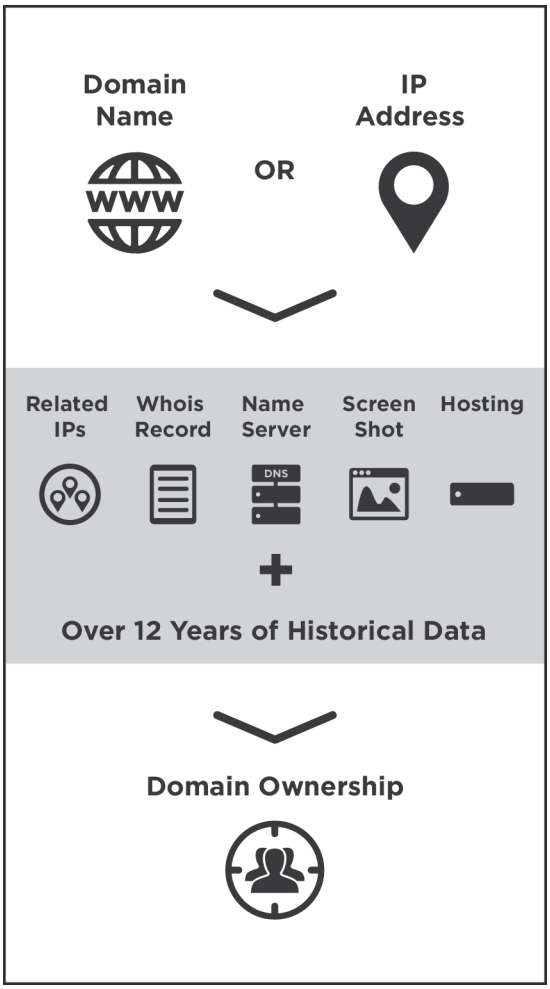
DOMAINTOOLS®

Beyond Whois

How ‘Extended Domain Profiles’ Can Yield Unexpected Insights

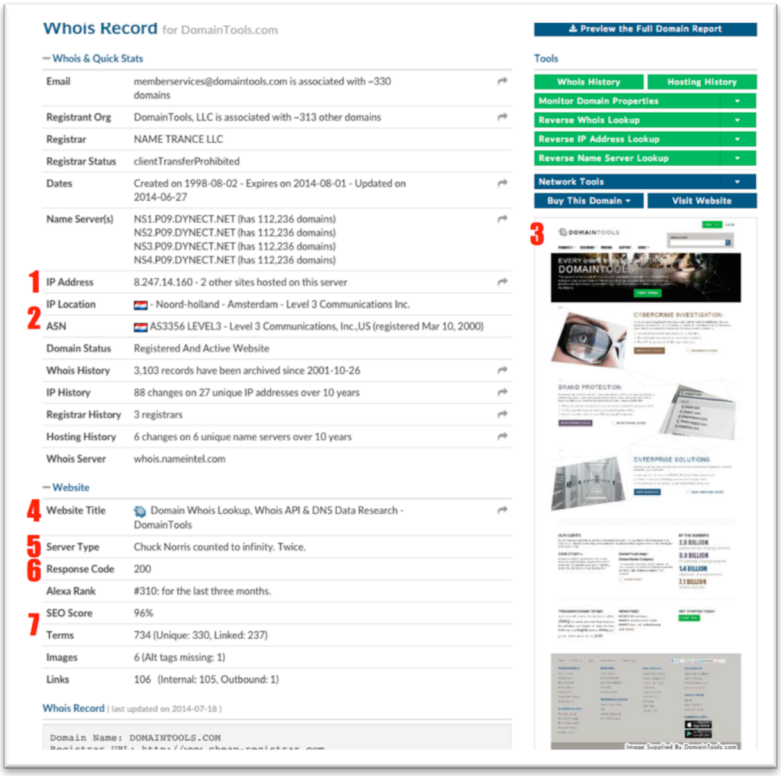
Like a marathon, every online investigation has a starting point. For the investigator tasked with researching a cybercrime or a security analyst assessing the risk level of traffic to or from an unknown domain, the starting point typically begins with a simple Whois search. This basic query takes just a few seconds but can quickly provide the foundation for a successful investigation, answering a number of essential questions such as: ‘is this domain legitimate?’ and ‘is the domain owned by a named individual or is it protected by Whois privacy?’

However, it’s often the data that does not reside in a Whois record that can be truly illuminating. For sake of an analogy, when buying a used car, a prospective buyer is often advised to secure a CarFax record which outlines a wealth of detailed data points related to the car that will ultimately arm the buyer with valuable information beyond just the name on the vehicle title. Likewise, there is a wealth of knowledge captured in current and historical domain registration records and other public data that can be invaluable to an investigator when leveraged appropriately.



Defining a Domain Profile

A Domain Profile represents an amalgam of disparate information and data points that relate to a single domain. Think of it as the difference between a sketch and a photograph: both convey the same information but a photograph offers it in far greater detail. A Domain Profile is a composite of data culled from Whois registration data and supplemented with other domain-related information. The data points that are used to build a Domain Profile include the following components:



Anatomy of a Domain Profile in DomainTools' Whois Result

1. **IP address:** some domains may not have one associated with them, but most do, even if it's just a parking site from the registrar.
2. **IP geo-location and Autonomous System Number (ASN):** this address provides second-level information about the network on which the domain resides.
3. **Screenshot:** current and historical screenshots of the site can show how a site was used over time and/or show similarity to other suspect sites.
4. **Website title:** code on the site that describes what the site is basically "about."
5. **Server type:** the server version the site is running may provide insight into the person(s) running the domain.
6. **Response code:** the code the web server sends back upon the initial HTTP connection–assuming the domain is attached to an operating website.
7. **SEO score, terms, GA codes, images and links information:** these provide insights into whether the domain owner has put any effort into making the domain visible and widely accessible.
8. **MX records:** these are not listed on the Whois results page, but can be found by conducting a **Reverse MX** search. An MX record is a type of resource record in the Domain Name System that specifies a mail server responsible for accepting email messages on behalf of a recipient's domain.

Each of these components represents a valuable slice of information in its own right. Together, the individual slices can be assembled into a body of knowledge by which an investigator or analyst can determine the overall status of a domain. This will enable them to answer fundamental questions from the outset of an investigation, such as:

- Is the domain associated with a working website? Does the site look like it's been professionally designed? Does it appear to have been updated recently?
- Does the domain reside on a dedicated IP address? Is the domain hosted on a "small" IP address (one with only a few domains on it), or is it hosted on a larger hosting site?
- Is there evidence that the domain's owner has tried to maximize the domain's presence, via SEO and other optimization techniques?
- How does the owner/webmaster describe the web site? What are they trying to tell the world (and search engine bots) about the site?

Using IP Addresses for Attribution and Enumeration

Of the domain elements that comprise the Domain Profile, the IP address is often one of the most reliable sources for attribution and enumeration, both of which are key methodologies for every investigator. Attribution refers to the process of naming the initiator of an activity, especially in cases where the person or organization being investigated does not want to be identified. A common example is finding the owner of a malicious domain. Enumeration, which is also known as "Forensic Domain Mapping," refers to the process of building a picture of the domains connected to a common data point or set of data points—often for the purpose of understanding the holdings of a particular person or organization.

Many different investigations boil down to one or both of these, and often one serves the other. For example, by enumerating the holdings of a digital 'John Doe', the investigator may discover that one of the associated domains has details leading to attribution that were not obvious in the initial lookup.

Take the scenario of an investigator trying to figure out the identity of a domain owner who used Whois privacy to mask their identity. If the website in question doesn't reveal any identifying information (i.e. no "About Us" or "Contact Us"), an investigator would use the Domain Profile to continue efforts to find identifying traits. Assuming that neither the Whois record nor Whois History yielded any useful information, the investigator could then take a closer look at IP addresses. On its own, an IP address probably won't reveal all that much. However, as part of a Domain Profile, using an IP address can help to:

- **Understand how many sites are hosted on a specific IP address** (via **Reverse IP**). Why is this important? If the IP address hosts thousands of domains, this will likely not be a productive path to follow. However, if the IP address in question hosts only a handful of sites, this increases the odds that there's some kind of connection between those sites (since a large hosting provider would be unlikely to allocate an address to a random handful of domains). By understanding the connection between sites hosted on the same IP address, one could then associate them via linking characteristics to the original target domain and then run a Whois query from one of these co-hosted domains to potentially identify the true domain owner.

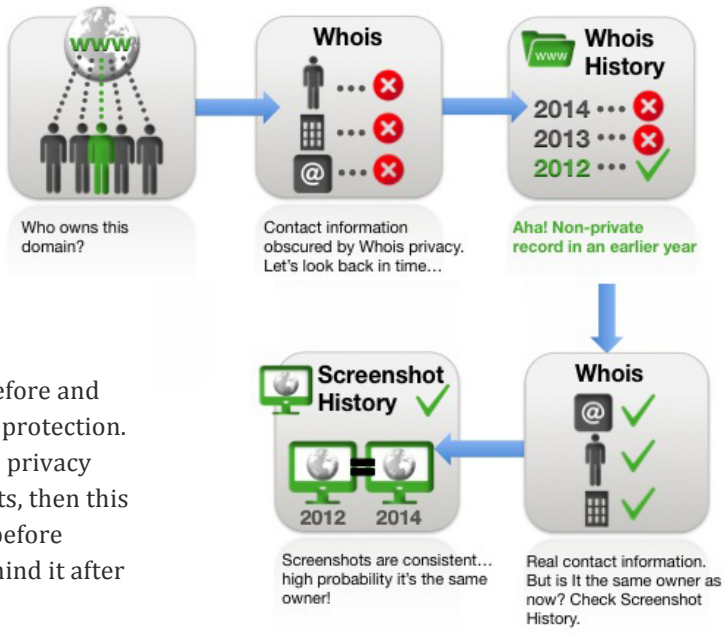
- **Isolate a domain to a specific geographic region.** For example, if a site that purports to sell official Seattle Seahawks merchandise is hosted in Uzbekistan yet represents itself to be "direct from Seattle," this can help characterize the domain—as very suspicious, in this hypothetical example. If there are several domains that could be connected, investigators can look for a related pattern where the locus of business and the locus of hosting don't make any rational sense. (*Note that the IP location often doesn't match a physical location, and in many cases, this is normal and innocuous. Your own intuition and judgment will be helpful in this regard. For instance, a Seahawks merchandise vendor with an IP address in Virginia is not nearly as suspicious from an investigative standpoint as one in Uzbekistan.*)

- **Identify the owner of a specific IP address.** The **IP Whois lookup** can be used to determine whether an IP address is owned by a reputable, big hosting provider or ISP, or a smaller one. So-called "bulletproof hosting" sites are of particular interest to cybercrime investigators, because these providers aim to shield their customers from prosecution, sinkholing and other law enforcement takedown activities. Many of the world's most prolific spammers and phishers use bulletproof hosting to mask their identity.

DomainTools' **screenshots** can be useful in documenting similarities between sites when using Reverse IP to map related sites hosted on the same IP address.

Looking Back to Look Forward

Data from history tools provide invaluable insight into how a particular domain has evolved over time, based on a combination of canonical Whois data points (name server and registrar) found in **Whois History** as well as DomainTools' proprietary IP address data (**Hosting History** and **Screenshot History** are other useful historical tools that can provide additional insights).



Screenshot History is especially useful for attribution, as it can be used to determine ownership before and after the threshold of when a domain went into privacy protection. If the "before" screenshots—those taken before domain privacy went into effect—closely matches the "after" screenshots, then this raises confidence that the visible owner of the domain before privacy is likely the same individual or organization behind it after privacy was enabled.

Conclusion

Whois records are an important jumping off point for cyber investigators. However, the sophisticated investigator will quickly recognize that while Whois represents a natural starting point, it is usually not enough. By using only Whois to conduct a digital forensics investigation, you're missing out on a wealth of other information that can be crucial to finding important answers. The datapoints that compose a Domain Profile go far beyond Whois, and have proven to be a critical asset in tracking down the owners of suspicious domains.

About DomainTools

DomainTools cyber threat intelligence solutions give organizations the ability to create a forensic map of criminal activity, assess threats and prevent future attacks. Customers and partners can also integrate this rich dataset into their existing solutions through an API. Fortune 500 companies, global government agencies, and many security and online fraud investigation vendors use the DomainTools platform as a critical ingredient in their threat investigation and mitigation work.