WMC GLOBAL

WORLDWIDE BRAND AND CUSTOMER PROTECTION AT SCALE

UrSULA
API DOCUMENTATION

v1.0.2 08.09.2024

WMC GLOBAL

TABLE OF CONTENTS

GETTING STARTED WITH UrSULA

Threat actors increasingly use mobile as their preferred attack channel, as customers become more dependent on “small screen” convenience and high open rates observed over SMS. The daily distractions of life make users especially vulnerable on their handheld devices. SMS firewalls cannot comprehensively protect customers from ongoing credential phishing attacks because of an inability to accurately and consistently differentiate credential phish from legitimate links at scale.

The URL Scanning & Universal Link Attribution (UrSULA) API is a high-volume synchronous URL analysis service for identifying credential phishing URLs. The API analyzes, classifies, and scores URLs and can process hundreds of millions of URLs a day to determine if they are live credential phishing attacks. The UrSULA API:

- Provides rapid identification, scoring, and verification of suspected credential phishing URLs.
- Analyzes URLs, redirects, and landing pages using machine learning and validated credential phishing attack intelligence
- Enhances security orchestration, automation, and response (SOAR) with an immediate analysis.
- Offers URL submission from any source email, social media, SMS, messages, or email gateway logs, firewall, and referral logs, to gain visibility on credential phishing.
- Actively circumvents phishing countermeasures such as JavaScript and user agent redirects.

With UrSULA you can:

- Identify credential phishing missed by other vendors.
- Utilize URL scoring for better security orchestration, automation, and response (SOAR) tuning
- Receive immediate submission feedback via API as well as monthly submission reporting.

WMC Global Website

DOCUMENT STATUS

This document supports v1.0 for the URL Scanning & Universal Link Attribution (UrSULA) API. The API and documentation are subject to change.

UrSULA: HOW DOES IT WORK?

SUBMIT A URL to UrSULA

- UrSULA accepts `HTTP` and `HTTPS` protocol specified in a submitted URL. If no identifiable protocol is submitted, UrSULA will default to `HTTPS`.
- See UrSULA Data Dictionaries

UNDERSTANDING URL CLASSIFICATION AND SCORING

- The UrSULA API scores and classifies submitted URLs to determine whether submitted URLs are credential phishing.
    - See UrSULA Classification and Scoring
- The UrSULA URL score measures suspicious credential phishing indicators on a submitted URL and its content. The score can be used in conjunction with the classification to dial in an organization's tolerance for suspicious content beyond classification labels.
- Some Media formats will not be scanned due to the nature of their format.

URL REDIRECTS

URL Redirects or URL forwarding is a technique used to forward visitors to a website via a different URL than the website's domain. There are many legitimate uses of URL forwarding, but threat actors also utilize it to hide credential phishing URLs from scanning. UrSULA follows URL redirects and returns analysis about the final landing page of the redirect.

ERROR RESPONSES

UrSULA API returns a JSON error response with an error code and an appropriate HTTP status code in the response header for error conditions. See UrSULA Responses

UrSULA CLASSIFICATION AND SCORING

To get a better understanding of how UrSULA handles different types of content please see the examples listed here.

URL SCORING

The UrSULA URL score measures suspicious credential phishing indicators on a submitted URL and its content. The score can be used in conjunction with the classification to dial in an organization's tolerance for suspicious content beyond classification labels.

URL CLASSIFICATION

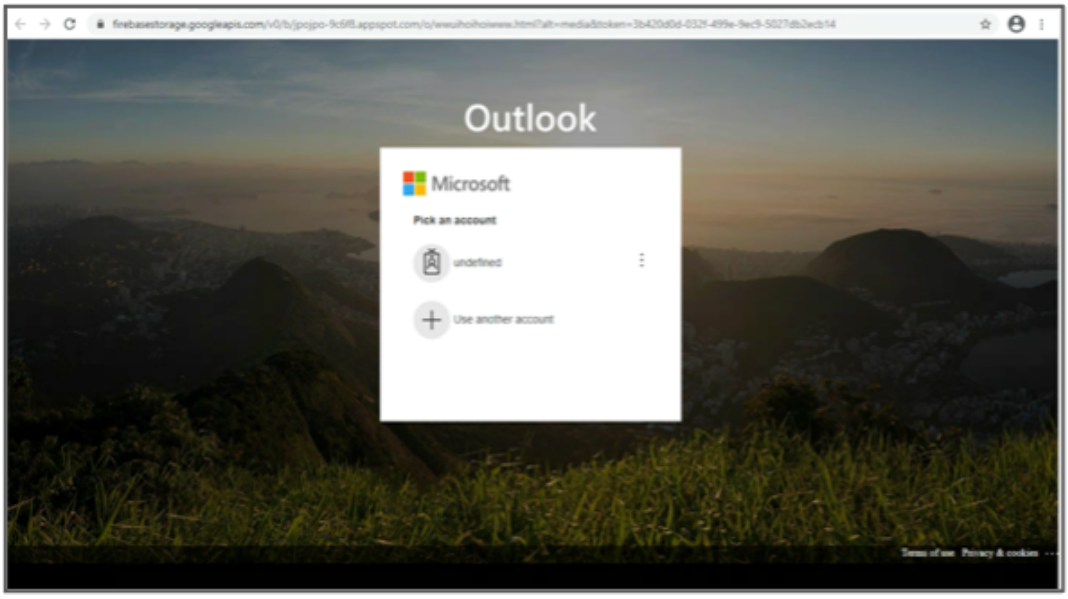
STATUS	CLASSIFICATION	SCORE	DESCRIPTION
Live	Phishing	65 to 99	UrSULA determined the URL as phishing.
Live	NotPhishing	49.99 and lower	UrSULA identified the URL as not phishing.
Live	Suspicious	50 to 64.99	UrSULA determined the page contains suspicious characteristics. It could not be confirmed as malicious.
BarelyLive	Suspicious	50 to 64.99	UrSULA determined the page contains suspicious characteristics. It may not be live yet or be taken down by the hosting provider.
SuspiciousRedirect	Suspicious	50 to 64.99	UrSULA determined the URL resulted in a redirect that did not result in a landing page.
Suspicious	Suspicious	50 to 64.99	UrSULA determined the page displays characteristics known to impact automated detection results.
Dead	Suspicious	39.99	UrSULA determined the URL could not be resolved.
WhiteListedDomain	NotPhishing	0	This page has been added to the pass/whitelist.

CLASSIFICATION EXAMPLES

!!! CLASSIFICATION EXAMPLES DISCLAIMER

IN THE EXAMPLES, THE HTTP:// AND HTTPS:// PROTOCOLS HAVE BEEN CHANGED TO PREVENT ACCIDENTAL TRAVERSAL. THE PHISHING LINKS WERE TO REAL PHISH. DO NOT ATTEMPT TO OPEN PHISHING URLS WITHOUT GUIDANCE FROM YOUR SECURITY TEAM.

PHISHING WEBSITE



Microsoft Phishing Landing Page

INPUT JSON

JSON

```
{
"url": "hxxps://firebasestorage.googleapis.com/v0/b/jpojpo-
9c6f8.appspot.com/o/wwuihoihoiwww.html? alt=media&token=3b420d0d-032f-499e-9ec9-
5027db2ecb14"
}
```

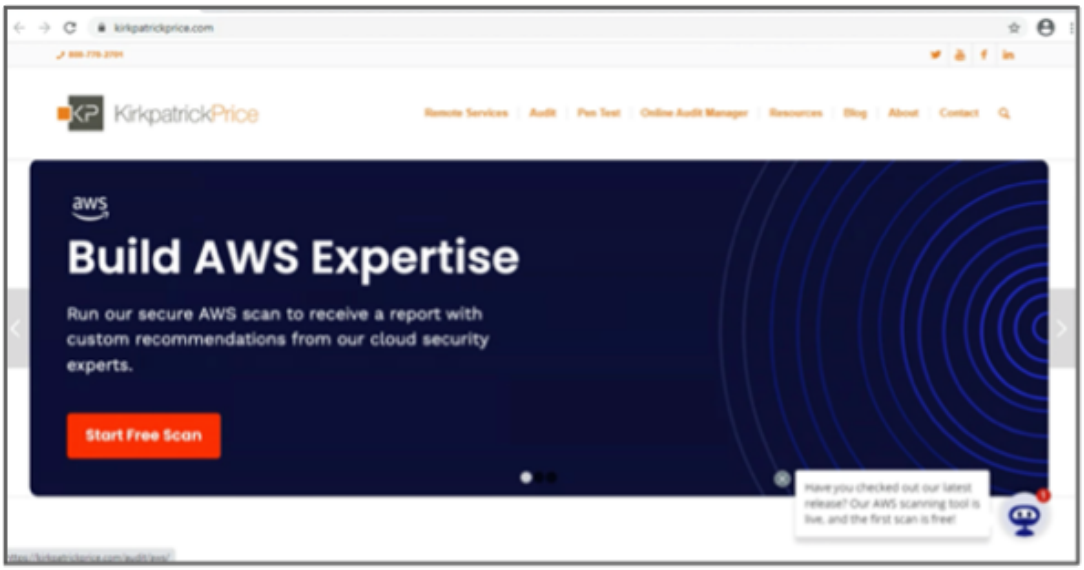
RESPONSE JSON

JSON


```
{
"url": "hxxps://firebasestorage.googleapis.com/v0/b/jpojpo-
9c6f8.appspot.com/o/wwuihoihoiwww.html?alt=media&token=3b420d0d-032f-499e-9ec9-5027db2ecb14",
  "submission_id": "c311sv8ujj8g00elmvh0",
  "classification_timestamp": "2021-06-10T14:15:25Z",
  "status": "Live",
  "score": 99,
  "classification": "Phishing"
}
```


The URL has been classified as phishing.

NONPHISHING WEBSITE



INPUT JSON

JSON

```
{
"url": "hxxps://kirkpatrickprice.com/"
}
```

RESPONSE JSON

JSON


```
{
  "url":
  "hxxps://kirkpatrickprice.com/",
  "submission_id":
  "c312f4je7reg00eiugl0",
  "classification_timestamp": "2021-06-
  10T14:54:10Z", "status": "Live",
  "score": 29,
  "classification": "NotPhishing"
}
```


The URL has been classified as not phishing.

SUSPICIOUS REDIRECT URL

INPUT JSON

JSON

```
{
"url": " hxxps://app.shopkeep.com/my/receipts/51215AE2-DB6E-4AE3-949E-49B152FA33B8"
}
```

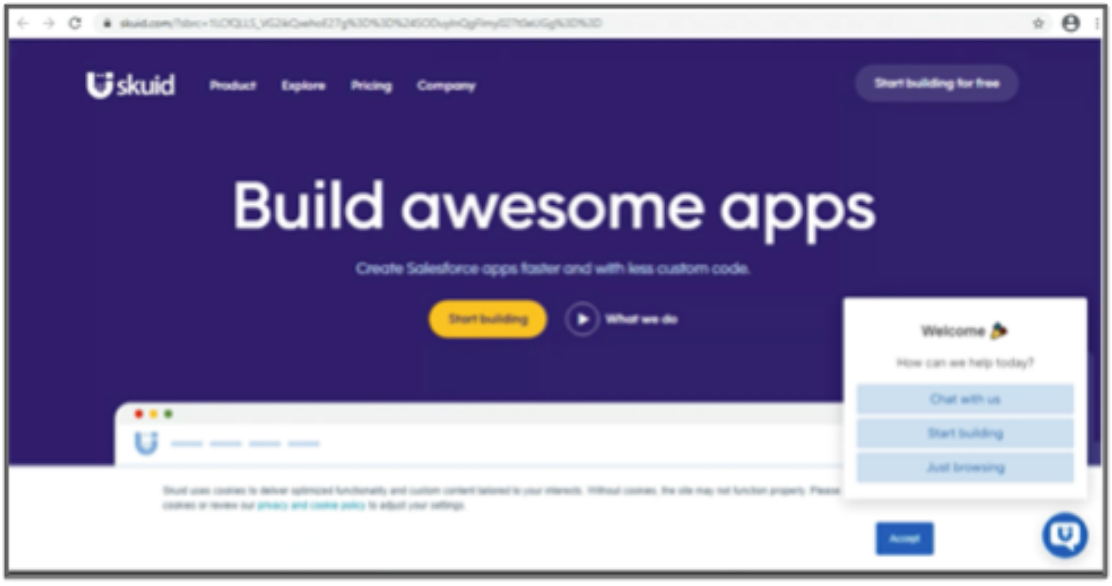
RESPONSE JSON

JSON


```
{
"url": "hxxps://app.shopkeep.com/my/receipts/51215AE2-DB6E-4AE3-949E-49B152FA33B8",
  "submission_id": "c314q0v5jlqg00djt3eg",
  "classification_timestamp": "2021-06-10T17:33:55Z",
  "status": "SuspiciousRedirect",
  "score": 11,
  "classification": "Suspicious"
}
```


The URL resulted in a redirect that did not result in a landing page.

REDIRECT URL



Note that this redirect points to the target `hxxps://www.skuid.com/?sbrc=1LCfQLLS_VG2ikQxehoE27g%3D%3D%24SODuy1nQgFlmy027t0eUGg%3D%3D`

INPUT JSON

JSON


```
{
"url": " hxxps://salesloft.skuid.com/t/103200/c/30f68a68-c49b-49bb-bb98-
28d955fea115/NB2HI4DTHIXS653XO4XHG23VNFSC4Y3PNU7XGYTSMM6TCTCDMZIUYTCTL5LEOMTJNNIXQZLIN5CTEN3
HEUZUIJJTIQS
skuid-com"
}
```


RESPONSE JSON

JSON


```
{
"url": "hxxps://salesloft.skuid.com/t/103200/c/30f68a68-c49b-49bb-bb98-
28d955fea115/NB2HI4DTHIXS653XO4XHG23VNFSC4Y3PNU7XGYTSMM6TCTCDMZIUYTCTL5LEOMTJNNIXQZLIN5CTEN3HE
UZUIJJTIQS
skuid-com",
  "submission_id": "c312e9ovd0pg009lehvg",
  "classification_timestamp": "2021-06-
  10T14:52:23Z", "status": "Live",
  "score": 17,
  "classification": "NotPhishing"
}
```


A redirect URL was submitted, traversed, and classified as not phishing.

SUSPICIOUS URL

INPUT JSON

JSON

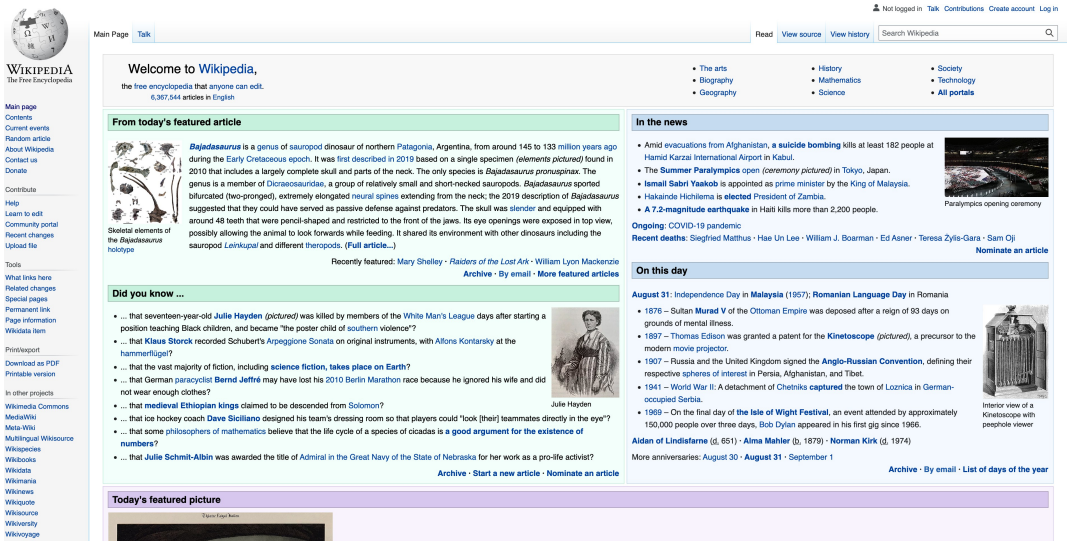

```
{
"url": " hxxp://customer.anywheremight.com/mw/2552d49dd0444d44aa5eb445970807cf.php"
}
```


RESPONSE JSON

JSON


```
{
  "url": "hxxp://customer.anywheremight.com/mw/2552d49dd0444d44aa5eb445970807cf.php",
  "submission_id": "c313ni56n8vg00ckq5pg",
  "classification_timestamp": "2021-06-10T16:20:24Z",
  "status": "Live",
  "score": 52,
  "classification": "Suspicious"
}
```

WHITELISTED PAGE



INPUT JSON

JSON

```
{
"url": "hxxps://en.wikipedia.org/wiki/Main_Page"
}
```

RESPONSE JSON

JSON


```
{
  "url":
  "hxxps://en.wikipedia.org/wiki/Main_Pa
  ge", "submission_id":
  "c313at5p9u6000cf6qcg",
  "classification_timestamp": "2021-06-
  10T15:53:24Z", "status":
  "WhiteListedDomain",
  "score": 0,
  "classification": "NotPhishing"
}
```

PAGE DEAD

INPUT JSON

JSON


```
{
"url": "hxxps://mailing.ing.com/api/query/Messenger/TrackerQuery?
Uri=HJ7SXwYidNgL9ZjF3mf3E29avanit9DptA6NmwziRI5odHRwczovL3d3dy5pbmcuYmUvZnIvcmV0YWlsL1BhZ2VzL3NlY3VyaXR5
}
```


RESPONSE JSON

JSON


```
{
  "url":
"hxxps://mailing.ing.com/api/query/Messenger/TrackerQuery?Uri=HJ7SXwYidNgL9ZjF3mf3E29avanit9DptA6NmwziRI5odHRwczovL3d3dy5pbmcuYmUvZnIvcmV0YWlsL1BhZ2VzL3NlY3VyaXR5LmFzcHg%3D&amp%3BData=8qEGiXyTnZXId0EmR3zIn55QNEHOgWhH55l%2Bc5W2kRE3NTcyMTA2NDMtOTQ1OS0wo",
  "submission_id": "c312ljdat8pg00ccus10",
  "classification_timestamp": "2021-06-10T15:07:57Z",
  "status": "Dead",
  "score": 50,
  "classification": "Suspicious"
}
```

PAGE CONTENT TAKEN DOWN

INPUT JSON

JSON

```
{
"url": " hxxps://rytujyj.lh0wt4knix.workers.dev/ "
}
```

RESPONSE JSON

JSON

```
{
  "url": "hxxps://rytujyj.lh0wt4knix.workers.dev/",
  "submission_id": "c313jvtkaqlg008l9mj0",
  "classification_timestamp": "2021-06-10T16:12:47Z",
  "status": "BarelyLive",
  "score": 50,
  "classification": "Suspicious"
}
```

UrSULA DATA DICTIONARIES

DATA DICTIONARY INPUT PARAMETERS

PARAMETER NAME	DESCRIPTION	DATA TYPE
x-api-key	User API key; a unique identifier used to authenticate a user issued by WMC Global.	string

DATA DICTIONARY SEARCH PARAMETERS

PARAMETER NAME	DESCRIPTION	DATA TYPE
url	Submit a URL for processing.	string

DATA DICTIONARY OUTPUT PARAMETERS

PARAMETER NAME	DESCRIPTION	DATA TYPE
url	Submitted URL.	string
submission_id	UrSULA submission identifier.	string
score	UrSULA phishing indicator score. A user can utilize this score in addition to the classification to adjust an organization's tolerance for phishing-related content.	string
status	The status of the referenced URL. Statuses are Live, Dead, BarelyLive, WhiteListedDomain, and SuspiciousRedirect.	string
classification	The result of UrSULA's automatic phishing detection. URLs are classified as Phishing, NotPhishing, and Suspicious	string
classification_timestamp	The time in which the URL was scored and classified.	string

Responses

BASE URL: api.phishfeed.com/ursula/v1

SCHEME: HTTPS

API: POST/url-class

CONTENT TYPE: application/JSON

RESPONSE CODE	DESCRIPTION
200	success

JSON

```
{
  "url": "www.definitelyphishing.ru",
  "submission_id": "c311sv8ujj8g00elmvh0",
  "classification_timestamp": "YYYY-mm-ddTHH:MM:SSZ",
  "status": "Live",
  "score": 90.1,
  "classification": "Phishing"
}
```

RESPONSE CODE	DESCRIPTION
214	no_classification

JSON

```
{
  "url": "www.definng.ru",
  "submission_id": "c311sv8ujj8g00elmvh0",
  "classification_timestamp": "YYYY-mm-ddTHH:MM:SSZ",
  "code": "1003",
  "message": "No classification"
}
```

RESPONSE CODE	DESCRIPTION
400	bad_request_parameters

JSON

```
{
  "code": "1025",
  "message": "HTTP Request body contains bad or incomplete URL or data"
}
```

RESPONSE CODE	DESCRIPTION
403	invalid_api_key

JSON

```
{
  "code": "1005",
  "message": "Missing or invalid API key"
}
```

RESPONSE CODE	DESCRIPTION
415	unsupported_media_type

Unsupported Media Examples

Links to certain direct types of media cannot be processed. See UrSULA Unsupported Media.

JSON

```
{
  "url": "www.definitelyphishing.ru",
  "submission_id": "c311sv8ujj8g00elmvh0",
  "classification_timestamp": "YYYY-mm-ddTHH:MM:SSZ",
  "code": "1015",
  "message": "Media type not supported"
}
```

RESPONSE CODE	DESCRIPTION
429	quota_exceeded

JSON

```
{
  "code": "1045",
  "message": "Too many requests"
}
```

RESPONSE CODE	DESCRIPTION
500	system_error

JSON

```
{
  "url": "www.definitelyphishing.ru",
  "submission_id":
  "c311sv8ujj8g00elmvh0",
  "classification_timestamp": "YYYY-mm-
  ddTHH:MM:SSZ", "code": "1002",
  "message": "A system error has occurred; please try again later or contact support"
}
```

UrSULA UNSUPPORTED MEDIA

Some URLs lead to specific content that cannot be processed for credential phishing due to its format. "Multipurpose Internet Mail Extensions" or MIME type is a standard that indicates the nature and format of a document, file, or assortment of bytes.

UrSULA will process content with text/ MIME type and attempt to process unknown media types.

UNSUPPORTED MIME TYPES AND EXAMPLES

MIME TYPE (Linked to full list)	DESCRIPTION	MEDIA TYPE EXAMPLES	UrSULA RESPONSE
image/	Bitmap, vector still images, and animated image formats.	PNG, JPEG, GIF	Media type not supported
audio/	Audio or music data.	MP3, AIFF, WAV, DSD, FLAC, MP4, WMA	Media type not supported
video/	Video data or files.	MP4, MOV, AVI, MPEG-4	Media type not supported
font/	Font/typeface data.	WOFF, TTF, OTF	Media type not supported
application/	Binary data that doesn't fall explicitly into one of the other types.	APK, EXE, JSON, EPUB, ZIP, GZIP, RAR, Z7	Media type not supported
multipart	A document comprised of multiple component parts, each of which may have its own individual MIME type; or, a multipart type may encapsulate multiple files being sent together in one transaction.	Email attachments	Media type not supported

message/	A message that encapsulates other messages.	Forwarded or replied-to message quoting	Media type not supported
model/	Model data for a 3D object or scene.	3MF VRML	Media type not supported
No type or unknown type	Media with no type set or with a new and unclassified type that UrSULA cannot process.		Status: Suspicious Classification: Suspicious

FULL TYPE LIST

The current IANA media type list is available at www.iana.org.

ADDITIONAL RESOURCES

Learn more about MIME Types on the Mozilla Developer Network.

UrSULA BEST PRACTICES

UrSULA TRAFFIC BEST PRACTICES

The Internet is a massive place with a lot of different content, and UrSULA's flexible configuration allows it to adjust to different environments to achieve the desired automation within budget. When deciding what to send to UrSULA, here are some considerations to minimize costs and maximize results.

Deduplicate URLs

UrSULA analyzes URLs in real-time, transversing redirects and getting to the link's landing page. When identifying what traffic to send to UrSULA, sending unique URLs will reduce the total traffic and cost of detection and focus detection on unique submissions, discovering more threats faster to improve automated downstream response. Some customers have reduced overall traffic by over 20%.

Exclude Media

Some URLs make it possible to determine what content they direct to. URLs linking directly to content like PDFs, images, or other media can be excluded from UrSULA submission. See UrSULA Unsupported Media for more information on what content is not supported.

Identify Known Credential Phishing URLs

Utilize PhishFeed to identify and block known credential phishing URLs and domains before submitting them to UrSULA for detection. PhishFeed captures thousands of new credential phishing URLs a day. Regularly pulling this information and caching before UrSULA submission can help substantially reduce the UrSULA submissions and speed up responses to credential phishing sites.

Identify Risky Traffic

Threat actors often use free or low-cost services to conduct phishing attacks. Identifying traffic that

might be at higher risk of abuse due to account age, costs, reputation, or lack of technical and regulator controls can help target phishing detection where it is needed most.
Traffic can also be "A to B" tested to compare phishing activity in different traffic segments to pinpoint where detection is needed.

Traffic Sampling

Not all traffic is equal. Sampling smaller portions of traffic can help spot-check traffic regularly for issues. Automation can help dial up or down submissions based on sampled traffic findings or other environmental factors to best fit.

UrSULA INTEGRATION BEST PRACTICES

Including the Response Code and Descriptions

The response code identifies system responses. A `200` response means that the URL was evaluated and classified. Other codes like `214`, `400`, `403`, `415`, `429`, and `500` indicate different system responses that have a variety of meanings. Including the response code can help with auditing and troubleshooting.

- **Include Response Codes**

  Including the codes will let analysts reference troubleshooting information, work directly with WMC Global support, and prevent them from involving engineering resources when working through issues.

- **Include Descriptions**

  Including the response code description can clue users into the meaning of the response without having to look up reference documentation or involve engineering resources. This can be helpful when analysts don't have access to the integration information and need additional context to the response they are getting.

- **Handle** `unsupported_media_type`

  Some URLs will lead directly to PDFs, images, or other file types. UrSULA will respond with `415` `unsupported_media_type` instead of a `classification` because the landing page's content cannot be assessed. It's important that this information be made available for analysts auditing the system as well as used as consideration for downstream automation.

See UrSULA Unsupported Media for more details.

- **Error Handling**

  Including the response code is essential for error handling within the API. The absence of a phishing `classification` or `score` does not indicate that a URL is safe. Things can go wrong with submissions. Error handling should identify how the system decides to handle URL submissions that return an error. It should also be made available for analysts to troubleshoot to reduce the involvement needed from engineering resources.

  See UrSULA Responses for more examples.

200 Successful Classifications

URL Classification Reference

STATUS	CLASSIFICATION	SCORE	DESCRIPTION
Live	Phishing	65 to 99	UrSULA determined the URL as phishing.
Live	NotPhishing	49.99 and lower	UrSULA identified the URL as not phishing.
Live	Suspicious	50 to 64.99	UrSULA determined the page contained suspicious characteristics. It could not be confirmed as malicious.
BarelyLive	Suspicious	50 to 64.99	UrSULA determined the page contains suspicious characteristics. It may not be live yet or be taken down by the hosting provider.
SuspiciousRedirect	Suspicious	50 to 64.99	UrSULA determined the URL resulted in a redirect that did not result in a landing page.
Suspicious	Suspicious	50 to 64.99	UrSULA determined the page displays characteristics known to impact automated detection results.

Dead	Suspicious	39.99	UrSULA determined the URL could not be resolved.
WhiteListedDomain	NotPhishing	0	This page has been added to the pass/whitelist.

Include the classification `status` field.

The URL status identifies if the site was `Live`, `Dead`, `Barelylive`, `Suspicious`, or `SuspiciousRedirect`, along with the `Classification` for better context. The response `status` should be available in the systems used for analysts or engineers auditing and troubleshooting the system UrSULA is integrated with. This will provide analysts with better information for troubleshooting and prevent them from involving engineering resources when working through issues.

For example:

- An organization that knows that all its URL traffic should be live may use automation to investigate accounts sending `dead` links.

See UrSULA Classification and Scoring more information.

Response Classification

The URL Classification is combined with the URL `status` to provide better context for the nature of the URL provided.

See UrSULA Classification and Scoring more information.

Response Score

The UrSULA URL `score` measures suspicious phishing indicators on a submitted URL and its content. The score can be used in conjunction with the classification to dial in an organization's tolerance for suspicious content beyond classification labels.

For example:

- An organization may determine they want to conduct automated blocking on phishing with a score over 70 but have phishing between `65-69` delayed, quarantined, or reviewed for additional actions.

See UrSULA Classification and Scoring for more information and content examples.

Include the UrSULA ID

The UrSULA ID is an identifier that can help analysts coordinate directly with WMC Global support for troubleshooting issues. A single URL can be submitted many times within a time period. To troubleshoot the exact submission, the UrSULA ID can help pinpoint the exact submission quickly to speed up analysis.

UrSULA FAQ

CLASSIFICATION

How does WMC Global determine a site is `dead` at the time of the scan?

UrSULA first sends a request to the website and evaluates the header response as part of the connection request. Based on the website response UrSULA routes the traffic to provide both the most timely and accurate information. When UrSULA receives a header response indicating the page is `dead`, it utilizes a secondary resource that attempts additional verification if it is `dead` or obfuscated.

Does UrSULA make distinctions between different types of phishing?

UrSULA examines the web page content to identify credential phishing. Phishing attacks use various tactics, techniques, and procedures (TTPs), including diverse hosting infrastructure, redirect chains, and technology, to deliver phishing to users. Because these TTPs often overlap between campaigns and phishing kits, UrSULA does not distinguish between targeted spear credential phishing and general credential phishing because the TTPs overlap.

Users can use the UrSULA responses to support their internal automation to examine for specific similarities or differences between URL structures to identify particular types of attacks.

URL SUBMISSION

Will trailing punctuation such as . or , cause detection issues?

Trailing punctuation may impact UrSULA's ability to execute a URL and receive a response. This can cause the system to believe the URL is inactive and return a 'Dead' classification.

How does UrSULA treat a submission if it is missing the protocol?

UrSULA will prepend submissions that do not have a protocol with `https://`. If a protocol has been added, the response will reflect the prepend `https://` with the URL submission.

How will UrSULA treat extra characters or whitespace submitted with the URL?

UrSULA processes legitimate URLs and returns responses based on an evaluation of the URL's content. UrSULA does not sanitize URL content due to the many upstream client environments and processes that generate URLs differently. Adding characters or whitespace anywhere in the URL submission can prevent the URL from being read accurately and impact classification.

How does UrSULA handle UTF and ASCII characters?

UrSULA accepts UTF-8 encoded submissions and utilizes the Internationalized Domain Names for Applications (IDNA) Framework to handle URL encoding translation to better mimic browser functionality and handling of URLs. Internationalized Domain Names (IDNs) were created to support non-Latin alphabets better. IDNs support arbitrary Unicode characters in hostnames in a backward-compatible way. User agents transform hostnames containing non-ASCII Unicode characters into an ASCII-only hostname. It is then sent to DNS servers.