DOMAIN MONITORING_

SOLUTIONS GUIDE & RESOURCES

digital shadows_

INTRODUCTION

Domain monitoring involves the tracking of domains that have slight variation or permutation of the target company name or brand they are impersonating. This can include a spelling error, switched characters, or additional keywords added to the domain name. It may even not have a similar domain name, but the contents of the site are visually similar and impersonating the target. Other popular snynoyms for this domain impersonations include doppelgänger, typosquatting, and URL hijacking.

These domains are subsequently leveraged for phishing attacks, BEC campaigns, and for selling counterfeit goods. Left undetected, these domains can lead to data breaches, credential theft, significant brand and reputational damage, and loss of revenue.

It's no longer economically feasible to buy all the domains similar to your own

With more than 1,200 Generic Top Level Domains recognized by ICANN, anyone with a credit card can buy any number of domains. Given this scale, defensive strategies that rely on buying similar domains no longer work. Organizations have now realized that they need to monitor for domains themselves. In this guide, you will learn how to detect domains that are registered to capitalize on your company's name and customer trust by impersonating your company's domain.

This guide outlines what potential data sources, detection methods, context, and remediation actions to consider if you want to effectively monitor domains and mitigate the risk of data loss, exposed credentials, and negative reputational impacts.

TERMINOLOGY MATTERS: DOMAIN SPOOFING

Please note that this guide does not extend to domain spoofing, SPF and DMARC. These are all incredibly important aspects of security that you can read more about in our Security Practitioners Guide to Email Spoofing.

CONTENTS

MOTIVES FOR DOMAIN IMPERSONATION

The barriers to entry for attackers registering these look-a-like domains is incredibly low, but the impact to the business can be financially devastating and irreversible.

36% OF BREACHES INVOLVE PHISHING
VERIZON DBIR 2020

Impersonating domains are most commonly used as a tool for harvesting credentials and phishing. Emails that appear to come from a similar looking domain have more credibility and therefore, a higher open rate and ability to nudge a recipient into opening an attachment or unknowingly proceeding to download malware. Similarly, when you visit a website that looks visually similar, a user is more likely to enter their credentials. Given that 36% of breaches involve phishing, it's no wonder that typosquats are so ubiquitous.

However, domain impersonation is not limited to phishing. These fake domains can be (and often are) used to spread malware to customers and employees, sell counterfeit or fraudulent products, redirect traffic to a competitor, or spread disinformation to damage public reputation. Sometimes the motivation is less nefarious, with individuals looking to make money–through selling the domain, advertising or affiliate links.

DISINFORMATION

In 2017, it was widely reported that Emmanuel Macron's presidential campaign was financed by Saudi Arabia. This story was shared by the website "lesoir[.]info", which was itself an impersonation of the Belgian news site lesoir[.]be. Despite being a spoof, the article was widely circulated across social media. Furthermore, an analysis of WHOIS registration data revealed at least 20 other spoofs of regional and global news outlets, such as timesoffisrael[.]com (a spoof of timesofisrael[.]com), bbc-arabic[.]com (a spoof of bbc[.]com/arabic), and bloomberq[.]com (a spoof of bloomberg[.]com).

COLLECT: DOMAIN SOURCES

For those wishing to detect instances of domain impersonation, the first step is to collect the right set of data. Here are the four areas you should consider.

1. REPORTED DOMAINS

Detecting these domains yourself is the ideal scenario, but domains publicly reported by others are not to be ignored. There's plenty of domains reported on Twitter and across various threat feeds – not to mention from phishing emails reported internally by employees.

2. NEWLY REGISTERED DOMAINS

The most useful data source for detecting domain impersonation is via a feed of newly registered domains. Different top level domains (TLD) such as .com, .gov, and .edu entities will provide different levels of data.

3. SSL CERTIFICATES

Certificate transparency logs are another great source of domain data. There are several free options you can turn to including CertStream, Google Transparency Report, and Crt.sh.

To learn more about certificate transparency logs, check out this great post on SANS ISC InfoSec Forums: Using Certificate Transparency as an Attack / Defense Tool

4. DNS DATA

Beyond the domains themselves, it's important to collect the DNS data associated with them. The DNS data can have vital information that helps you to assess the associated risk, and identify broader trends. Within the umbrella of DNS Data, the associated data are:

- NS records (Name server)
- A records (Address)
- MX (Mail Exchange)
- CNAME (Canonical Name)
- TXT (Text - a catch all for other information)

COLLECT: TOP CHALLENGES

TLD COVERAGE AND STANDARDIZATION

Unfortunately, there is no one provisioner of domains. In order to gather domain registration data, you will need to gather these from different top level domains. Be aware that there is no standardized format for these, so challenges can arise when you begin to analyze the data.

HISTORICAL DATA

It's one thing getting the right DNS data and context, but another accessing the historical data and tracking it over time. Attackers may change WHOIS information in order to hide links to other campaigns, so going back to view previous details can be highly valuable. Some security teams use Archive.org's Wayback Machine to get an idea of what the domain has looked like previously (https://archive.org/web/web.php).

ONGOING MONITORING AND STORAGE

Oftentimes, security teams will want to capture screenshots, analyze the contents of domains, and store historical DNS data. This type of historical data can be vital for quickly responding to risks associated with domain impersonation. It would be cost prohibitive to store all domains and their contents for all time, so security teams should be clear about how much data they wish to pay to store.

DOMAIN COLLECTION ON A SHOESTRING

Newly Registered Domains

- WHOISXML: https://newly-registered-domains.whoisxmlapi.com/

DNS Records

- MX Toolbox: https://mxtoolbox.com/

Certificate Transparency Logs

- **CertStream:** https://certstream.calidog.io/
- **Google Transparency Report:** https://transparencyreport.google.com/https/certificates
- **Crt.sh:** https://crt.sh/

Content Analysis

- Wayback Machine: https://archive.org/web/web.php

DETECT:
TYPES OF DOMAIN IMPERSONATION

Once you have collected your various sources, you can then analyze the domain names for those attempting to impersonate your own.

In this section, you'll read about the most common impersonation techniques that are used by threat actors and, spoiler alert, there are a lot.

Given all of these possible permutations, the obvious challenge for most security teams is to balance how many variations you search for with the noise it can generate.

GET STARTED WITH DNS TWIST

If you have never monitored for impersonating domains before, you can generate permutations with DNS Twist. This will give you a good idea of how many similar domains exist.

README.md

dnstwist

See what sort of trouble users can get in trying to type your domain name. Find lookalike domains that adversaries can use to attack you. Can detect typosquatters, phishing attacks, fraud, and brand impersonation. Useful as an additional source of targeted threat intelligence.

```
homoglyph     githulb.com  NS:ns1.dnsimple.com
homoglyph     gitlhub.com  162.255.119.242 NS:dns1.registrar-servers.com MX:eforward1.registrar-servers.com
homoglyph     github.com   !ServFail !ServFail NS:!ServFail
homoglyph     qithub.com   23.20.239.12 NS:nsg1.namebrightdns.com
homoglyph     gitihub.com  173.239.5.6 NS:ns1.expiereddnsmanager.com MX:mx7.gitihub.com
homoglyph     github.com   46.101.110.218 NS:ns1.reg.ru
homoglyph     glthub.com   3.234.181.234 NS:ns1.namebrightdns.com
hyphenation   gi-thub.com  NS:!ServFail
hyphenation   git-hub.com  192.30.253.167 NS:ns1.p16.dynect.net
insertion     gitbhub.com  67.227.226.240 NS:ns1.parklogic.com MX:mx156.hostedmxserver.com
insertion     giythub.com  199.59.242.153 NS:ns1.bodis.com MX:mx76.m2bp.com
insertion     gigthub.com  199.59.242.153 NS:ns1.bodis.com MX:mx76.m2bp.com
insertion     gitjhub.com  23.82.12.31 NS:ns1.thednscloud.com
insertion     gituhub.com  95.211.117.215 NS:ns1.hastydns.com
insertion     giuthub.com  95.211.219.67 NS:ns1.hastydns.com
insertion     githgub.com  199.59.242.153 NS:ns1.bodis.com MX:mx76.m2bp.com
insertion     githujb.com  66.96.162.128 NS:ns1.domain.com MX:mx.githujb.com
insertion     gitfhub.com  199.59.242.153 NS:ns1.bodis.com MX:mx76.m2bp.com
insertion     gijthub.com  199.59.242.153 NS:ns1.bodis.com MX:mx76.m2bp.com
insertion     guithub.com  78.41.204.33 NS:ns1.torresdns.com MX:mail.h-email.net
insertion     gikthub.com  199.59.242.153 NS:ns1.bodis.com MX:mx76.m2bp.com
insertion     gitrhub.com  5.79.79.209 NS:ns1.tacomadc.com
insertion     githuyb.com  173.239.22.42 NS:ns1.mnsdnsmart.com
```

DNS Twist: https://github.com/elceef/dnstwist

COMMON TECHNIQUES

There are two high level types of domain impersonation, typosquats and combosquats, but there are at least 10 different strategies used by threat actors.

Furthermore, security professionals should also consider non-Roman alphabet. Characters substitutions can also include Unicode characters, such as those from Cyrillic, Greek, or other international character sets.

OMISSION

A character is missing. In this case, the “i” is missing.

ADDITION

A character is added to the word. substituted for another character. In this case, another “s” added.

TYPOSQUAT: SUBSTITUTION

A character is substituted for another character. In this case, the “i” is swapped for the number “1”.

TRANSPOSITION

Multiple characters are swapped around. In this case, the first “s” and first “l” are swapped.

HYPHENATION

A hyphen is added to break up words.

HOMOGLYPH

A character is substituted for another similar-looking character. In this case, the roman alphabet “h” is swapped for the Cyrillic “н”.

HOMOPHONE

A similar sounding word is used in the impersonation i.e. which vs witch.

SUBDOMAIN

An attacker-owned subdomain references your company or assets. In this case, “digitalshadows” has been appended to the domain accounts-payable.com.

TLD SWAP

The same domain name is registered under a different top-level domain. In this case, it is .biz.

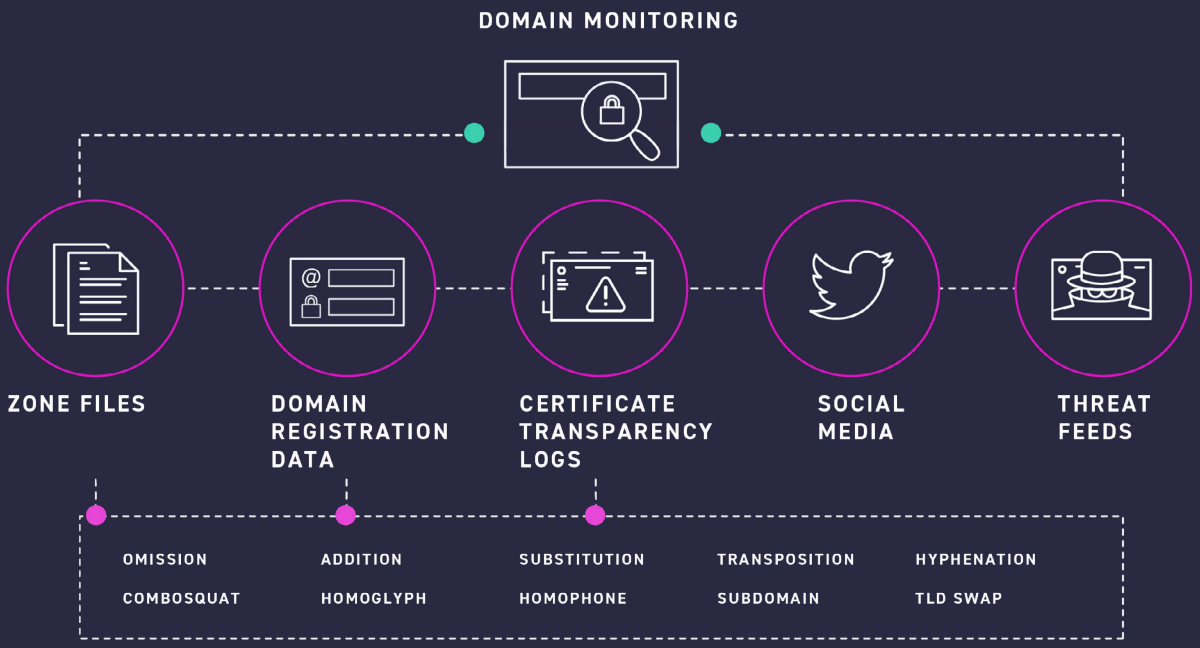
ADDED KEYWORDS

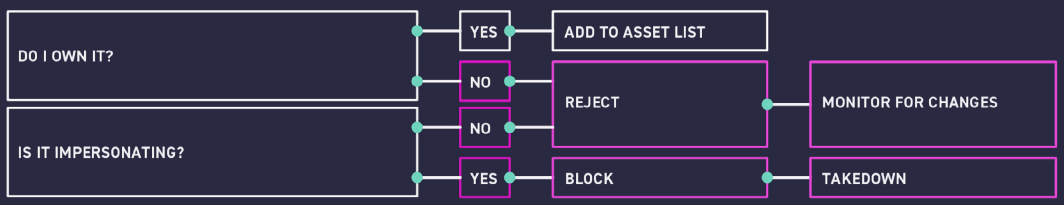
Another word is added on to the url. In this case, “login” has been added. Note: these are often industry-specific.

ANALYZE:
ASK THE KEY QUESTIONS

Once you have collected domains from a comprehensive variety of data sources and detected impersonations of your own domains, the next step is to analyze if those domains pose a legitimate threat to your business.

Domain ownership	Content Analysis	Threat Analysis
Has the domain been registered by your organization?	Is the content mimicking my website?	Could the domain send emails?
Is the WHOIS or DNS consistent with corporate owned websites?	Is the website selling counterfeit goods?	Is it attempting to capture credentials?
Is the WHOIS contact a member of your organization?	Is it just a parking page?	Has the domain been flagged on a threat feed before?





REMEDIATE

BLOCK THE DOMAIN

Block the domain at available Internet and email gateway(s) to prevent users browsing to the domain or receiving email from the domain. You should also consider sending the domain to Google Safebrowsing and Microsoft SmartScreen.

TAKEDOWN THE CONTENT

If the domain is hosting content that defrauds the organization, you may request to have the site taken down.

INVESTIGATE & CORRELATE RELATED ACTIVITY

Investigate the alert to identify associated activity: Correlate with other log sources to identify any activity related to the domain or IP Review other alerts related to the registrar or with the DNS data If there is associated activity, analyze it for potential increased threat. If additional domains are identified, it may be necessary to raise a new incident to manage the remediation.

PROACTIVE DOMAIN REGISTRATION

Consider purchasing possible variations of your domains proactively.

CONSULT WITH ICANN

If one of your domains is typosquatted, consult ICANN for inquiries into filing a Uniform Domain Name Dispute Resolution Policy (UDRP)

CONTACT APWG

Email US-CERT partners with the Anti-Phishing Working Group (APWG) to collect phishing email messages and website locations to help people avoid becoming victims of phishing scams.

GET STARTED

COLLECT

Open Phish: https://openphish.com/phishing_feeds.html

PhishThank: http://phishtank.org/developer_info.php

Blocklist.de: https://www.blocklist.de/en/index.html

UrlHaus: https://urlhaus.abuse.ch

CertStream: https://certstream.calidog.io/

Crt.sh https://crt.sh/

Google Transparency Report https://transparencyreport.google.com/https/certificates

MX ToolBox https://mxtoolbox.com/

Wayback Machine: https://archive.org/web/web.php

DETECT

DNS Twist
https://dnstwist.it/ or https://github.com/elceef/dnstwist

URLCrazy
https://github.com/urbanadventurer/urlcrazy

Phishing Catcher
https://github.com/x0rz/phishing_catcher

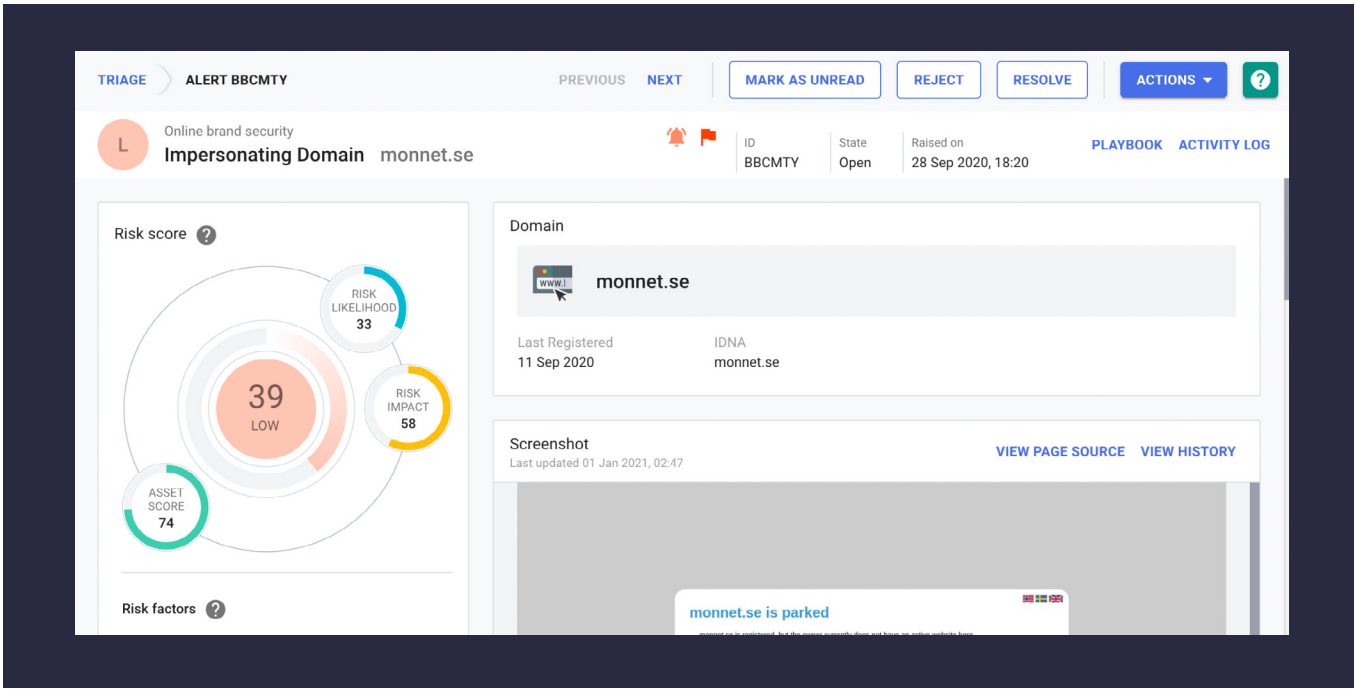
REMEDIATE

Report to Google Safebrowsing:
https://safebrowsing.google.com/safebrowsing/report_general/

Report to MS Smartscreen:
https://www.microsoft.com/en-us/wdsi/support/report-unsafe-site-guest

Report to US-CERT and the Anti-Phishing Working Group (APWG):
Email phishing-report@us-cert.gov.

THE DIGITAL SHADOWS APPROACH



SearchLight provides the most comprehensive, relevant and trusted domain monitoring capability. Triage effortlessly with superior classification and automated actions for domain alerts

AUTOMATIC IDENTIFICATION AND ANALYSIS OF RISK FACTORS

- **Parking page**
- **Has an MX Record**
- **Has DNS Record**
- **Hosting Content**
- **Has assets in content**
- **Has logos in content**
- **Domain in Threat Feed**
- **Newly registered when raised**
- **First seen recently**
- **Referencing website content**

ADDITIONAL CONTEXT

- **Screenshots**
- **DNS Records**
- **WHOIS information**
- **Timeline of changes**
- **HTML**

WHAT'S INCLUDED IN SEARCHLIGHT DOMAIN MONITORING:

ALERT TYPES

Impersonating Domain

Impersonating Subdomain

Phishing Webpage

COLLECTION SOURCES

Newly registered domains

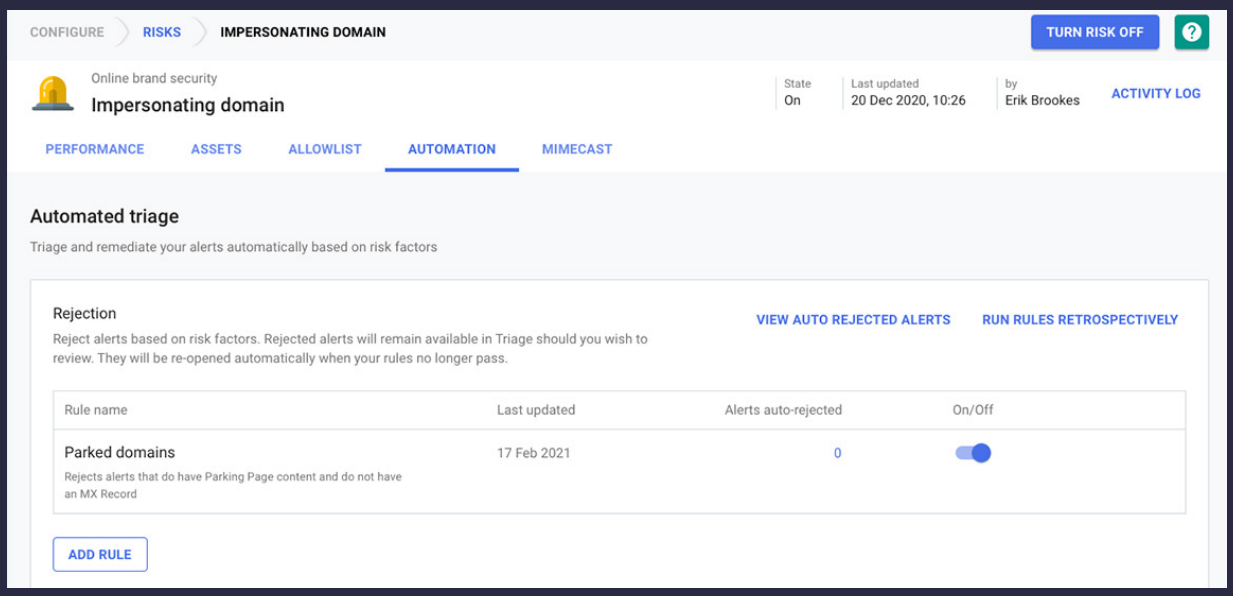
Certificate transparency logs

Phishing feeds

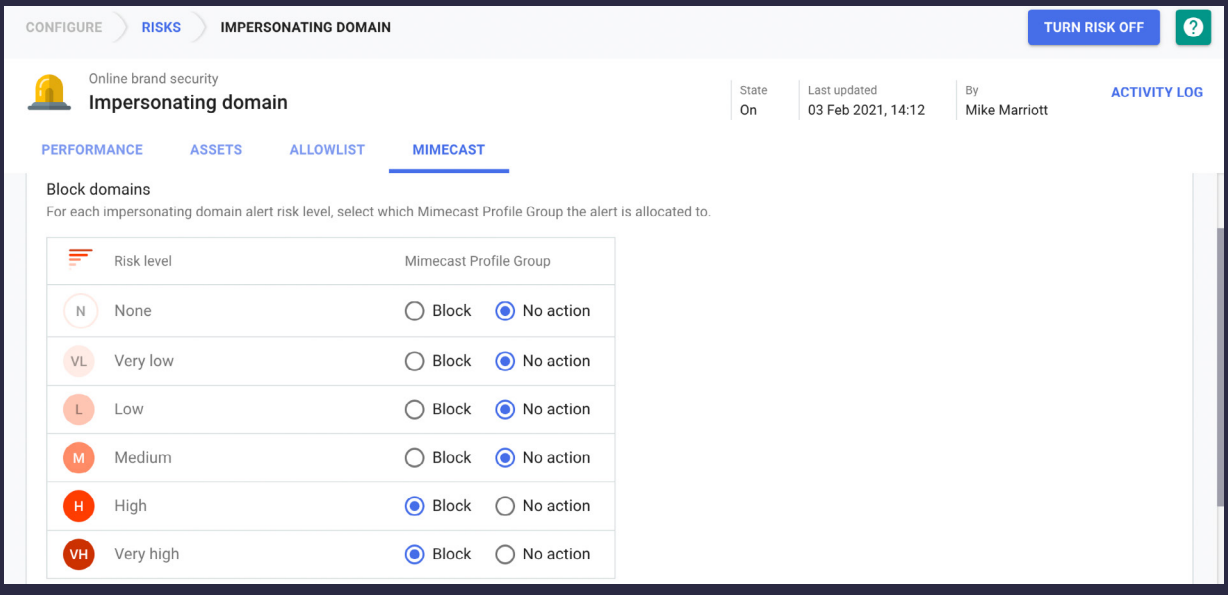
Social media

AUTOMATE 75% OF ALERT TRIAGE VIA TEMPLATED REJECTION RULES.

With SearchLight, you can automate the triage based on defined risk factors. For example, you can automatically reject all parking pages, saving you valuable time.



Automated triage, enabling users to automatically reject the domains of least interest



Automated domain blocking with Digital Shadows' Mimecast integration

THANK YOU_

digital shadows_

About Digital Shadows

Digital Shadows minimizes digital risk by identifying unwanted exposure and protecting against external threat. Organizations can suffer regulatory fines, loss of intellectual property, and reputational damage when digital risk is left unmanaged. Digital Shadows SearchLight™ helps you minimize these risks by detecting data loss, securing your online brand, and reducing your attack surface.

To learn more and get free access to SearchLight™, visit

www.digitalshadows.com

London
Columbus Building, Level 6,
7 Westferry Circus,
London, E14 4HD
+44 (0) 203 393 7001

San Francisco
235 Pine St. Suite 1050,
San Francisco, CA 94104
+1 (888) 889 4143

Dallas
5307 E. Mockingbird Ln.
Suite 200
Dallas, TX 75206