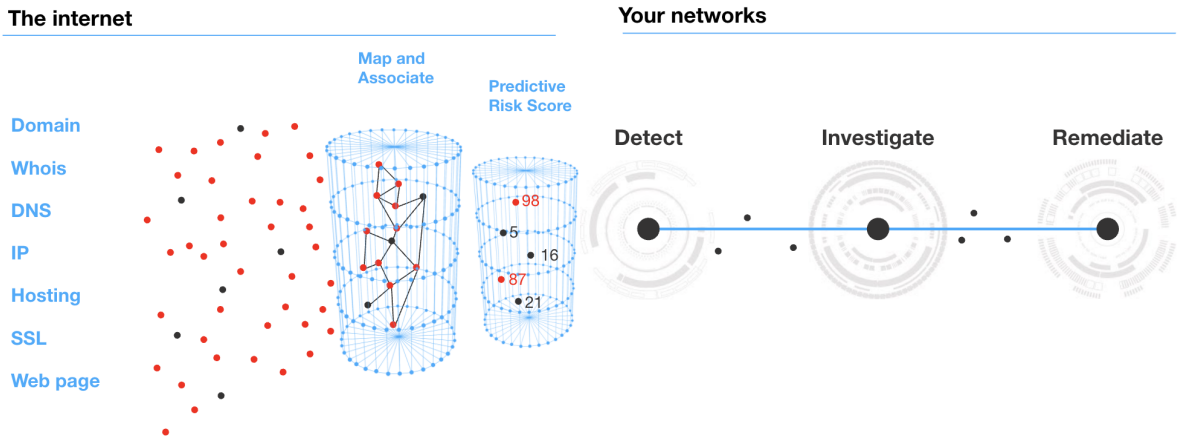
Technical Brief

DomainTools Hotlists and Feeds

DOMAINTOOLS®

Pinpoint and characterize the most dangerous infrastructure on the Internet

When it comes to network defense, **a major part of the risk involves traffic from the protected environment to threat-actor-controlled assets**. Connections from trusted users to hostile domains or IP addresses enable malware downloads or command and control, data exfiltration, espionage, and other threat activities. Preventing users from reaching dangerous infrastructure, while supporting necessary business functions, is a major component of any network and endpoint defense strategy. Therefore, security teams need reliable inputs on the risk level of the domains and IP addresses seen in their traffic flows, in order to improve situational awareness and to ward off incursions that may be underway.



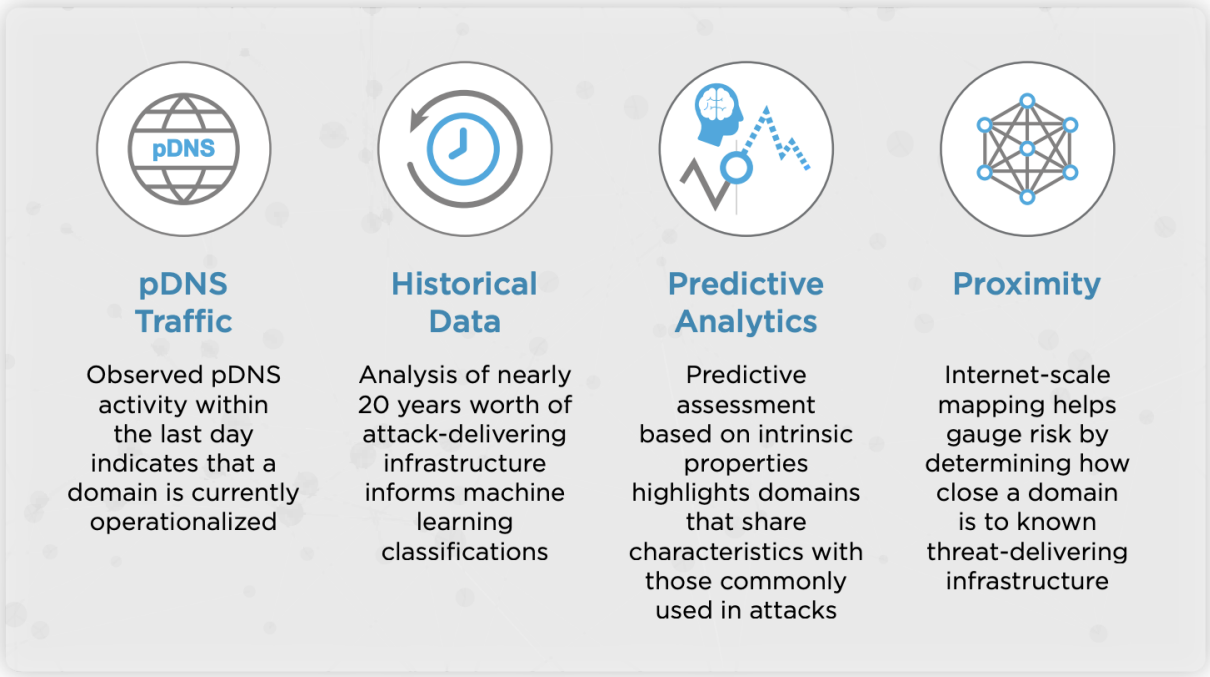
DomainTools offers a variety of feeds, each with a specific area of focus, to help with these needs:

- **Domain Hotlist**: a daily feed of high-risk domains that are observed to be active within a 24 hour time window
- **Domain Risk Feed**: a daily feed of high risk domains, regardless of observed traffic
- **IP Hotlist**: a daily feed of high-risk IP addresses hosting hostile domains that are observed to be active within a 24 hour time window, with risk scores and other enrichment data
- **Hosting IP Risk Feed**: a daily feed of all IP addresses known to be hosting domains, with risk scores and other enrichment data
- **Domain Discovery Feed**: a daily feed of all newly registered and newly observed domains
- **5-Minute Domain Whois Feed**: a feed of the most recently registered or changed domains as processed on a 5-minute basis
- **5-Minute IP Whois Feed**: A feed of the most recently updated IPv4 Whois records as processed on a 5-minute basis

Domain Risk Products

Domain Hotlist

The **Domain Hotlist** is designed to identify the riskiest population of domains, based on a combination of Risk Score and recent activity, as observed in passive DNS (pDNS) records. To qualify for the Hotlist, a domain must have a Proximity score of 70+ or a Threat Profile score of 90+, and must be observed in pDNS to have received traffic within the preceding 24 hours. Typical Hotlist size is ~900,000 domains.



Domain Hotlist Specifications

Updated: Daily

Inclusion Thresholds: Proximity score of 70+ or Threat Profile score of 90+; pDNS activity within 24 hours

File format: gzip-compressed tab separated text file

Data format: one domain per line with component risk scores

Feed size (typical): ~900,000 domains, ~3.5 MB compressed

Field	Description
Domain name	Pinpoint spoof or otherwise suspicious domain names
Phishing	Machine learning classifier prediction for phishing
Malware	Machine learning classifier prediction for malware
Spam	Machine learning classifier prediction for spam
Proximity	Indicates shared registration or infrastructure with known-bad domains
Overall	Equals highest of the component scores

The **Domain Risk Feed** is broader in scope than the Domain Hotlist, and does not include the criterion of pDNS activity. This feed can help identify infrastructure that is in the preparation phases but is not yet operationalized by a threat actor (which is why these domains may not have activity recorded in pDNS). This feed includes all domains with risk scores above 70.

Domain Risk Feed Specifications

Updated: Daily

Inclusion Threshold: Combined Domain Risk Score of 70 or higher

File format: gzip-compressed tab separated text file

Data format: one domain per line with component risk scores

Feed size (typical): ~30 million domains, ~400MB compressed

Field	Description
Domain name	Pinpoint spoof or otherwise suspicious domain names
Phishing	Machine learning classifier prediction for phishing
Malware	Machine learning classifier prediction for malware
Spam	Machine learning classifier prediction for spam
Proximity	Indicates shared registration or infrastructure with known-bad domains
Overall	Equals highest of the component scores

Who Uses The DomainTools Risk Feeds?

As users who demand high-confidence domain risk assessments at scale, the following are all users of the Risk Feeds:

- Online platforms with user communities in the hundreds of millions
- Cyber security technology companies
- Financial services organizations
- National and international-level law enforcement

Domain Discovery Feed

Thousands of malicious domains are registered and used every day for phishing, ransomware, credential harvesting, fraud, and more. As a result, many security teams now use a domain's age as a signal of risk, with brand-new domains standing out for extra scrutiny. The Domain Discovery Feed is a simple text file of domain names. This gives you maximum flexibility for using the new domain information to create alert or block rules for network or host defenses. Security Information Event Management (SIEM) platforms, Threat Intelligence Platforms (TIP), and a variety of other log and event aggregation sources can capture domains accessed from the protected environment; scripts which check these domains against the Domain Discovery Feed can then raise alerts when traffic to matching domains is observed. In some environments, a zero-trust policy toward new domains is employed; in such cases, the Domain Discovery Feed can enable the creation of automatic blocking rules for most traffic, or quarantine/inspection rules for SMTP and other protocols that can accommodate various dispositions.

Domain Discovery Feed Specifications

Updated: Daily

File format: gzip-compressed csv file

Data format: one domain name per line

Feed size (typical): ~375,000 domains, ~2.5GB compressed

Ideal Applications - Domain Risk Products

Whereas lower-volume, ad-hoc lookups can call the DomainTools Risk Score API, many applications operate at scales that demand the immediacy of an on-site database. These include:

- DNS RPZ implementations (also known as DNS Firewalling)
- IDS/IPS
- Fraud detection systems
- Online platforms supporting user-contributed URLs (such as social media platforms)

DomainTools makes it easy to put the power of Domain Risk Score to work in any technology stack. With various ways to access the data (including feed and API services), white-glove customer support, and access to DomainTools engineering and data science experts, any organization requiring high-confidence, Internet scale risk assessment can benefit from the DomainTools Risk Feed.

> *"This has been a very successful threat source for Quad9. We are very selective and DomainTools quickly established itself as one of our top-producing data sources. [I]t's clearly been a big win for helping to keep our users safe."*
>
> - John Todd, Executive Director, Quad9

IP Risk Products

Illuminating Blind Spots

Existing IP reputation feeds have important limitations in both accuracy and coverage. Many IP addresses host domains that have been registered with malicious intent, but which have not yet been observed on industry blocklists. These IP addresses may represent a risk to the organization, but a blind spot in traditional IP-based defenses or intelligence sources. **Unlike traditional IP reputation lists, the DomainTools IP risk products leverage the fine-grained, predictive assessments of the popular DomainTools Domain Risk Score, for any domains hosted on an IP.** Because the Domain Risk Score reliably predicts how likely a given domain is to be malicious, even before the domain has been weaponized, an aggregate Risk Score of all domains on a given IP address provides a high-confidence view into the risk level of the IP.

IP Hotlist

The **IP Hotlist** is designed to identify the riskiest population of hosting IP addresses. Two main criteria define this list: the average Domain Risk Score of the hosted domains, and the level of traffic the address is receiving, as measured in Internet-wide passive DNS collection. The Hotlist is an ideal database for high-confidence blocklist and detection rule creation.

IP Hotlist Specifications

Updated: Daily

Inclusion Thresholds: More than 50% of domains on the IP have proximity score of 70+ or Threat Profile score of 90+; pDNS activity on malicious domains within 24 hours

File format: gzip-compressed tab separated text file

Data format: one IPv4 per line with percentages of phishing, malware, and spam metrics for domains hosted in the IP

Feed size (typical): 40-50,000 IP addresses, ~1MB compressed

Field Category	Use Case
Threat Type	Understand the risk category of domains on the IP
ISP and Geolocation	Confirm geographical attributes and ownership
Domain Stats	Measure the IP's reach and scale
Confirmed Threats	Obtain more granular details on "convicted" domains on the IP
Predicted Threats	Threat predictions for domains not yet found on industry blocklists
pDNS and Zerolisted Metrics	Avoid false positives with allow-listed domains; scope traffic activity as recorded by worldwide passive DNS sensors

Hosting IP Risk Feed

The **Hosting IP Risk Feed** is a daily feed of all IP addresses found to be hosting at least one domain. As with the Hotlist, a risk score is given to the IP address based on the population of domains it hosts. Unlike the Hotlist, however, this feed includes *any* actively-hosting IP, regardless of its risk level, and the **Hosting IP Risk Feed also contains detailed data fields enriching the IP**. This makes it ideal for users who wish to apply their own criteria to evaluate IP addresses for risk or characterize them for other purposes.

Hosting IP Risk Feed Specifications

Updated: Daily

Inclusion Threshold: IP is actively hosting one or more domains (regardless of risk level)

File format: gzip-compressed tab separated text file

Data format: one IPv4 per line with the fields given below

Feed size (typical): 15-20 million IP addresses, ~200MB compressed

Field Category	Use Case
Threat Type	Understand the risk categories of domains on the IP
ISP & Geolocation	Confirm geographical attributes and ownership
Domain Stats	Measure the IP's reach and scale
Confirmed Threats	Obtain more granular details on convicted domains in the IP
Predicted Threats	Threat predictions for domains not yet found on industry blocklists
pDNS and Allow-listed Domain Metrics	Scope traffic activity; avoid false positives

Domain and IP Whois Feeds

Increasingly, security teams are enriching the domain names and IP addresses found in log and event sources with relevant metadata to help identify threats to the organization. At smaller scales, this enrichment can be done via API calls, but at large scale, it can be more efficient to query an on-premises file. The feeds listed below provide a variety of data types, updated at intervals ranging from 5 minutes to 24 hours, to enable various enrichment, alerting, and investigative use cases.

5-Minute Whois Feeds

For the most up-to-date data on the newest or most recently-changed infrastructure on the Internet, DomainTools provides domain and IP Whois data, updated on a five minute interval. These feeds are useful for enrichment of infrastructure being rapidly provisioned and "burned" by fast-moving threat actors. Because the data in these feeds comes solely from Whois records, they do not include Risk Scores, nor any of the hosting or content data that are found in the Iris investigation platform and APIs. They do, however, give the analyst the best method of gathering at least preliminary information about the very newest domains in existence.

5-Minute Whois Feed Specifications (Domain and IP Whois)

Updated: Every 5 minutes

Inclusion Threshold: all domains or IPs processed since the previous update

File format: gzip-compressed tab separated or JSON text file

Data format:

- **Parsed** Domain Whois feed: one domain per line followed by each of the Whois data fields given below
- **Parsed** IP Whois feed: one ASN per entry followed by each of the IP Whois data fields given below
- **Raw** Whois feeds: unparsed Whois records *(**NOTE**: subscription to the parsed Whois feeds also provides raw records. The raw-records-only feeds are not recommended for most use cases; please contact DomainTools for inquiries)*

Parsed Domain Whois Feed size (typical): ~18,000 domains

.tsv ~8MB compressed

.json ~10MB compressed

Raw (unparsed) Domain Whois Feed size (typical): ~18,000 domains

.gz file ~7MB compressed

Parsed IP Whois Feed size (typical): ~100-250 ASNs

.json ~70-150KB compressed

Raw (unparsed) IP Whois Feed size (typical): ~18,000 domains

.gz file ~40-75KB compressed

Files retained minimum 7 days

Parsed Whois Data Fields

Domain Whois
Domain name
Parse success (y/n)
Server (Whois)
Lookup Date
Lookup Time
Create Date
Updated Date
Expires Date
Registrar Name

Registrar Abuse Contact: Phone
Registrar Abuse Contact: Email
Registrar IANA ID
Registrar URL
Registrar Whois Server
Admin Name
Admin Org
Admin Street
Admin City
Admin State/Province
Admin Postal Code
Admin Country
Admin Phone
Admin Fax
Admin Email
Billing Name
Billing Org
Billing Street
Billing City
Billing State/Province
Billing Postal Code
Billing Country

Billing Phone
Billing Fax
Billing Email
Registrant Name
Registrant Org
Registrant Street
Registrant City
Registrant State/Province
Registrant Postal Code
Registrant Country
Registrant Phone
Registrant Fax
Registrant Email
Technical Name
Technical Org
Technical Street
Technical City
Technical State/Province
Technical Postal Code
Technical Country
Technical Phone
Technical Fax

Technical Email
Name Server
Registrar Status
Raw Whois Data Blob

IP Whois
RIR queried
Net Range
CIDR
Net Name
Net Handle
Parent
Net Type
Origin AS
Organization
RegDate
Updated
Org Name
Org ID
City
State
Postal Code

Country
RegDate
Updated
Ref
Referral Server
OrgAbuseHandle
OrgAbuseName
OrgAbusePhone
OrgAbuseEmail
OrgAbuseRef
OrgTechHandle
OrgTechName
OrgTechPhone
OrgTechEmail
OrgTechRef
OrgNOCHandle
OrgNOCName
OrgNOCPhone
OrgNOCEmail
OrgNOCRef
Comments
Raw IP Whois data blob

About Domain Risk Score

The Risk and Hotlist feeds are informed by the risk perspective of the DomainTools Risk Score. **Unlike traditional IP reputation lists, the DomainTools Hotlists and Risk Feeds reflect the fine-grained, predictive assessments of the DomainTools Domain Risk Score**. The Domain Risk Score reliably predicts how likely a given domain is to be malicious, even before the domain has been weaponized, based on characteristics the domain carries from its inception. Machine learning classifiers score each domain for phishing, malware, and spam profiles, while the Proximity score indicates how closely connected a given domain is to other domains already proven malicious.