaced characters. In total, just over 3% of the dataset featured homoglyph domains, i.e., those incorporating non-ASCII characters (characters other than the Latin alphabet and other standard characters).

Findings

3. Observations on frequently occurring features of unique domains

A. Top TLDs

The following are the most popular TLDs represented within the dataset of 8,480 third-party domains.



As has been seen with previous studies[5,6,7], this list of extensions is dominated by popular TLDs (e.g., .COM), and new gTLDs (e.g., .XYZ and .TOP), which are proven to be popular with infringers.

5. cscdbs.com/blog/branded-domains-are-the-focal-point-of-many-phishing-attacks/
6. circleid.com/posts/20210908-credential-hinting-domain-names-a-phishing-lure
7. unit42.paloaltonetworks.com/top-level-domains-cybercrime/

B. Brand variants consisting of transposed characters

Within the dataset of unique domains, the following are the most observed brand variants where a pair of adjacent characters have been swapped. In total, 253 of the domains were found to incorporate transposed characters.

Brand variant	No. domains
googel	29
appel	25
goolge	23
micorsoft	17
faecbook	15
amzaon	14
mircosoft	13
amazno	13
gogole	13
amaozn	11

Findings

A B C D E F G H I J K L M N

C. Most popular character replacements

The dataset of unique domains encompassed more than 3,000 individual replaced characters[8].

The top 10 most common replacements with Latin alphabet characters were:

Character replacement	No. instances
l → i	93
n → m	59
o → i	52
a → o	47
c → k	47
o → a	45
a → e	42
o → g	40
g → d	38
g → b	37

The top 10 most common replacements with other characters (non-Latin characters and numbers) were:

Character replacement	No. instances
o → 0	78
l → 1	24
o → õ	20
o → ó	17
o → ö	15
l → ł	15
e → 3	13
e → é	12
z → -	10
o → o	10

In many of these cases, the characters have been replaced with visually similar alternatives to create a convincing lookalike domain name. There are instances, however—notably in the Latin alphabet character replacements—where characters have been replaced with characters that are adjacent to them on a standard QWERTY keyboard (e.g., n → m, o → i, g → b, etc.) presumably to try to catch misdirected traffic from browser requests containing common typing errors. It's also worth noting that some of the common character replacements observed in the dataset may relate to true third-party brand use or unrelated terms (e.g., all the observed g → m replacements in the dataset appear as the term 'moogle', which could refer to a creature in the Final Fantasy gaming series).

8. In this analysis, we exclude examples where the replaced-character version forms an alternative word in its own right (e.g. 'apply' or 'ample' for 'apple'), and all replacements for the Visa brand (since the small number of characters means that many of the variations are significantly different from the brand name and may pertain to unrelated third-party use), as the replaced versions in these cases may not be intended to be deceptive variants of the brand name in question.

Findings

D. Top registrant and registrar characteristics

Among the dataset, we found WHOIS (domain registration) information for **4,513 domains**, of which two thirds (3,007) used domain privacy services or had redacted registration information. Use of anonymization services demonstrates a domain owner's attempt to mask their identity and could indicate nefarious intentions[9].

The following graphics show the top organizations and countries given in the sets of registration contact details for the domains.

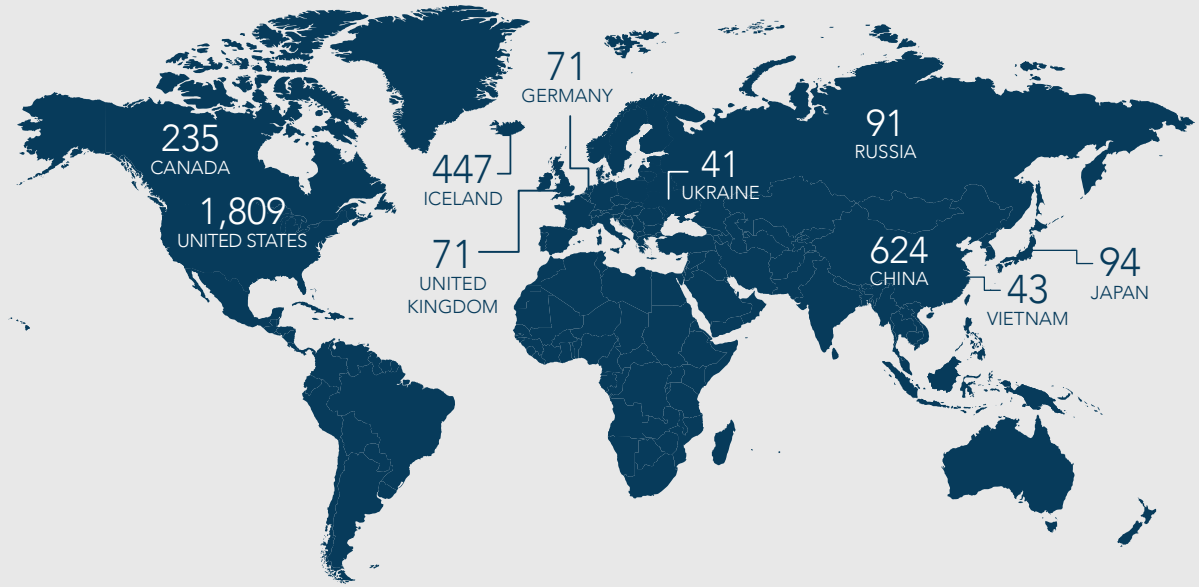
Most common registrant organizations:

Organization	No. domains
Domains By Proxy, LLC	752
Privacy service provided by Withheld for Privacy ehf	401
See PrivacyGuardian.org	245
Contact Privacy Inc. Customer 7151571251	153
Private by Design, LLC	134

two thirds

of domains with WHOIS data used domain privacy services

Most common registrant countries:



9. cscdbs.com/en/resources-news/supply-chain-report/

Findings

The following table shows the top registrars through which the domains in the dataset were registered.

Most common registrars:

Registrar	No. domains
GoDaddy.com, LLC	839
Namecheap, Inc.	485
NameSilo, LLC	286
Dynadot LLC	263
Porkbun LLC	176
Alibaba Cloud Computing Ltd. d/b/a HiChina	148
DNSPod, Inc.	142
Name.com, Inc.	111
PDR Ltd. d/b/a PublicDomainRegistry.com	106
Google LLC	97

The registrar landscape within the dataset is dominated by consumer-grade providers, a trend which has also been previously seen in other CSC studies of domains registered for potentially infringing use[10].

10. cscdbs.com/en/resources-news/impact-of-covid-on-internet-security/

Findings

4. Cybersecurity, fraud protection, and brand protection observations

Of the 8,480 unique third-party domains in the dataset, 4,552 (54%) were still registered at the time of analysis (i.e., those domains where the most recent activity event was a registration or re-registration). Just less than a third of the domains—equivalent to 56% of the registered domains—were found to produce a live website response[11]. Furthermore, 1,590 domains (19% of the dataset, or 35% of the registered domains) were configured with active MX records, indicating that they can send or receive emails (e.g. for use in a phishing attack). This is possible even when there is no live site content. Additionally, as noted in previous studies[12], dormant domains also have potential to be used fraudulently in the future, so monitoring for changes to configuration status and site content is advisable.

Live sites in the dataset featured a range of content, including lookalike sites; websites using potentially unauthorized official branding; third-party sites using similar branding; sites featuring content relating to third parties operating in a similar business area as the infringed brand; and gambling-related or adult material. Additionally, many displayed pay-per-click links—as a means of monetizing the web traffic—or holding pages offering to sell the domain names, i.e., potential cybersquatting. Many displayed browser warnings indicating threatening content is or has been present.

11. Those returning an HTTP status code of 200
12. unit42.paloaltonetworks.com/strategically-aged-domain-detection/

Findings

Figure 3 shows examples of some of the most significant infringements identified within the dataset, with the associated SLD (domain-name string) shown in square brackets. This includes websites trying to pass off as the brand in question (as part of a fraudulent brand-impersonation attack, for example), unauthorized use of official intellectual property (IP)—particularly concerning if the content provides an undesirable brand association—or generating revenue for a third party operating in a similar business area through misdirecting web traffic that's arguably intended for the official brand.

Figure 3: Examples of brand infringements identified within the dataset

a) Lookalike site [googe]

b) Potential phishing [mc-donalds]

c) Potential unauthorized branding use [aamazon]

d) Similar branding [armazon]

e) Third-party content in related business area [fgoogle]

f) Other brand infringement [fakebook]

Findings

Across the dataset, 3% of the domains included at least one character outside a standard Latin (ASCII) character set (i.e., domains that can be represented in Punycode[13] format). Many of these are visually almost identical to the official domain name for the brand in question, with no additional, missing or transposed characters, and therefore could be used as highly convincing attack vectors. Around 10% of these were configured with active MX records, indicating email capability.

Examples of highly convincing lookalike domain names (replaced characters highlighted in bold):

- amazo**ɳ**.com
- amaz**ƶ**on.com
- a**þþ**le.com
- **ɑ**pple.com
- faceboo**ĸ**.com
- **f**acebook.com
- googl**ɬ**e.com
- g**oo**gle.com
- mcdonald**ƨ**.com
- micr**o**soft.com
- m**ı**crosoft.com

13. en.wikipedia.org/wiki/Punycode

As shown earlier in Figure 2, the majority of domains incorporating character replacements featured a single replaced character. In part, this reflects the fact that the fuzzy searches explicitly focused on matches only one character different. In the dataset, we identified 33 domains where at least half of the characters were replaced by homoglyphs, e.g., gõõgłé.com and fäcëböök.com.

The full list includes a number of examples where all of the character replacements make use of non-Latin versions that look almost identical to the original characters (e.g., □□o□le.com), or where the alternative version simply appears visually to be just in a different font (e.g., F□□□□□□□.com). These could form the basis of extremely deceptive infringements.

We found 183 distinct characters (including non-Latin homoglyphs from numerous alphabets and character sets) being used as character replacements across the full dataset, as shown here.

0 1 2 3 4 5 6 7 8 9 - l a ᴀ á à ȧ â ä ǎ ă ā
ã å ą ạ ɑ ɒ æ b ʙ ḃ ƀ ḇ ь c ᴄ ć ċ ĉ ç ƈ d e
ᴇ é è ė ê ë ě ē ę ɇ ẻ ẹ ę ḛ ə ɘ ɜ ɛ f ꜰ ḟ ƒ g
ɢ ǵ ġ ĝ ǧ ğ ḡ ģ ɠ g h i ɪ í ì ǐ ī į ɨ ỉ ị ι j k ᴋ ĸ
ḱ ķ ƙ ḵ l Ĺ ł ḷ ɬ ⌊ ḻ ꟾ m ṃ n ń ñ ņ ŋ ṇ o ᴏ ο ó
ò ȯ ô ö ǒ ō õ ő ɵ ø ỏ ơ ọ ở œ p ṗ q r ʀ s ƨ
f t ť f ṭ ţ þ u ü v w ƿ ᴍ x y z ᴢ ź ž ƶ ȥ Ꝋ a
e i ï p A E L P B P G K C Z F M N L A E O

Summary

CSC's findings highlight the degree to which trusted brands can be targeted by bad actors, which presents significant threats to their security postures, revenue, and reputations. It highlights the need for domain intelligence as part of a layered security approach, including comprehensive domain monitoring as the foundation of a holistic brand protection initiative.

CSC's 3D Domain Security & Enforcement solution (powered by our DomainSec platform) can provide this overview, enabling brand owners to have visibility of domain activity encompassing detection of brand variants, including fuzzy matches like typos and homoglyph domains, and soundalike variations, across a range of domain name extensions. The system also incorporates MLDS technology to intelligently modify monitoring based on previously identified infringements.

The closeness of the match of a third-party domain name to that of the brand owner's official name is also one key input into algorithms that can help determine the threat level that may be posed in the future by that domain. These concepts are central to the idea of threat scoring, which can be used to prioritize infringements for analysis, future monitoring, and enforcement.

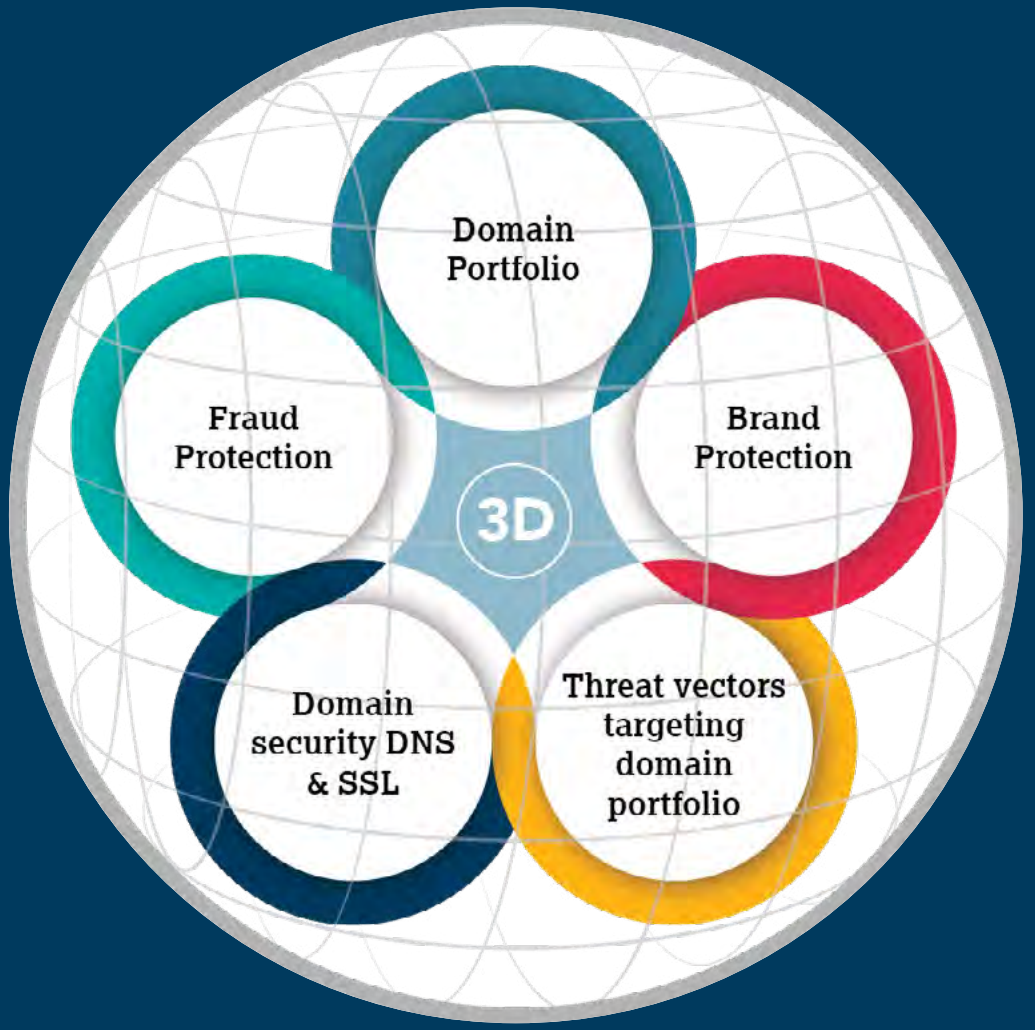
For more information on CSC's domain security, and brand and fraud protection solutions, visit cscdbs.com.

About CSC's DomainSec platform

CSC's 3D Domain Security and Enforcement solution has been created harnessing the power of CSC's DomainSec platform.

DomainSec is a SaaS cybersecurity platform invented by CSC, and is the industry's first holistic approach for securing and defending brands' domain ecosystems. It uses proprietary MLDS technology and combines machine learning, artificial intelligence, and clustering technology to identify leading indicators of compromise.

DomainSec brings CSC's domain management and domain security, along with brand protection and fraud protection solutions, into one platform—meaning we can offer exponentially better protection and help organizations refine their Zero Trust security model, going beyond just safeguarding perimeters.



CSC is the trusted cybersecurity and threat intelligence provider of choice for the Forbes Global 2000 and the 100 Best Global Brands® in enterprise domain names, domain name system (DNS), digital certificate management, as well as digital brand and fraud protection. As global companies make significant investments in their security posture, CSC can help them understand known cybersecurity oversights that exist and help them secure their online digital assets and brands. By leveraging CSC's proprietary technology, companies can solidify their security posture to protect against cyber threat vectors targeting their online assets and brand reputation, helping them avoid devastating revenue loss, and significant financial penalties because of policies like the General Data Protection Regulation (GDPR). CSC's online brand protection services—the combination of online brand monitoring and enforcement activities—take a holistic approach to digital asset protection, along with fraud protection services to combat phishing. Headquartered in Wilmington, Delaware, USA, since 1899, CSC has offices throughout the United States, Canada, Europe, and the Asia-Pacific region. CSC is a global company capable of doing business wherever our clients are—and we accomplish that by employing experts for every business we serve. Visit **cscdbs.com**.

12012022