

[Trust and verification in an age of misinformation Module 3.3](#)

Okay time for lesson number three here in week number three, hope you're finding everything useful. We're gonna dive in a little bit on investigating websites. So, there's a couple of basic things that you want to do when you're interested in a website for a story and you want to learn more about it, um one is you want to try and figure out any ownership information about it, and the second is you want to see if it's connected to other websites through its Google Analytics code or Google AdSense ID, I'll explain a little bit more about that in a second.

Let's talk about finding a website owner, um so this is a really great website it's free there are some better paid tools, but I want to focus on free tools in this. It's called domain big data, so that's domain big data.com, again free tool. So, there's a search out there called a Whois search, anytime a website domain name, so domain big data.com for example is registered, um it's recorded in a universal database out there. So, at any given time you can do a who is search and at the very least you're gonna be able to find out when a domain name was was created and also potentially when it expires.

Now if you're really lucky there also may be public information about the person who owns it contact information for them, that's if you're really lucky, but people are able to pay to keep their owner identification information private. So, sometimes you're gonna run a Whois search and you're gonna show you're gonna come with information that says like privacy protect LLC, and that's going to tell you that somebody has paid the extra you know five or ten dollars to get that privacy protection.

All right, so that's our overview we're gonna look in a good scenario here where we actually do get the information about an owner. So let's run our search, we run a search here in domain big data. This is the top of the results page here now I just decided to do a search for Newsweek.com, here's what we get in the upper section you got you can see I have the the arrow there it's giving us the date creation of the domain name. So, this domain name was created in in 1994 been around for a long time, which you would expect for Newsweek magazine and below that you know it's just sort of calculating automatically the age of it.

So, just like social media accounts that are new are a little more suspicious really new domain names can also tend to be a little more suspicious, especially if they're claiming to be a well-established brand of some sort. So, that's the first thing is we're gonna look at when was it created, second thing that we want to look at here where the arrow is is in the registrant information. So, here is you know the person or entity that has bought this domain name and is currently holding it. In this case you know we're really lucky because we've got you know the organization Newsweek LLC and and domain big data

is great because it's telling us that same organization is associated with 98 total domains, and if you were to click on Newsweek LLC domain big data it's gonna list them for you.

So, then suddenly we can see alright what else does this company own, and below that we also have the email associated with the domain name registration. In this case it's domains at Ivy times which is associated with more than a hundred domain names. So, I'm gonna look at both of those and I'm gonna get a sense of the universe that these companies own, and if I had done more research I would see that Newsweek and IB Times are part of the same company.

So we don't have a case where you know the organization and then the email address are linked to different companies, which would be a confusing thing we'd need to look more into. In this case they really are connected with the same company but we have a different amount of domain names, one with the organization name and two with the email. So, you need to look at both because people often register things under different names or different email addresses and you want to get a whole picture of the universe of all the domain names that they own.

Alright, now the other things I just want to note here is you know you get a phone number, it tells you the city you get an address. This is a really important thing I'm going to emphasize it right here, is the email address. So, when someone buys a domain name, they can put in whatever name they want whatever address they want whatever phone number they want they can put in and make up whatever information they want, so you always have to take this with a grain of salt. But the email, of all of the information that you would find in a public facing domain registration record, the email is often the most interesting. And the reason for that is even if you want to make up a name and fake stuff to throw people off the scent you still in order to manage your domain name kind of need an email address where the company you bought it from can contact you. So, a lot of times you know the email address maybe your best leave, now it's possible someone just created an email address they'll never use again and it's only for that domain name, sometimes it's a dead end, but it's also possible that the that the email address will actually enable you to get in touch with the person who owns the domain name. So, the email is potentially the most useful piece of information in a domain name registration, if you think some of the stuff might be fake. And just to give you a sense of what the typical registration, you know listing looks like this is it. Domain big data is nice because it organizes stuff a little bit nicer, but down below that it gives you you know the basic list. And the thing I want to note here is that I've highlighted the section, which has the registrar and the registrant, but sometimes you'll have a one name in the registrants section and one name in the administrator section, which is the

more technical section. So, sometimes you have a company an entity that owns the domain name, the registrant, and then they have contracted a separate entity to be the administrator of it. So, that's why you might find different information between these two.

Alright, last thing I want to note is a tip about IP address. So, the IP address tells you the location on the internet that the server this website is hosted on is, and sometimes if you if you click on the IP address it's going to tell you that there's you know thousands or tens of thousands of websites hosted on the same ip, that's not really useful to you. However, if you were to click on IP address and you only see like five or ten or 20 or 30 or 40 or 50 sites hosted on the same ip that's really interesting. Those potentially could be connected to the same website, to the same website owner. So, when you get a domain name record and there's all the information is private or it's a dead end look at the IP, because maybe there's only a few other sites on that same

server, and maybe those other sites have public domain registration information that's a good clue. And, you need to buttress that with other things, but if there's only a few sites on the same ip there's a decent chance they may be connected, it's not for sure but they could be connected.

All right, let's move on here I mentioned Google Analytics and Google AdSense. So, Google Analytics is a free program that tons and tons of websites use to track the traffic coming to their website, and Google AdSense is a really lucrative advertising network that a lot of websites use to put ads on their sites and earn money. The fact that these two things are so popular and so ubiquitous on the web means that it's a great tool for you to try and connect websites to each other. So, I'm going to describe the process and then show it to you. The first thing you're going to do is when you come to a website you're going to go up and into the the menu in your browser and you're gonna open up the source code of the web page, I'll show you how to do that. Then you're going to look on that page for the tell-tale signs that they're using Google Analytics and/or Google Ad sets. In this case the letters UA and a dash or the letters pub and a dash and then after those letters would be a unique number and the reason this is interesting is that in order to use game of Google Analytics or Google AdSense you have to be assigned a unique ID that's linked to your account. So, if we find the same unique ID across multiple web sites that's in the extremely strong signal that they are connected together. All right, and then we're gonna use a tool called spy on web where we paste in the analytics or AdSense ID and see which other websites are using it.

One thing to note is that you can actually just type in the domain name in to spy on web but I discourage you from doing that because sometimes the data in spy on web is old and it may pull up an old analytics or AdSense ID that's not actually on the site anymore. All right, so what am I talking about here let's start in your browser you go up to the view

menu then you go to developer and you go to view source on any web page. This will open up a new tab with the source code for that particular webpage, and once you have that web source code open that's what this bottom section is

down here you're then going to go up to the edit menu find, find and then this little box opens up here. And, for example I typed in UA and - to find the Google Analytics code and lo and behold there was one on the page for this particular website, UA - five four - 609 eight 9-7. So, that tells us that this website is using Google Analytics and that's great because then we can go to spy on web and take that number right there search on spy on web, remember it's spy on web com and we get a result.

So, the website I was on was conservative Tribune com, it's telling us the domain name, the IP address, it had it has or had two different Google Analytics IDs on it, which is common to find more than one it gives us a little other information about traffic sources for it, which is a neat little thing to look at like where is the traffic coming for this website, how are they getting an audience. But the thing that we're really interested in is that we can see other domain names that use these exact same Google Analytics

IDs, and if you look at these this is an excellent indication that these are all part of the same network. So, we're able to take disparate websites and suddenly connect them with the same analytics or AdSense ID.

Okay, let me just summarize a few of the key things not only in this section, but the previous two, is well lessons that are really really important for you to remember. The first thing is, I'm showing you a bunch of tools, I'm showing you a bunch of techniques, but it does not replace traditional journalism, it does not replace talking to people interviewing them asking, them questions bringing the information and I have shown you to them to get them to answer for it, okay really important. When we talked about social media accounts, we talked about investigating the network in the history of an account don't just take it for granted look at where it exists within the network, try to find other accounts. As well an overall thing these last two here I just want to emphasize to you that you know when you're on a website when you're looking at a social media account you've really just got to have this mentality of I'm gonna look at everything, I'm gonna click on every link if I hit a roadblock and the Whois information is private I'm gonna look at the IP address, I'm gonna look here, I'm gonna look there, I'm gonna take sections of the text on a website and plug it into Google to see if other websites have the same text you've got to face these roadblocks and these dead ends and say I'm gonna get around it I'm gonna find another way. So, that's the mentality for doing this kind of work. And then the last thing here, is I really want you to write this down and bookmark this HTTP colon slash slash bitly slash verification tools, and tips com, this is

a place where I am constantly gathering and updating tutorials tools the stuff you've seen here plus a whole bunch I didn't get into.

So, this is a free resource for you that is always gonna be there and always updated with stuff I'm finding. So, this is this is me giving you a personalized guide to it, you can ignore my background falling again, it keeps doing that. So, that's that's a quick summary of a few things I wanted to reinforce and again I want to reinforce, that if you take this URL and you bookmark it you're gonna have access to great tools and tutorials that I'm always gathering because I'm obsessed with this stuff.

Alright, so that wraps up our verification piece here but I've got two other things two other quick lessons that I want you to see, so looking forward to seeing you in those.