

# Acquisizione forense di pagine web

Obiettivi, approcci e metodologie

---

ANDREA LAZZAROTTO



ANDREALAZZAROTTO.COM

# A proposito di me

- Consulente informatico forense e sviluppatore
- Gli interessi includono la WhatsApp forensics e anti-forensics (manipolazione delle chat) e l'acquisizione e analisi dei metadati dei profili Instagram
- Autore di alcuni strumenti open source, inclusi **RecuperaBit** per la ricostruzione di NTFS e **Carbon14** per datare le pagine web (entrambi si trovano in CAINE)



# Agenda

## INTRODUZIONE E OBIETTIVI

Cos'è un'acquisizione web forense e perché ci interessa

## APPROCCI E STRUMENTI

I software principali e gli elementi da acquisire

## METODOLOGIE OPERATIVE

Una procedura da seguire e alcuni punti critici

# Introduzione e obiettivi

COS'È UN'ACQUISIZIONE WEB FORENSE E PERCHÉ CI INTERESSA



# Definizione

Vogliamo ottenere l'acquisizione forense di una pagina web, vale a dire la copia "cristallizzata" di un contenuto presente online, seguendo modalità forensi.

“Forense” significa innanzitutto che l’operazione deve essere svolta **seguendo criteri scientifici e con procedure verificabili** ma anche (idealmente) ripetibili. Il contenuto dovrebbe essere il più possibile completo e conforme all’originale.

Questo viene fatto spesso per utilizzare il risultato a scopo di tutela legale o in ambito giudiziario.



## ALCUNI AMBITI DI APPLICAZIONE



### Proprietà intellettuale

Documentare la violazione di brevetti, marchi o copyright, o dimostrare l'anteriorità



### Diffamazione online

Tutelarsi in seguito alla diffusione di insulti o contenuti lesivi sui social network



### Bullismo

Acquisire prove (anche video) su atti di bullismo avvenuti o diffusi in forma digitale

# PROPRIETÀ INTELLETTUALE

≡ ELLE ENTERTAINMENT MODA BEAUTY VIVERE MEGLIO ATTUALITÀ

ABBONATI NEWSLETTER IT

1 LE BORSE PIÙ BELLE DELLE SFILATE AI 2021-2022  
2 MICHELLE PFEIFFER BELLISSIMA SENZA TRUCCO  
3 LA MODA CHE VERRÀ: I VESTITI AUTUNNO-INVERNO 21-22  
4 I CAPPOTTI PIÙ BELLI DALLE SFILATE AI 2021-2022  
5 70 ANNI DI MAX MARA CON I CAPPOTTI AI 2021-2022

## Gucci e Facebook fanno causa congiunta contro la contraffazione di borse e capi in vendita sui social

—*La maison fiorentina e il colosso social hanno un "solido piano strategico" per difendere la proprietà intellettuale*

ELLE DI FEDERICA CAIAZZO 27/04/2021

Non è un gioco, ma un business molto sporco di certo sì. Dopo essere sceso in campo con Kering per i diritti delle donne, il marchio Gucci torna a manifestare il proprio impegno a sostegno della proprietà intellettuale. Stavolta però lo fa insieme al più grande colosso dei social media, Facebook. Missione: combattere la contraffazione online, dopo l'ultimo (eclatante) caso di vendita di capi Gucci contraffatti tramite le piattaforme Instagram e Facebook.

## Gucci e Facebook contro la contraffazione: l'accusa

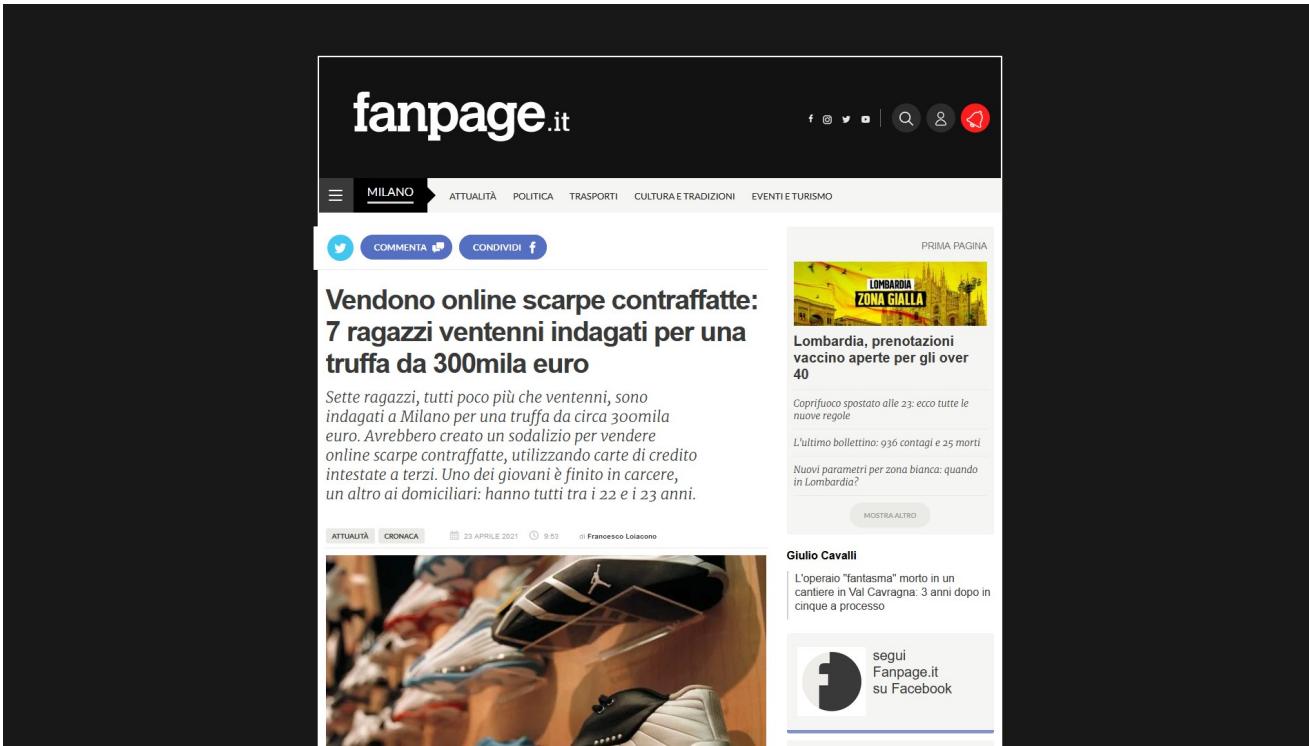
### + RELATED STORY



Gucci apre la terza Osteria da Massimo Bottura

Oltre 150 profili Instagram e 5 account per riuscire nell'intento di vendere falsi "firmati" Gucci. Finte scarpe Gucci, borse Gucci false and so on... A riportare la notizia in

# PROPRIETÀ INTELLETTUALE



# DIFFAMAZIONE ONLINE

**SiracusaNews**  
mercoledì 19 Maggio - 2021 Aggiornato alle 17:24

Cronaca	Politica	Sport	Attualità	Cultura	Editoriale	Acchiappavip	Pubblired	Video
Home	Edizioni Locali ▾	Diventa Reporter	Necrologi	Offerte Lavoro	Social ▾	Contattaci	Pubblicità	Cerca
12 MAGGIO 2021 CRONACA   LENTINI								
0 1 MINUTO DI LETTURA								
<p>UN 30 ENNE</p> <p><b>f o</b> Insulta gli agenti di Polizia su Facebook, <b>✉ t</b> identificato e denunciato</p> <p><b>✉</b> L'uomo è stato denunciato per i reati di vilipendio delle istituzioni e diffamazione aggravata a mezzo social</p> 								
<p>Agenti del Commissariato di Pubblica Sicurezza di Lentini hanno denunciato un uomo di 30 anni, lentinese, per i reati di vilipendio delle istituzioni e diffamazione aggravata a mezzo social.</p> <p>A seguito di un'operazione di Polizia, il denunciato, non erarendo l'intervento delle Forze</p>								

# DIFFAMAZIONE ONLINE

Mercoledì, 19 Maggio 2021 ☀ Sereno o poco nuvoloso

citynews | ☀ 🔍 ⌂ Accedi

## ≡ TRIESTEPRIMA

**Redazione**  
17 dicembre 2020 09:29

[f](#) [t](#) [g](#)

Si parla di  
**difamazione online**  
**Polizia Postale**

### CRONACA

#### Diffamazione online: 56enne denunciato dalla Polizia Postale

Il reato prevede la reclusione fino a un anno o la multa fino a oltre mille euro



**TRIESTEPRIMA**

**L**a Polizia Postale e delle Comunicazioni ha deferito nei giorni scorsi all'autorità giudiziaria un triestino del 1964 per il reato di diffamazione. Questo reato prevede la reclusione fino a un anno o la multa fino a 1032 euro per chiunque, in assenza della persona lesa offenda l'altrui reputazione.

© Riproduzione riservata

[f](#) [t](#) [g](#)

**I più letti**

- CRONACA**  
1. Ancora far west con risse e violenze in pieno centro: indagini a tappeto, denunciata una ventenne
- CRONACA**  
2. Coprifumo alle 23, riapertura palestre, matrimoni: il nuovo Decreto riaperture
- CRONACA**  
3. Covid: in Friuli Venezia Giulia zero decessi, non succedeva da ottobre
- CRONACA**  
4. Addio a Alessandro Talotti, stella italiana del salto in alto
- CRONACA**  
5. Seminuda e ubriaca molesta i passanti con una bottiglia, poi crea scompiglio al pronto soccorso: denunciata

# BULLISMO

LA STAMPA

Insultata su Instagram e al telefono perché omosessuale: la denuncia di una diciottenne

La ragazza di Bari ha inutilmente provato a confrontarsi l'uomo che si nascondeva dietro un profilo social femminile



VALERIA D'AUTILIA  
14 Maggio, 2021

BARI. «Pensa che merda avere una figlia come te». Prima i messaggi, poi la telefonata. Insultata perché omosessuale. Una serie di violenze verbali partite da un account Instagram e indirizzate ad una ragazza di Bari. Lei, una diciottenne prossima alla Maturità. Una ragazza semplice, che sta vivendo una storia d'amore con una sua coetanea. E forse proprio le foto della coppia poste sui social negli ultimi giorni possono aver dato fastidio a qualcuno. Improvvisamente, è stata contattata dal profilo di una donna. Minacce e aggressioni via chat. «Fai schifo, scimmia», «Morite, depravati, vi bruceret nei fornii». Poi una serie di telefonate anonime nel cuore della notte. Per giorni. All'inizio non ha mai risposto. «Rispondi se hai il coraggio. Oppure sai essere solo una schifosa depravata lurida?». A quel



# BULLISMO

Mercoledì, 19 Maggio 2021  Nubi sparse e schiarite

citynews | ☀️ 🔍 🔍 |  Accedi

≡ ROMATODAY

  
Veronica Altimari  
videomaker  
12 maggio 2021 17:24  
  

Si parla di  
risse tra ragazzi

**ATTUALITÀ**  
**Dalle minacce sui social alla "resa dei conti": la storia di Alice, picchiata dalla "banda dei diciassette" a soli 11 anni**  
Il racconto del papà che, arrivato sul posto appena in tempo, è riuscito a fermare le violenze. Ora la piccola è seguita da uno psicologo



Sullo stesso argomento  
  
**ATTUALITÀ**  
Appuntamento sui social:

“**V**ieni qui a chiarire”, è la richiesta avanzata per far arrivare la vittima sul posto. Dopo settimane di vessazioni attraverso i social solo perché ha da poco iniziato un’amicizia con un bambino del quartiere. Una volta li vlene presa, trascinata in un punto appartato, ma al centro di una vera e propria arena di spettatori. Accerchiata, spintonata, buttata a terra. Il tutto di fronte ad una folla di coetanei, alcuni poco più grandi di lei, che con lo smartphone in mano riprendeva tutto.

La disavventura di Alice (nome di fantasia) è finita con **17 giorni di prognosi accertati al**

**I più letti**

- ATTUALITÀ**  
1. Vaccini a Roma e nel Lazio: quando, dove, come prenotare. Tutte le informazioni
- ATTUALITÀ**  
2. Coronavirus a Roma e nel Lazio, il bollettino dei nuovi contagi: i dati del 16 maggio 2021
- ESQUILINO**  
3. Atac, torna il controllore a bordo dei mezzi. Lo slogan: “Con 1 euro e 50 paghi meno e stai sereno”
- ATTUALITÀ**  
4. Coronavirus a Roma e nel Lazio, il bollettino dei nuovi contagi: i dati del 17 maggio 2021
- ATTUALITÀ**  
5. Coronavirus a Roma e nel Lazio, il bollettino dei nuovi contagi: i dati del 18 maggio 2021

## SI PARLA ANCHE DI HACKINBO!

The screenshot shows a news article from the website 'Domani'. The header includes a search bar, social media links (Twitter, Instagram, Facebook), and navigation links: Stoglia il giornale, Abbonamenti, Sostieni le inchieste, Newsletter, and Accedi. Below the header, the 'Domani' logo is displayed in red. A horizontal menu bar lists categories: Ambiente, Fatti, Politica, Giustizia, Longform, Economia, Commenti, Voci, Cultura, Tecnologia, Video, Covid-19, Recovery Plan, Areale, Mafie, Merica, Deutsche Vita, In contraddittorio, and L'Arte di Domani. The main content features a red 'TECNOLOGIA' tag above the title: 'Andrea Lazzarotto presenterà ad HackInBo le tecniche per acquisire le prove dal web'. Below the title is the date '20 maggio 2021 • 11:59' and social media sharing icons. The central image is a blurred photograph of a person speaking at a podium, with the HackInBo 2021 logo overlaid.

In caso di diffamazione online o altri contenuti dannosi, è importante raccogliere le prove sul web in modo corretto: il semplice screenshot non è sufficiente

Andrea Lazzarotto, un giovane consulente informatico forense 29enne, sarà relatore ad HackInBo 2021 per presentare un



*L'ultimo che avete visto potrebbe anche  
non essere proprio vero...  
Così come qualsiasi altro screenshot.*

— ANDREA LAZZAROTTO

# La “soluzione” del notaio

Per soddisfare la necessità di “certificare” il contenuto di una pagina web, alcuni pensano che sia una buona idea richiedere a un notaio **il rilascio di una “copia conforme” di una pagina web.**

Tale documento assumerebbe l’efficacia probatoria di cui all’articolo 2712 c.c. (che riguarda le riproduzioni meccaniche, fotografiche o informatiche).

Dal punto di vista tecnico, **il notaio non può certificare cosa è contenuto in una pagina web, ma al limite può certificare ciò che sta vedendo sullo schermo:** due cose del tutto diverse.



# Obiettivi

## ACQUISIRE IL CONTENUTO

Vogliamo documentare e riprodurre un contenuto presente sul web affinché venga introdotto come elemento di prova.

## PREVENIRE LA CANCELLAZIONE

Il contenuto potrebbe venire rimosso da un momento all'altro.

## DATARE UNA PAGINA

La data di pubblicazione può essere determinante.

## ACQUISIRE IL CONTENUTO



Testo visibile

Teoricamente vale anche una stampa, **se si può verificare la corrispondenza con la fonte**



Tutto ciò che non si vede

Il codice della pagina, i metadati e altri elementi risultano di grande importanza

# Prevenire la cancellazione

Quando ci sono situazioni in cui il contenuto potrebbe venire modificato oppure cancellato in futuro (praticamente sempre), lo screenshot da solo **risulta totalmente inadeguato**.

Testi, foto e video illeciti **potrebbero venire cancellati** da chi li ha pubblicati oppure dal gestore della piattaforma.

Pertanto, non è sufficiente una mera riproduzione meccanica di una pagina, ma è necessario collegare il contenuto alla posizione (es. URL) e a una data in cui esso era accessibile.



# Datare una pagina

Nelle attività di OSINT o di tutela della proprietà intellettuale è importante riuscire a stabilire ragionevolmente **quando una pagina è stata pubblicata.**

La maggior parte dei CMS permette di nascondere il valore della “data di pubblicazione”, o ancora peggio **modificarla a piacimento** (per esempio WordPress).

Tuttavia esistono altri elementi statici la cui data di modifica è più difficile da alterare (o impossibile senza accesso FTP).



# Approcci e strumenti

I SOFTWARE PRINCIPALI E GLI ELEMENTI DA ACQUISIRE



# Tipologie di strumenti

## IN LOCALE

Vengono installati sul computer e sono totalmente sotto il controllo di chi effettua l'acquisizione.

## REMOTI

Si basano sui servizi forniti da una terza parte, certamente estranea al procedimento. Il livello di controllo è inferiore.

## PROGRAMMI IN LOCALE



Controllo totale

Generalmente sono flessibili, consentono di usare proxy e accedere a pagine nella LAN



Possibili contestazioni

Qualcuno potrebbe obiettare che l'acquisizione sia potenzialmente modificabile in locale

## STRUMENTI REMOTI



Minore controllo

Non richiedono configurazione ma c'è meno flessibilità, spesso non sono interattivi



Maggiore autorevolezza

Risulta difficile contestarne la credibilità accusandoli di non essere "terzi"

# Elementi registrati

Si può lavorare su diversi livelli:

- Rappresentazione grafica o registrazione video
- Codice sorgente HTML, CSS, JS e immagini
- Richieste HTTP e relativi header
- Traffico di rete (pacchetti)
- Configurazione del sistema sottostante



## RAPPRESENTAZIONE GRAFICA



Screenshot

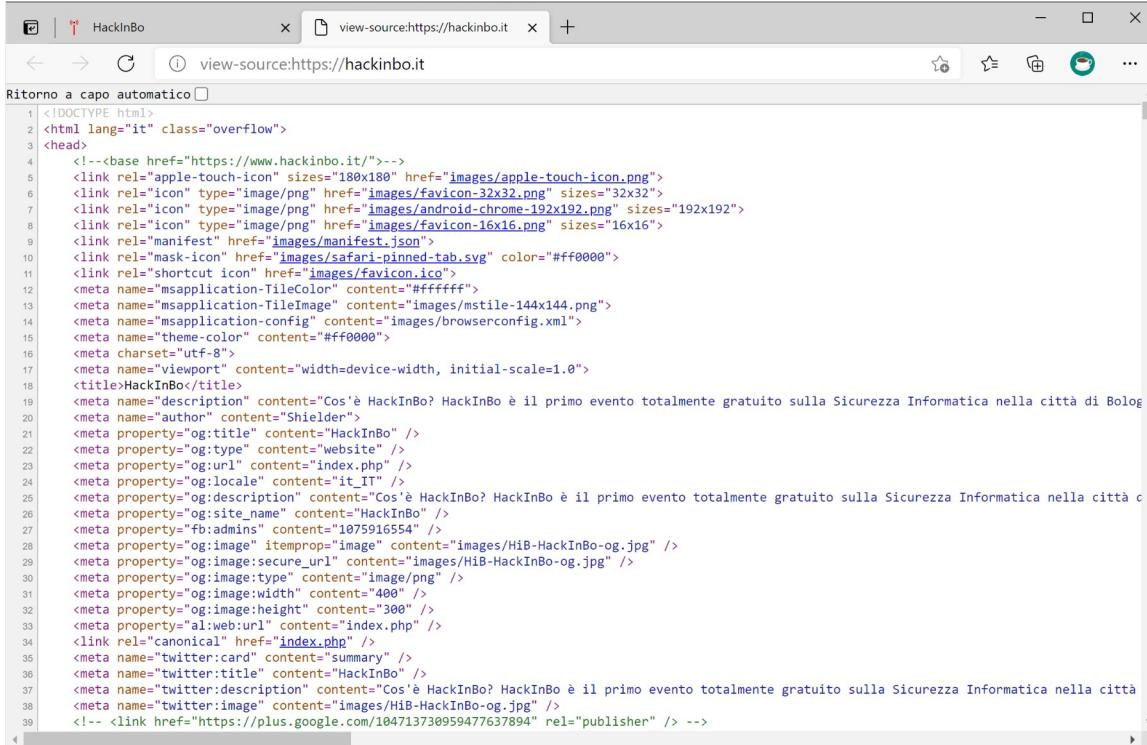
Serve come “colpo d’occhio” ma può essere comodo da inserire in una relazione scritta



Video

Mostra in tempo reale le operazioni effettuate e la navigazione sulla pagina di interesse

## CODICE SORGENTE



The screenshot shows a browser window with the title "HackInBo" and the URL "view-source:https://hackinbo.it". The main content area displays the raw HTML source code of the website. The code includes standard HTML tags like <html>, <head>, and <body>, along with various meta tags for SEO and styling, such as <meta name="description"> and <meta name="viewport">. There are also links to external resources like favicons and touch icons.

```
1 <!DOCTYPE html>
2 <html lang="it" class="overflow">
3 <head>
4   <!--<base href="https://www.hackinbo.it/"-->
5   <link rel="apple-touch-icon" sizes="180x180" href="images/apple-touch-icon.png">
6   <link rel="icon" type="image/png" href="images/favicon-32x32.png" sizes="32x32">
7   <link rel="icon" type="image/png" href="images/android-chrome-192x192.png" sizes="192x192">
8   <link rel="icon" type="image/png" href="images/favicon-16x16.png" sizes="16x16">
9   <link rel="manifest" href="images/manifest.json">
10  <link rel="mask-icon" href="images/safari-pinned-tab.svg" color="#ff0000">
11  <link rel="shortcut icon" href="images/favicon.ico">
12  <meta name="msapplication-TileColor" content="#ffffff">
13  <meta name="msapplication-TileImage" content="images/mstile-144x144.png">
14  <meta name="msapplication-config" content="browserconfig.xml">
15  <meta name="theme-color" content="#ff0000">
16  <meta charset="utf-8">
17  <meta name="viewport" content="width=device-width, initial-scale=1.0">
18  <title>HackInBo</title>
19  <meta name="description" content="Cos'è HackInBo? HackInBo è il primo evento totalmente gratuito sulla Sicurezza Informatica nella città di Bolognina!>
20  <meta name="author" content="Shielder">
21  <meta property="og:title" content="HackInBo" />
22  <meta property="og:type" content="website" />
23  <meta property="og:url" content="index.php" />
24  <meta property="og:locale" content="it_IT" />
25  <meta property="og:description" content="Cos'è HackInBo? HackInBo è il primo evento totalmente gratuito sulla Sicurezza Informatica nella città di Bolognina!>
26  <meta property="og:site_name" content="HackInBo" />
27  <meta property="fb:admins" content="1075916554" />
28  <meta property="og:image" itemprop="image" content="images/HIB-HackInBo-og.jpg" />
29  <meta property="og:image:secure_url" content="images/HIB-HackInBo-og.jpg" />
30  <meta property="og:image:type" content="image/png" />
31  <meta property="og:image:width" content="400" />
32  <meta property="og:image:height" content="300" />
33  <meta property="al:web:url" content="index.php" />
34  <link rel="canonical" href="index.php" />
35  <meta name="twitter:card" content="summary" />
36  <meta name="twitter:title" content="HackInBo" />
37  <meta name="twitter:description" content="Cos'è HackInBo? HackInBo è il primo evento totalmente gratuito sulla Sicurezza Informatica nella città di Bolognina!>
38  <meta name="twitter:image" content="images/HIB-HackInBo-og.jpg" />
39  <!-- <link href="https://plus.google.com/104713730959477637894" rel="publisher" /> -->
```

Le pagine web sono formate principalmente da codice HTML, CSS, JS e immagini.

Senza questo, in sostanza non è un'acquisizione.

Fate attenzione alle pagine dinamiche.

# RICHIESTE HTTP

The screenshot shows the Network tab in Google Chrome DevTools for the URL <https://hackinbo.it/>. The tab bar includes Elementi, Console, Origini, Rete (selected), Prestazioni, Memory, Application, Security, and Lighthouse. The network timeline shows several requests, with a red vertical line indicating the current request for the main page. The left sidebar lists resources under 'Nome': `hackinbo.it`, `styles.css`, `bootstrap.min.css`, `lightslider.min.css`, `lightgallery.min.css`, `lg-transitions.min.css`, `classic-10_7.css`, `main.css`, `android-chrome-192x192.png`, `HIB-logo-no-scritta.png`, `1473873258_1473442507_1473376634_IMG_8345.png`, `1473699752_LK5_6566.jpg`, `1619723831_relatore_antonio_brogi.jpg`, `1619723783_relatore_claudio_canavese.jpg`, `1619723942_relatore_fabio_carletti.jpg`, `1620230256_relatore_andrea_lazzarotto.jpg`, and `1620032915 Massimo Bozza.jpg`. The main panel displays the request details for `hackinbo.it`:

**URL richiesta:** <https://hackinbo.it/>  
**Metodo di richiesta:** GET  
**Codice stato:** 200  
**Indirizzo remoto:** 94.23.64.24:443  
**Criterio referrer:** strict-origin-when-cross-origin

**Intestazioni risposte**

- `access-control-allow-origin: https://hackinbo.it`
- `cache-control: no-store, no-cache, must-revalidate, post-check=0, pre-check=0`
- `content-encoding: gzip`
- `content-length: 11141`
- `content-type: text/html; charset=UTF-8`
- `date: Thu, 20 May 2021 23:04:52 GMT`
- `expires: Thu, 19 Nov 1981 08:52:00 GMT`
- `pragma: no-cache`
- `server: Apache`

115 richieste 727 kB trafeiti 23.8 MB risorse Fine: 5.97 s DOMContentLoaded

# TRAFFICO DI RETE

Cattura da Ethernet0

File Modifica Visualizza Vai Cattura Analizza Statistiche Telefonia Wireless Strumenti Aiuto

Aplica un filtro di visualizzazione ... <Ctrl-/>

No.	Time	Source	Destination	Protocol	Length	Info
2988	63.559652	204.79.197.222	192.168.218.129	TCP	60	443 → 49738 [ACK] Seq=10311 Ack=1892 Win=64240 Len=0
2989	63.578667	204.79.197.222	192.168.218.129	TLSv1.2	244	Application Data, Application Data
2990	63.578870	192.168.218.129	204.79.197.222	TCP	54	49738 → 443 [ACK] Seq=1892 Ack=10501 Win=65535 Len=0
2991	63.826854	192.168.218.129	172.217.19.195	TCP	55	[TCP Keep-Alive] 49718 → 80 [ACK] Seq=409 Ack=703 Win=63538 Len=1
2992	63.827328	172.217.19.195	192.168.218.129	TCP	60	[TCP Keep-Alive ACK] 80 → 49718 [ACK] Seq=703 Ack=410 Win=64240 Len=0
2993	64.020488	VMware_c0:00:08	Broadcast	ARP	60	Who has 192.168.218.2? Tell 192.168.218.1
2994	64.027959	192.168.218.129	93.184.220.29	TCP	55	[TCP Keep-Alive] 49719 → 80 [ACK] Seq=402 Ack=800 Win=63441 Len=1
2995	64.028059	93.184.220.29	192.168.218.129	TCP	60	[TCP Keep-Alive ACK] 80 → 49719 [ACK] Seq=800 Ack=403 Win=64240 Len=0

```
> Frame 1: 84 bytes on wire (672 bits), 84 bytes captured (672 bits) on interface \Device\NPF_{71DF18E2-F0D5-4411-94AA-CB74FC658319}, id 0
> Ethernet II, Src: VMware_92:e2:e0 (00:0c:29:92:e2:e0), Dst: VMware_ed:f9:15 (00:50:56:ed:f9:15)
> Internet Protocol Version 4, Src: 192.168.218.129, Dst: 192.168.218.2
> User Datagram Protocol, Src Port: 55535, Dst Port: 53
> Domain Name System (query)
```

```
0000  00 50 56 ed f9 15 00 0c 29 92 e2 e0 08 00 45 00  .PV... )....E.
0010  00 46 96 83 00 00 80 11 00 00 c0 a8 da 81 c0 a8  .F..... .....
0020  da 02 d8 ef 00 35 00 32 36 19 7a 96 01 00 00 01  ....5-2 6-z....
0030  00 00 00 00 00 00 0c 64 65 74 65 63 74 70 6f 72  .....d etectpor
0040  74 61 6c 07 66 69 72 65 66 6f 78 03 63 6f 6d 00  tal-fire fox.com
0050  00 01 00 01  .....  
Ethernet0: <live capture in progress> || Pacchetti: 2995 · visualizzati: 2995 (100.0%) || Profilo: Default
```

## CONFIGURAZIONE DEL SISTEMA



### Proxy o VPN

Utilizzo di un proxy o altre modalità per alterare il traffico



### Certificati

Configurazione di autorità di certificazione estranee

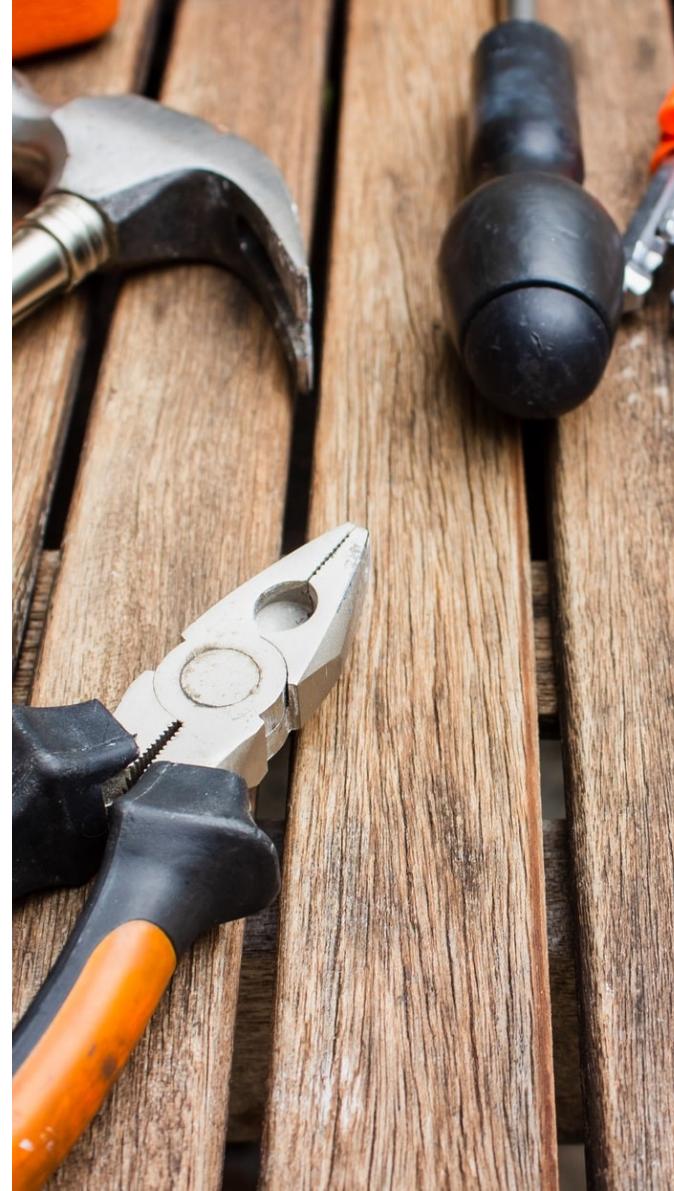


### DNS

Modifiche al puntamento del DNS o al file degli host

# Strumenti principali che vedremo

- FAW (Forensic Acquisition of Websites)
- Magnet Web Page Saver
- OSIRT Browser
- ArchiveWeb.page



# Strumenti remoti

## SITI DI ARCHIVIAZIONE

Nascono essenzialmente per supportare il nuovo lavoro archivistico digitale, naturale evoluzione della preservazione della conoscenza attuata dai bibliotecari.

Proveremo Archive.org e Archive Today.

## ALTRI SERVIZI SPECIFICI

Esistono alcune (pochissime) soluzioni remote specifiche per l'acquisizione forense interattiva, per esempio Cliens o LegalEye.

## ALTRI STRUMENTI



mitmproxy

Eccellente per analizzare e registrare flussi HTTP



Wireshark

Consente di registrare il traffico di rete e ispezionare i protocolli



Carbon14

Serve a stimare la data di pubblicazione di una pagina

# Metodologie operative

UNA PROCEDURA DA SEGUIRE E ALCUNI PUNTI CRITICI



# Partire dallo strumento

Questo atteggiamento purtroppo si riscontra comunemente quando si parla di diverse tematiche legate alla *digital forensics*.

Partendo da un argomento molto generico e variegato, a volte c'è la tendenza a offrire **una soluzione immediata e “universale”** anche quando ogni istanza del problema è particolare.

Questa metodologia di lavoro è errata e può portare a commettere degli errori.





*I suppose it is tempting,  
if the only tool you have is a hammer,  
to treat everything as if it were a nail.*

— ABRAHAM MASLOW  
"LAW OF THE INSTRUMENT"

# Metodologia

## PIANIFICARE

Prima è opportuno capire cosa si vuole acquisire, quali sono gli elementi importanti e qual è l'obiettivo per il quale si sta effettuando il lavoro.

## ESEGUIRE

Solo successivamente si può procedere a effettuare l'acquisizione in modo adeguato e procedere a verbalizzare le operazioni svolte.

# Elementi da valutare

Alcuni aspetti che devono essere soppesati:

- Tipologia di accesso al contenuto
- Dati relativi alle richieste HTTP
- Nomi e percorsi dei file
- Presenza di contenuti “complessi” o non standard



## TIPOLOGIA DI ACCESSO



Pubblico

Probabilmente non avrò problemi a usare strumenti esterni e/o non interattivi



Privato

Dovranno essere inserite delle credenziali, preferibilmente cercando di proteggerle

## RICHIESTE HTTP



### Codice di stato e header

Contengono elementi preziosi quali i cookie, la data di modifica e la natura della risposta



### Dati POST

Il contenuto potrebbe comparire solamente compilando un modulo in un certo modo

## NOMI DEI FILE

```
▼ <article class="QBXjJ M9sTE L_LMM JyscU ePUX4 " role="presentation" tabindex="-1"> event flex
  ▶ <header class="Ppjfr UE9AK wd0qh"> ... </header> flex
  ▶ <div class="MEAGS"> ... </div> flex
  ▼ <div class="_97aPb wKWK0"> flex
    ▼ <div class="ZyFrc" role="button" tabindex="0"> event flex
      ▶ <div class="eLAPa kPFhm">
        ▼ <div class="KL4Bh" style="padding-bottom: 100%;">
           event
        </div>
        <div class="_9AhH0"></div> flex
      </div>
```

Alcuni strumenti **non preservano i nomi e i percorsi** originali dei file, per esempio Archive Today.

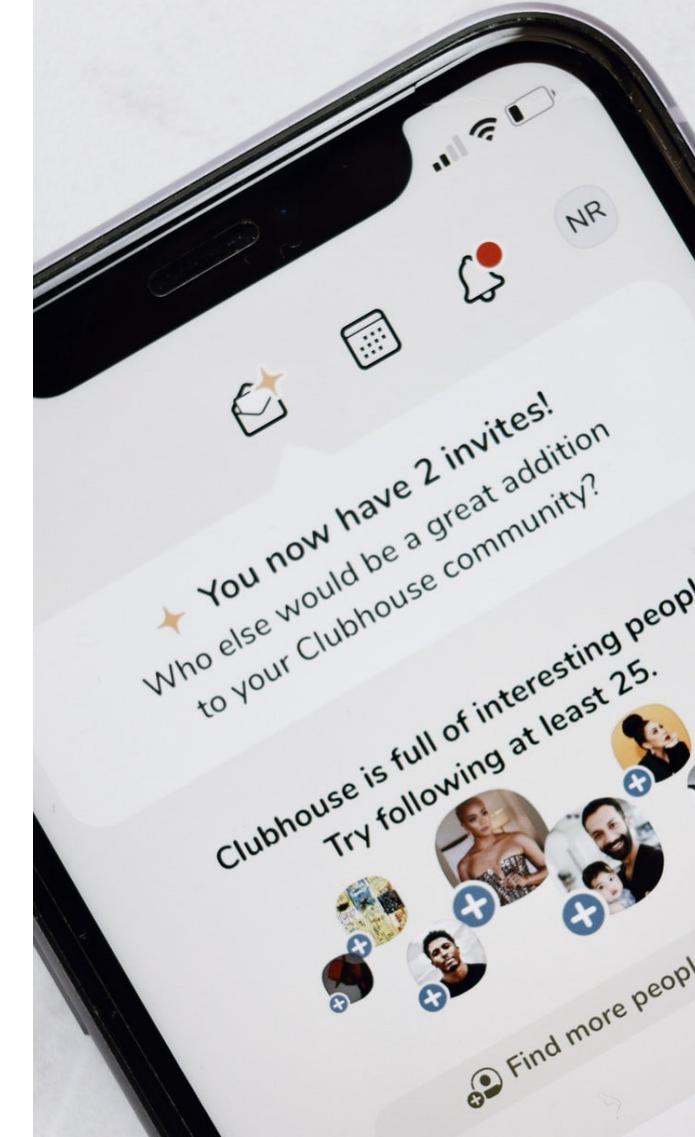
Essi potrebbero essere determinanti in alcune circostanze.

# Contenuti più “complessi”

In alcuni casi potremmo trovarci di fronte alla necessità di dover acquisire dei contenuti complessi, come i video pubblicati sulle piattaforme di streaming.

Alcuni contenuti distribuiti tramite tecnologie web di fatto **non sono neppure fruibili del tutto tramite browser**, ma solo con modalità specifiche.

Si pensi alle applicazioni mobili sviluppate con tecnologie “ibride” o quelle che fanno uso di API di tipo REST. In questi casi potrebbe essere necessario “intercettare” lo smartphone con mitmproxy.



## CONTACT

### Web

[andrealazzarotto.com](http://andrealazzarotto.com)

### GitHub

Lazza

### Twitter

@thelazza



# Demo

LABORATORIO CON ESEMPI PRATICI

