MODULO 3 - Pratica 2. Raccolta informazioni

Esercizio: Utilizzare i comandi di Google Hacking per raccogliere informazioni su un sito web.

Istruzioni:

- 1. Aprire un browser web e accedere a Google.
- 2. Utilizzare i seguenti comandi di Google Hacking per raccogliere informazioni sul sito web:
- "site:nome-del-sito.com" per visualizzare tutte le pagine indicizzate di quel sito.
- "inurl:nome-del-sito.com" per visualizzare tutte le pagine con l'URL contenente il nome del sito.
- "intext:'parola chiave' site:nome-del-sito.com" per visualizzare tutte le pagine che contengono la parola chiave specificata nel testo del sito.
- "filetype:estensione site:nome-del-sito.com" per visualizzare tutti i file con l'estensione specificata presenti sul sito.
- 3. Utilizzare i risultati **per identificare eventuali informazioni sensibili o vulnerabilità** presenti sul sito.
- 4. Utilizzare queste informazioni per valutare la sicurezza del sito e prendere le misure necessarie per proteggere le informazioni sensibili.

SVOLGIMENTO

SITO WEB ANALIZZATO: aslroma2.it

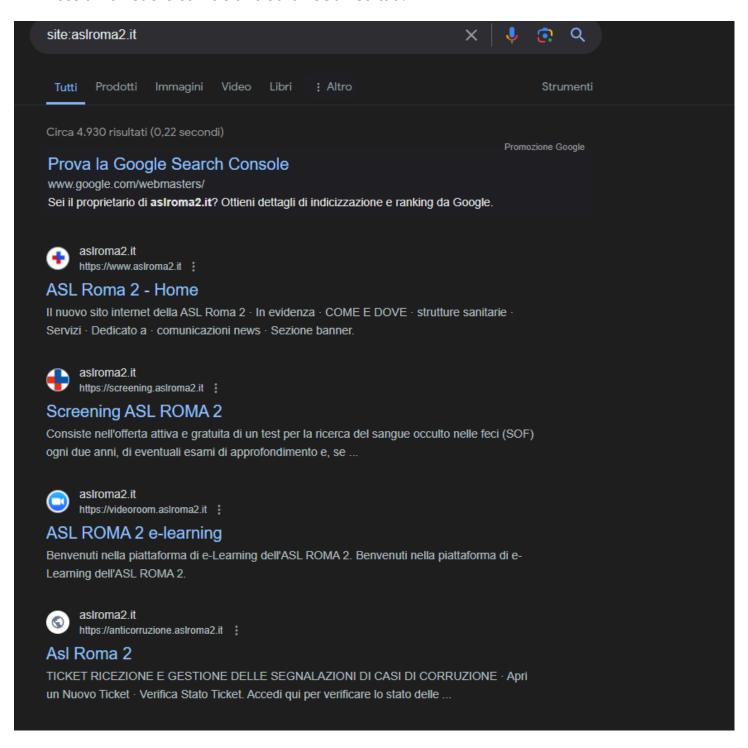
MOTIVAZIONE: La scelta di valutare la sicurezza di questo sito è legata all'importanza che esso detiene; si tratta di un sito che si collega a diverse altre piattaforme ed applicativi, come nel caso di AREAS e UNICA, e che permette inoltre di visionare anche file personali e privati legati al personale medico e all'organizzazione stessa.

ANALISI CON GOOGLE DORKS

PRIMA RICERCA: Google Dork "site:".
Utilizzando "site:" in un comando di ricerca verranno forniti risultati solo dal sito Web specifico menzionato.

"site:aslroma2.it":

Possiamo vedere come siano sorti 4930 risultati.

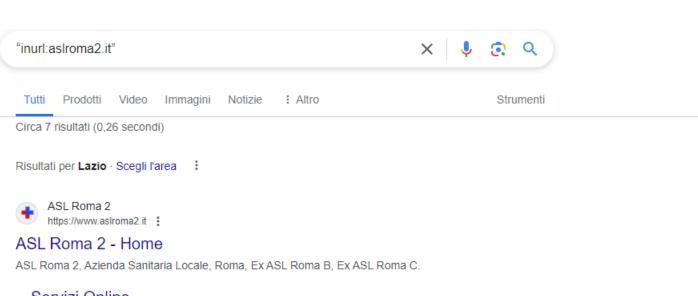


2- SECONDA RICERCA: Google Dork "inurl:".

Questo Dork specifico permette di visionare tutti gli URL in cui è presente il nome del sito preso in esame.

"inurl:aslroma2.it"

Otteniamo 7 soli risultati.



Servizi Online

ASL Roma 2, Azienda Sanitaria Locale, Roma, Ex ASL Roma B ...

Concorsi

Inserimento 18/12/2023. Ultima modifica 18/12/2023. AVVISO ...

Speciale vaccini

è possibile richiedere una ricerca sull'Archivio Storico Vaccinale ...

Azienda

L'azienda sanitaria ASL Roma 2 · Modello assistenziale per ...

Distretti

ASL Roma 2, Azienda Sanitaria Locale, Roma, Ex ASL Roma B ...

Altri risultati in aslroma2.it »



asl roma 2 - amministrazione trasparente

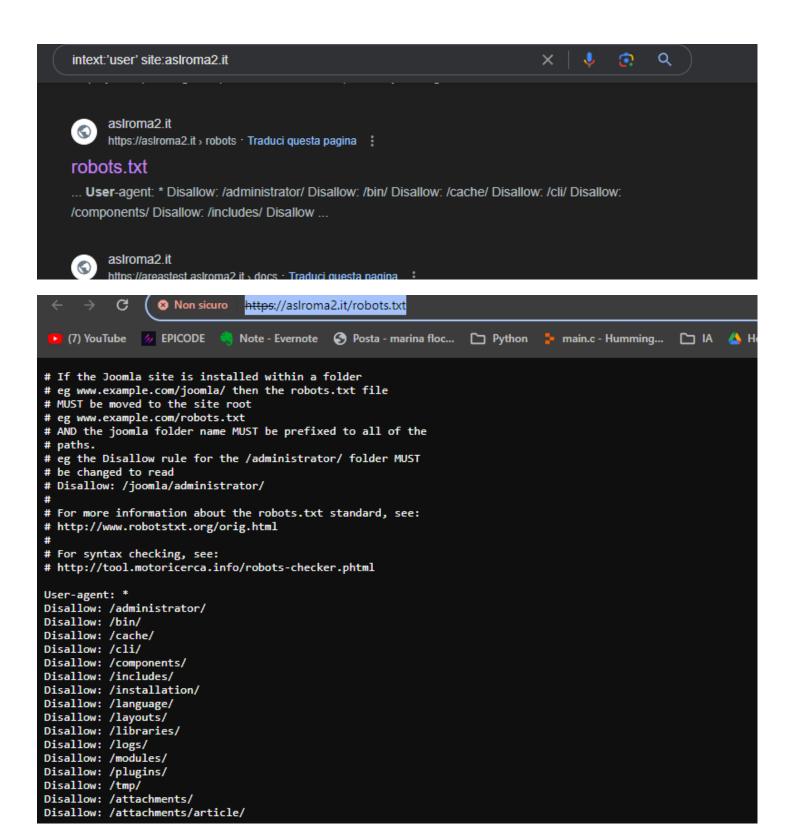
Questa è una sezione del sito ufficiale della ASL Roma 2 relativa all'Amministrazione Trasparente. La trasparenza è accessibilità totale ai dati e ai ...



3- TERZA RICERCA: Google Dork "intext: 'parola chiave'.

Questo servirà a ridurre i risultati di ricerca andando a trovare solo le pagine contenti al loro interno il termine specificato.

intext:'user' site:aslroma2.it



Adesso abbiamo individuato un primo elemento importante, ovvero il file "robots.txt".

Esso contiene informazioni sulla struttura di un sito web, e può essere utilizzato da un utente malintenzionato per conoscere risorse che non possono essere raggiunte semplicemente scansionando ripetutamente i collegamenti ipertestuali.

Se seguiamo le comuni pratiche di sicurezza durante la costruzione di un server web, dobbiamo sicuramente aver disabilitato l'elenco delle directory e creato alcune regole per l'accesso alle

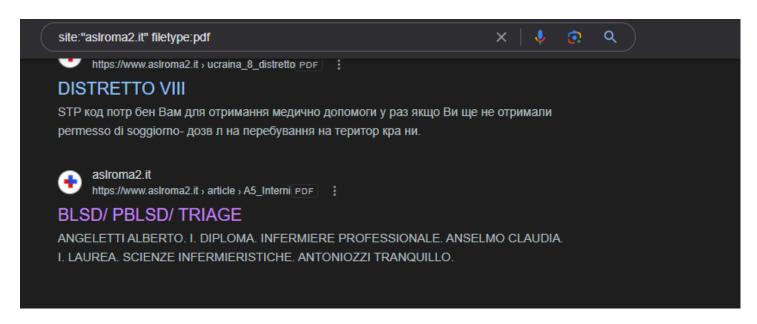
risorse.

Esiste tuttavia il rischio che gli aggressori approfittino del file robots per conoscere la struttura del nostro server web.

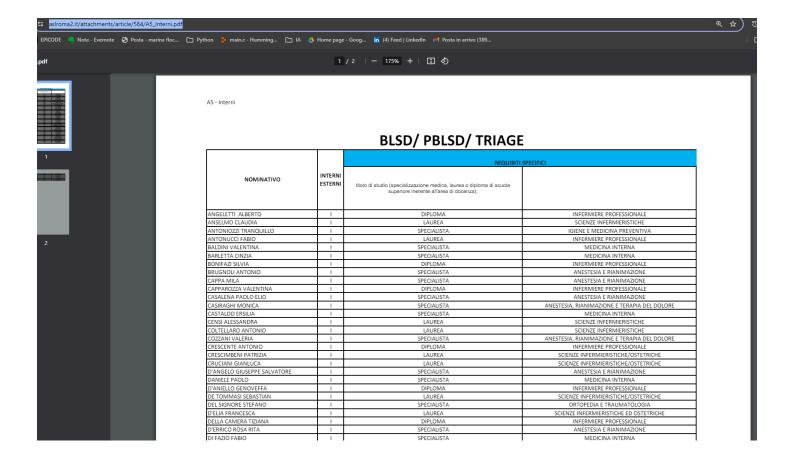
4- QUARTA RICERCA: Google Dork "filetype:estensione".

Con esso si mira a trovare ed ottenere file di un formato particolare, come ad esempio file Word o PDF presenti in un sito specifico.

site:aslroma2.it filetype:pdf



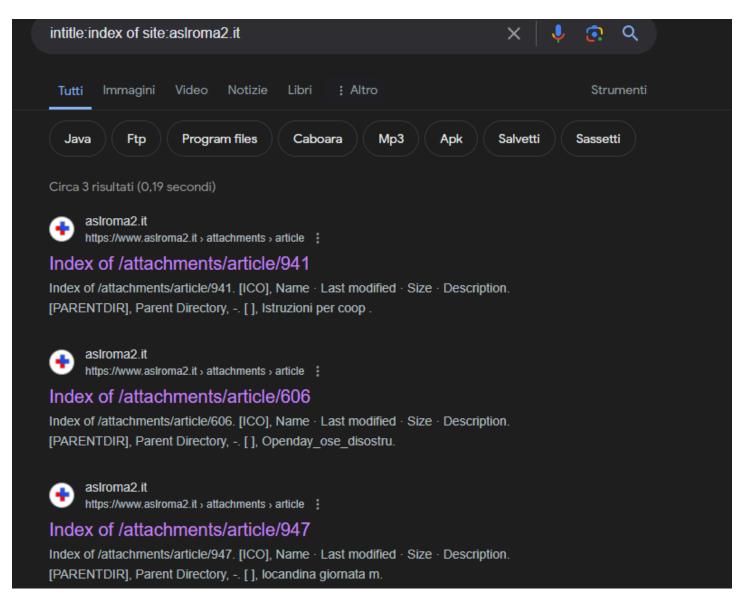
Troviamo diversi tipi di file pdf disponibili per la visione, tra cui anche un documento contenente nominativi e titoli di studio di alcuni membri dello staff medico del TRIAGE della ASL Roma 2.



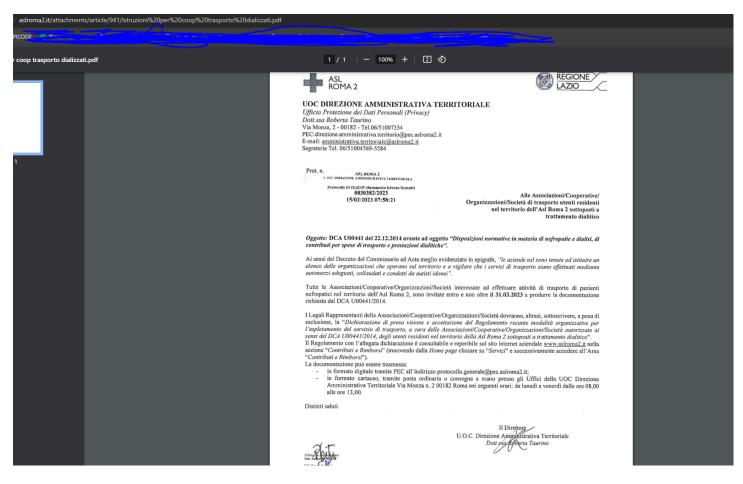
Proviamo ad effettuare qualche altra ricerca.

A questo punto ci interessa provare anche il Google Dork "index of".

-intitle:index of site:aslroma2.it



Con questa ricerca otteniamo la visibilità di diversi file inseriti sul sito, ma che non dovrebbero essere accessibili pubblicamente. Un esempio si trova nell'immagine sottostante, che mostra una comunicazione interna della ASL alle società e cooperative territoriali.



Inoltre è possibile risalire fino alla parent directory di questa sezione del sito, la quale mostra altri diversi articoli.

Index of /attachments/article

	<u>Name</u>	Last modified	Size Description
Par	ent Directory	Ĺ	-
238	<u>8/</u>	2016-09-01 15:30	-
239	9/	2016-09-01 15:34	_
228	<u>3/</u>	2016-09-01 16:38	-
241	<u>l/</u>	2016-09-09 13:24	-
229	9/	2016-09-09 13:53	-
230	<u>)/</u>	2016-09-09 13:59	-
242	2/	2016-09-13 11:29	_
243	<u>3/</u>	2016-09-19 17:03	-
221	<u>l/</u>	2016-09-29 12:57	-
248	<u>3/</u>	2016-09-30 10:42	-
249	<u>9/</u>	2016-10-03 17:36	-
<u>236</u>	<u>5/</u>	2016-10-06 16:03	-
252	2/	2016-10-06 18:04	_
105	5/	2016-10-12 15:47	-
<u>128</u>	<u>3/</u>	2016-10-21 14:45	-
<u>108</u>	<u>3/</u>	2016-11-10 11:08	-
122	2/	2016-11-15 14:00	-
246	<u>5/</u>	2016-11-23 10:22	-
<u>183</u>	<u>3/</u>	2016-12-01 14:04	-
<u>267</u>	7/	2016-12-14 10:19	-
<u>177</u>	7/	2016-12-14 10:39	-
205	5/	2017-01-03 18:52	-
206	<u>5/</u>	2017-01-03 18:53	-
270	<u>)/</u>	2017-01-16 13:34	-
<u>191</u>	<u>l/</u>	2017-01-20 14:18	-
166	<u>5/</u>	2017-01-31 17:49	-
285	5/	2017-02-03 08:57	-
287	7/	2017-02-13 13:13	-
280	a /	2017-02-23 14:49	-

VULNERABILITA' INDIVIDUATE O POTENZIALI TALI

- 1. La prima vulnerabilità che abbiamo potuto notare in questa piccola ricerca è stato il file di testo robots.txt. Esso infatti fornisce informazioni utili per comprendere le directory più sensibili, lasciando così indicazioni rilevanti ad un potenziale attaccante a caccia di dati sensibili.
- La seconda vulnerabilità che abbiamo rilevato è quella riguardante gli articoli pubblicati nel sito della ASL Roma 2. Vi sono comunicazioni interne e pubblicazioni che potrebbero fornire materiale per attacchi su diversi fronti, non solo verso la ASL, ma anche verso i partner, impiegati e aziende collaboratrici.

RIMEDIO

1. Se una determinata risorsa o directory non è accessibile da remoto, non deve essere posizionata su una macchina esposta a Internet oppure l'accesso deve essere limitato tramite le regole di configurazione del server web. Rimuovere tutti i collegamenti che puntano ad esso da qualsiasi altra pagina Web non è sufficiente.

Possiamo anche evitare di suggerire all'aggressore quali sarebbero gli obiettivi più preziosi **non enumerando le risorse a cui i crawler non dovrebbero accedere** . Supponiamo di non consentire generalmente la scansione (indicando *Disallow: /* nel file robots) ma di consentire l'accesso a singole risorse o directory (utilizzando il tasto *Consenti*). In tal caso non menzioneremo esplicitamente i percorsi più sensibili. Possiamo anche limitare del tutto l'utilizzo del file robots includendo le indicazioni rilevanti per i crawler all'interno dell'intestazione HTML di una pagina, con:

<meta name="robots" content="noindex" />

2- Per evitare che il file robots.txt possa risultare utile a degli utenti malintenzionati, si potrebbe modificare il file per disabilitare l'indicizzazione delle directory che desideri mantenere private. Per fare ciò si possono aggiungere istruzioni nel file robots.txt per impedire ai motori di ricerca di indicizzare determinate directory o tutto il sito, se necessario.

Se il sito utilizza Apache come server web, con il file .htaccess si può impedire la visualizzazione dell'elenco di file all'interno delle directory senza un file index.html o index.php.

Aggiungendo ad ogni directory del sito un file index.html o index.php si impedirà al server web di elencare i file contenuti nella directory quando non è presente un file index predefinito.

Le directory sensibili si possono inoltre proteggere con delle password, o impostando nel modo giusto le autorizzazioni dei file.