DOMAINTOOLS®

HOW WHOIS DATA ENSURES A SAFE AND SECURE INTERNET

A DETAILED LOOK AT HOW PUBLIC DOMAIN OWNERSHIP DATA DRIVES THREE ESSENTIAL CYBERSECURITY WORKFLOWS

INTRODUCTION

Each year, millions of individuals, businesses, organizations and governments register domain names in the publicly accessible, global domain name system (DNS). ICANN, the organization responsible for the security and stability of the Internet, enforces a level of trust and transparency by requiring these entities provide contact information in exchange for listing in the domain name system. That data is then made publicly available through a network protocol known as Whois which functions as a kind of domain name White Pages that anyone on the Internet can access if they have problems with a domain or the services it is hosting.

At least, that was the expectation three decades ago when the Whois and DNS protocols were conceived. Today, most Internet users are unaware that Whois data is available as a way to protect themselves from malware, fraud, ransomware and other types of nefarious online activities. They are also unaware that security professionals and cybercrime investigators consider information on domain name registrants vital to their daily efforts to keep Internet users safe and their organizations secure.

"Microsoft includes Whois data with our security intelligence insights to provide additional context in investigations and threat detections. This context helps us more quickly triage security issues and implement protections for Microsoft and our customers."

– Jeremy Dallman, Cloud + Enterprise Threat Protection, Microsoft

This paper aims to educate Internet users, policy makers and the ICANN community on the consumer protection and network defense use cases for publicly available Whois data. It strives to close a knowledge gap on why registrant data matters to the security and stability of the Internet. Ongoing discussions on the future of publicly available Whois data make this topic timely and of general interest, especially in the context of the forthcoming European General Data Protection Regulation (GDPR).

DOMAIN RISK ASSESSMENT AND MITIGATION

What is it?

Risk assessment and mitigation is the ongoing everyday duty of the systems and people tasked with network defense and consumer protection. Since domain names are so fundamental to the operation of the Internet, they factor in nearly every attack, and teams in security operation centers must closely analyze domain name registration data to know if an alert represents a credible threat. Applied at scale, risk assessment of domain names using registrant data can identify a signal from noise and bring to the surface otherwise unknown attacks early in their lifecycle.

In 2016, Facebook noticed two websites that were active on the Facebook network, com-video[.]net and account-login[.]net, that were surreptitiously registered with Facebook's corporate contact information. By monitoring Whois for actors trying to impersonate Facebook, and joining this alert to domains active in the Facebook ecosystem, the Facebook security team was able to quickly elevate the threat level of these domains and drive a heightened remediation and takedown process.

Why does it matter?

Most Internet users expect their networks to permit uninhibited communication with every domain on the Internet, be it inbound email or outbound web traffic, and they tolerate blocking only reluctantly and in limited scenarios – on their work computer, for example. Though consistent with the ideals of the Internet and the founding principles of the domain name system itself, this stands in stark contrast with most corporate firewall and physical security policies that adopt a "default deny" policy, where unknown traffic or persons are not permitted until proven trustworthy. Uninhibited communication creates scenarios where network defenders can assess the risk of a domain name only after it has been active on a network, when emails from the domain have already appeared in inboxes, attachments opened, and links clicked.

This reactive approach, coupled with vast amounts of network traffic and limited security staff, makes it essential that teams deliver informed analytics on a potential threat as quickly and accurately as possible. Understanding where a security team needs to focus resources, and knowing which alerts should be escalated, can make the difference between stopping an intrusion early and permitting a full-scale breach. The key factor in many cases can be found in Whois data.

How does Whois help?

Whois data provides information on the ownership of a domain name, including where it was registered, by whom, and for how long. These factors enable rapid risk assessment by human analysts and drive detection models in security alert systems. For example, an analyst can quickly check the Whois record on the source domain for a suspicious email that claims to be "from the CEO" to learn whether the registration patterns are consistent with corporate policies.

Scaling an assessment and alerting process to thousands of indicators per hour simply cannot be handled manually, so the next step is to apply these proven principles at scale. Top security operation centers in the US defense industrial base and financial sector were some of the first to enrich domain names on their network with Whois data, because it helped them detect new threats early and then efficiently hunt for previously unknown compromises. Today, organizations around the world, including leading cyber security companies and global security operation centers are building sophisticated technologies and effective machine-learning models that use Whois data to achieve similar outcomes.

"Like other companies, Facebook uses Whois data in conjunction with our security technology and systems to help protect people from a range of abuse, spam, and other risks. For example, we have used Whois data and related DNS infrastructure to identify and take down tech support scams operated by spammers who make fraudulent use of domain names, phone numbers, and websites."

- Denise Michel, Domain Name System Strategy & Management, Facebook, Inc.

ENUMERATION AND CORRELATION

What is it?

Most Internet users are unaware that the domain name system lacks protections to block known bad actors from registering additional domains to further their campaign. Cybercrime investigators and security analysts must have other ways to link new domain registrations to established intelligence and known indicators. Their efforts can be described as ***enumeration*** and ***correlation***.

Security researchers at ThreatConnect recently published an analysis of threat actor infrastructure targeting the Olympic Games in South Korea, likely in retaliation for the banning of Russian athletes for doping violations. Correlating Whois data with other data sets, ThreatConnect is able to tie the actors to previous efforts in 2016 targeting anti-doping agencies, and enumerate other domains being used to target both the International Olympic Committee as well as numerous anti-doping agencies worldwide.

Enumeration is the act of pivoting from a single domain or identity record through related indicators of compromise (IOCs) to map the full infrastructure controlled by a domain registrant. ***Correlation*** means associating those indicators with other data sources, be it external intelligence from reports on established cyberattack campaigns, indicator lists from trusted peers, or attributes seen in previous attacks against the organization itself. Simply put, a single point of Whois data can often become a breadcrumb leading investigators down a path of discovery to an entire set of malicious content and criminal activities.

Why does it matter?

Legitimate entities and malicious actors alike rarely control only one domain name, and yet investigators often lack that vital context when they see only a single domain in an alert console or an abuse complaint. Risk assessment in isolation can only take them so far; they must surface the full list of domains associated with the entity, and then perform risk assessment on the entire set. For example, knowing a malicious actor has targeted a phishing campaign at other organizations in the same industry reveals sophistication and directs the analyst to elevate the threat level.

Attackers and criminals are free to register batches of domains each day, deploying them incrementally or stockpiling them until they are needed. Enumeration enables network defenders to proactively block future attacks by revealing these "standby" domains. It also helps investigators build cases and prioritize resources by revealing the intent and the scope of the criminal activity.

Enumeration is key to helping organizations maximize value from their investments in threat intelligence. Intel vendors and sharing organizations routinely deliver their intelligence with lists of domains names, IP addresses, and other IOCs. If those IOCs are observed on a network, the entire set of context provided by that intelligence source can be applied to an investigation.

The challenge comes when attackers set up unique infrastructure dedicated to a specific campaign or aimed at a certain organization, which they nearly always do. That's when ***correlation*** becomes essential: if the entire set of an actor's domains, identities, and web hosting characteristics can be enumerated and analyzed, there is a strong likelihood some of those IOCs will match those provided from intelligence sources.

How can Whois data help?

A Whois record for a domain or IP address provides pivot points for enumeration and correlation. The more data available, the more likely an analyst or system will get a match, which speeds their ability to block access with confidence and perform a precise, efficient, and comprehensive investigation.

Enumeration of domains with Whois data at Internet scale requires gathering, indexing, and presenting Whois records from thousands of disparate sources. This is because no central repository exists for registration data of domain names, despite how essential such information is to network defense and consumer protection. Organizations and criminal investigators must initiate programs and dedicate resources to gather Whois data for domain names on their own, or elect an industry partner with proven experience making Whois data accessible for these use cases.

"Whois information is invaluable in minimizing the damage a compromised domain can do to its owner, those who rely on it, and Internet users in general. With the information provided in Whois, an incident response technician can quickly contact responsible parties. This is a standard practice that happens tens of thousands of times daily around the world. From a single network operator working to defend their own network to incident response teams who work at scale, Internet defenders rely upon the Whois data for a domain as the first and often most useful tool to quickly abate attacks, protect their own customers and Internet users in general, and also protect the interests of the domain owner and users."

– Rod Rasmussen, Former VP of Cybersecurity, Infoblox

ATTRIBUTION AND REMEDIATION

What is it?

Attribution is the process of positively identifying the individual or organization behind an attack. ***Remediation***, in the context of domain names, aims to remove malicious content from the Internet and stop criminal behavior, either with legal takedowns or notifying domain owners of compromised sites.

The US Retailer Target suffered a major breach in 2013 that resulted in the disclosure of millions of credit card numbers, many of which were listed for sale on underground forums. Brian Krebs, a security blogger who first reported on the Target breach, tried to unmask the operator of one of those forums, but got stuck at an alias. Using historical Whois records, Krebs was able to uncover a name, Andrey Hodirevski, and an address in southwestern Ukraine.

Why does it matter?

Network defenders may question the value of attribution and seldom seek remediation, as the effort and cost far outweigh the benefits. For example, a domain that mimics a popular consumer products login page in a credential harvesting attack can be registered for less than $10, but it could cost thousands in legal fees to get it taken down. Still, top teams know that understanding who you are dealing with can inform the tactics, techniques and procedures used in the attack and help prioritize incident response resources.

The situation is quite different for law enforcement, brand owners, governments and top security researchers. These entities pursue attribution of attacks to the individual, actor group, and nation-state levels, challenging the perception of impunity held by many online criminal actors and raising awareness of global espionage campaigns. They also architect large-scale takedowns of botnets and criminal infrastructure with global arrest warrants and coordinated network actions.

How does Whois help?

Whois data is, by its very nature, focused on who owns and controls resources on the Internet, and therefore its orientation towards attribution is obvious. An accurate Whois record, publicly accessible, in the clear, can tell anyone the name, email address, phone number, and physical address of the person or organization behind a suspect domain. Even simpler data points such as the country of the registrant or the country in which the name server is registered can give initial indications of attribution and target takedown requests with precision.

It is common, especially in more sophisticated campaigns, for attackers to attempt to mask their identities. It is in precisely these situations where Whois data can be especially powerful. Even privacy-protected domains can reveal data in Whois or DNS records that can be used to pivot and enumerate attacker infrastructure. Thorough evaluation of Whois data across the full map of weaponized or dormant attacker DNS infrastructure can yield critical clues to the identity of the perpetrator. Investigators have cracked many cybercrime cases when they found an email address in a single historic Whois record on one domain they linked to the attack after performing Whois-based enumeration.

Like real estate, domain names have ownership histories, and understanding that chain of ownership can reveal identities or guide legal action toward registrars and web hosting companies that supported the domain earlier in its lifecycle. As with enumeration and correlation, the right solution provider becomes critical here, because registrars and registries are not required to provide archived Whois data. Historical data is particularly useful in pursuit of attribution.

"Financial institutions are among the largest and most common targets for cyberattacks. The Cyber Defense Alliance works collaboratively with many UK banking firms to share indicators of compromise and turn that information into actionable threat intelligence and coordinated incident response. The CDA often works closely with law enforcement in our security operations and domain name Whois data is critical to our daily work."

- Maria Vello, Chief Executive Officer, Cyber Defense Alliance

KEY TAKEAWAYS

- Whois data informs risk assessment of domain names, both by humans investigating suspect activity and by systems trying to detect it.
- Investigators need access to a complete list of a bad actor's domain names and related infrastructure. Enumeration gives them this list, but it depends on key fields in Whois data to be effective.
- Correlation among threat actor domains, network traffic, and external threat intelligence yields powerful outcomes, especially when registrant data is used to link records together.
- Investigators and network defenders need clear, open pathways to attribution of threat actors and remediation of problematic domains. Publicly-available Whois data has enabled these outcomes in support of an open Internet for decades, and it will continue to do so as long as it remains accessible.

REFERENCES

ThreatConnect: https://threatconnect.com/blog/duping-doping-domains/

Krebs on Security: https://krebsonsecurity.com/2013/12/whos-selling-credit-cards-from-target/