Digital Forensics Tool Validation: String Search Analysis with Autopsy

Progetto di Tool Validation
per il Corso di **Digital Forensics**

Lucrezia Ilvento

Laurea magistrale in **Informatica**

Anno Accademico **2023/2024**

Contents

Chapter 1

Introduzione

1.1 Contesto e Obiettivi

Nel campo della digital forensics, la validazione degli strumenti è essenziale per garantire l'accuratezza e l'affidabilità dei risultati ottenuti durante le indagini. Gli strumenti di digital forensics devono essere in grado di analizzare e recuperare dati in modo preciso e riproducibile, per supportare le indagini legali e garantire la giustizia. La validazione degli strumenti assicura che questi funzionino correttamente e producano risultati coerenti con le aspettative e le specifiche dichiarate dai produttori.

La validazione della funzionalità di ricerca di stringhe è particolarmente cruciale, poiché molti casi di digital forensics si basano sulla capacità di identificare e analizzare stringhe di testo all'interno di file e dispositivi. Le stringhe di testo possono contenere prove cruciali che devono essere recuperate e analizzate accuratamente. Pertanto, è fondamentale che gli strumenti utilizzati siano rigorosamente testati e convalidati.

Il progetto *Computer Forensics Tool Testing (CFTT)* del *National Institute of Standards and Technology (NIST)* è un'iniziativa chiave volta a garantire la qualità e l'affidabilità degli strumenti di forensics. Questo progetto mira a stabilire una metodologia standardizzata per il test degli strumenti di forensics informatici, sviluppando specifiche generali per gli strumenti, procedure di test, criteri di valutazione e set di test. I risultati forniscono informazioni essenziali ai produttori per migliorare i loro strumenti, agli utenti per fare scelte informate e ad altre parti interessate per comprendere le capacità degli strumenti.

In questo contesto, il presente progetto si concentra sulla validazione della funzionalità di ricerca di stringhe nel tool *Autopsy*. L'obiettivo è valutare se Autopsy soddisfi i requisiti di precisione e affidabilità necessari per un uso efficace in digital forensics, e confrontare le sue prestazioni con quelle di altri strumenti open source nel settore.

Chapter 2

Descrizione degli Strumenti e dei Dataset

2.1 Autopsy

Autopsy è un software open source progettato per l'analisi forense di dispositivi e file system. Dotato di un'interfaccia grafica intuitiva e di strumenti modulari, Autopsy supporta le indagini digitali attraverso diverse funzionalità chiave come il recupero di dati cancellati, l'analisi di file e metadati, e la visualizzazione e filtraggio dei contenuti per ottenere informazioni rilevanti. Inoltre, Autopsy consente la generazione di report dettagliati basati sui risultati delle analisi.

2.2 Dataset Utilizzato

Il dataset fornito dal NIST, all'interno del programma *Computer Forensics Tool Testing (CFTT)*, è stato progettato per testare le capacità dei tool forensi nella ricerca e gestione di stringhe di testo. Le tipologie di test presenti nel dataset includono:

- **Ricerca ASCII e Unicode**: Verifica la capacità del tool di trovare stringhe in formati di testo ASCII e Unicode, con e senza distinzione tra maiuscole e minuscole.
- **Ricerca di Substringhe**: Testa l'accuratezza del tool nella ricerca di corrispondenze parziali all'interno delle parole.
- **Operazioni Logiche**: Include test che utilizzano operatori logici come *AND*, *OR* e *NOT* per combinare più condizioni di ricerca.
- **Test per Stringhe Unicode con Segni Diacritici**: Valuta la capacità del tool di gestire caratteri Unicode con segni diacritici e differenti rappresentazioni dello stesso carattere (NFC, NFD).
- **Stringhe in Lingue Diverse**: Include stringhe in varie lingue (latino, ideogrammi cinesi e giapponesi, cirillico, arabo) per testare il supporto multilingue di Autopsy.

- **Ricerche su File Particolari**: Ricerche in contesti particolari come file frammentati, aree inaccessibili e metadati.

Questo dataset offre un'ampia gamma di scenari che consentono di valutare in dettaglio l'accuratezza e l'affidabilità di Autopsy nella ricerca di stringhe.

Chapter 3

Setup e Procedura del Test

3.1 Preparazione dell'Ambiente

La preparazione dell'ambiente di test è stata fondamentale per garantire che Autopsy fosse correttamente configurato per eseguire l'analisi delle stringhe Unicode. Prima di iniziare l'analisi, è stato necessario caricare il dataset fornito dal NIST e selezionare i moduli essenziali per il processo di ingestione e analisi.

3.1.1 Installazione di Autopsy

Autopsy versione 4.21.0 è stato installato su un sistema operativo Windows 11, dotato di processore Intel Core i7 e 16 GB di RAM, per garantire una potenza di calcolo sufficiente a gestire i file immagine di grandi dimensioni e le analisi complesse.

3.1.2 Caricamento del Dataset

Il dataset fornito dal NIST è stato caricato come un'immagine disco nel formato `.dd`, che rappresenta una copia forense di un disco rigido. Per caricare correttamente il file immagine in Autopsy, è stato selezionato il tipo di sorgente dati *"Disk Image or VM File"*. Successivamente, sono state configurate le opzioni di analisi per indicizzare i file e le partizioni presenti nell'immagine disco.

3.1.3 Selezione dei Moduli di Ingestione

Sono stati abilitati solo i moduli essenziali per la ricerca di stringhe, riducendo il carico computazionale. Tra questi::

- **File Type Identification**: Per identificare i tipi di file presenti nell'immagine.
- **Keyword Search**: Per abilitare la ricerca di stringhe, sia in formato ASCII che Unicode.
- **Embedded File Extractor**: Per estrarre i file incorporati in archivi compressi o contenuti all'interno di altri file.
- **Archive Extractor**: Per analizzare gli archivi compressi come ZIP e RAR.

I moduli come il *File Carving* e l'*E-mail Parser* sono stati esclusi per ridurre i tempi di analisi.

3.2 Esecuzione dell'Analisi

L'analisi dell'immagine disco è stata avviata per indicizzare i file e prepararli per la ricerca. Durante l'analisi, il processo si è fermato temporaneamente al 98%, ma, dopo aver verificato le risorse di sistema, è stato completato con successo.

3.2.1 Verifica degli Indici di Ricerca

Una volta completata l'analisi, è stato confermato che gli indici necessari per la ricerca fossero stati correttamente creati. L'indicizzazione è stata effettuata utilizzando il motore Solr, che consente ricerche efficienti.

3.3 Ricerca delle Stringhe

Una volta completata l'analisi e creati gli indici, è stata avviata la fase di ricerca delle stringhe. Ogni test case è stato eseguito singolarmente per garantire che tutte le opzioni richieste fossero configurate correttamente. In alcuni casi, dove era necessario cercare più stringhe contemporaneamente, è stata creata una lista di parole chiave. Le stringhe e le impostazioni per ogni test case sono state fornite dal NIST. La configurazione delle ricerche ha riguardato varie opzioni, tra cui la sensibilità al maiuscolo/minuscolo (*match case*), la ricerca di stringhe ASCII o Unicode, e l'utilizzo di operatori logici (*AND*, *OR*, *NOT*).

3.3.1 Impostazione della Ricerca

Per ogni test case, è stata configurata la ricerca in *Autopsy* utilizzando il modulo *Keyword Search*. Di seguito vengono descritte le configurazioni principali adottate per ciascun test case:

- **Case Sensitivity**: Nei test case che richiedevano la sensibilità al maiuscolo/minuscolo, è stata abilitata l'opzione *Regular Expression*, poiché *Autopsy* è case-insensitive per impostazione predefinita.
- **Exact Match vs Substring Match**: Nei test case in cui era necessario cercare solo parole intere, è stata abilitata l'opzione *Exact Match*. Nei casi in cui era richiesto includere anche sottostringhe, è stata selezionata l'opzione *Substring Match*.
- **Operatori Logici**: Poiché *Autopsy* non supporta direttamente operatori logici avanzati (*AND*, *OR*, *NOT*), sono state eseguite ricerche separate per ciascuna stringa, e i risultati sono stati combinati manualmente.

3.3.2 Casi di Test

Di seguito si riporta un elenco dei casi di test eseguiti, insieme alle rispettive configurazioni:

- **FT-SS-01**: Questo test case richiedeva la ricerca della stringa *DireWolf* con l'opzione *Regular Expression* attiva e l'abilitazione della ricerca in formato *ASCII*.
- **FT-SS-02**: Questo test case richiedeva la ricerca della stringa *wolf*, disabilitando la sensibilità al maiuscolo/minuscolo e attivando l'opzione *Substring Match* per consentire la ricerca di sottostringhe.
- **FT-SS-03**: Questo test case richiedeva la ricerca della stringa *Wolf*, disabilitando la sensibilità al maiuscolo/minuscolo e attivando l'opzione *Exact Match* per trovare solo corrispondenze esatte.
- **FT-SS-04**: In questo caso, è stata testata la ricerca logica *AND* per i termini *panda* e *fox*. L'opzione *Exact Match* era attiva, ma il confronto per individuare file che contenessero entrambi i termini è stato effettuato manualmente.
- **FT-SS-05**: Questo test case richiedeva la ricerca logica *OR* per i termini *Were* o *DireW*. L'opzione *Regular Expression* era attiva. Le ricerche sono state eseguite separatamente, e i risultati sono stati successivamente combinati manualmente.
- **FT-SS-06**: In questo caso, la ricerca logica *NOT* è stata eseguita cercando la stringa *fox*, escludendo il termine *tiger*. La ricerca è stata effettuata per *fox*, e i risultati sono stati filtrati manualmente per escludere quelli contenenti *tiger*.
- **FT-SS-07-Latin**: Questo test case si concentrava sulla ricerca di caratteri Unicode con segni diacritici. La ricerca Unicode è stata eseguita senza necessità di modifiche, poiché *Autopsy* supporta Unicode di default.
- **FT-SS-07-CJK-char**: In questo caso è stata testata la ricerca di ideogrammi CJK (*Cinese*, *Giapponese*, *Coreano*). La ricerca Unicode è stata eseguita senza necessità di modifiche, in quanto supportata da *Autopsy* di default.
- **FT-SS-07-Norm**: Per questo caso è stato utilizzato il file `ft-ss-07-Norm-strings.txt`, contenente stringhe con diacritici normalizzati nelle forme *NFC* e *NFD*. Sono state eseguite ricerche separate per ciascuna versione della stringa per valutare come *Autopsy* gestisce le differenze nella rappresentazione Unicode.
- **FT-SS-08-Email**: Questo caso di test richiedeva una ricerca predefinita di indirizzi email. Non è stata inserita una stringa specifica, ma è stata selezionata l'opzione *Email Address* all'interno del motore di ricerca di *Autopsy*.
- **FT-SS-09-Doc**: Questo test case richiedeva la ricerca di stringhe come *flintlock*, *rifle*, *shotgun*, *revolver*, *longbow*, *crossbow*, *peroxide*, e *nitroglycerin* all'interno di file formattati come `.doc`, `.docx`, e `.html`. È stata attivata l'opzione *Exact Match*.

- **FT-SS-10-Hex**: Questo test case richiedeva la ricerca di stringhe in formato esadecimale. Poiché *Autopsy* non prevede nativamente questa funzionalità, la ricerca non è stata eseguita.

- **FT-SS-10-Regex**: In questo caso è stata utilizzata un'espressione regolare per cercare il pattern `[DW]..eWolf`, che rappresenta una stringa che inizia con "D" o "W" e ha due caratteri seguenti prima di terminare con *eWolf*. Questa ricerca è stata eseguita con l'opzione *Regular Expression*.

Chapter 4

Risultati

Il set di dati e i casi di test utilizzati per questo report si concentrano sugli aspetti più comunemente riscontrati nella ricerca di stringhe testuali in un contesto forense. Anche se non tutte le funzionalità sono state coperte, i test eseguiti hanno abbracciato una vasta gamma di caratteristiche fondamentali. Le principali funzionalità testate con *Autopsy 4.21* includono:

- **File System**: I file system testati includono ambienti *MS Windows* (FAT, exFAT, NTFS).
- **Posizione della stringa**: Le stringhe sono state cercate in file attivi, file cancellati (recuperabili), spazi non allocati e meta-dati.
- **Metodo di ricerca**: Sono state eseguite ricerche indicizzate utilizzando *Autopsy*, per confrontare i risultati con il set di dati fornito dal NIST.
- **Codifica della stringa**: Le stringhe sono state codificate in *ASCII*, *UTF-8*, *UTF-16BE* e *UTF-16LE*.
- **Unicode normalizzato**: I test hanno verificato come *Autopsy* gestisce forme alternative di rappresentazione dei caratteri, come legature e segni diacritici.
- **Lingue**: Sono state testate stringhe in diverse lingue, tra cui segni diacritici (tedesco, francese, spagnolo), caratteri non latini (russo), testo in scrittura da destra a sinistra (arabo) e lingue asiatiche (cinese, giapponese, coreano).
- **File frammentati**: Sono state testate stringhe che si estendono su più frammenti di file.
- **Operatori logici**: Sono state eseguite ricerche utilizzando operatori logici come *AND*, *OR*, e *NOT*.
- **Ricerca Stemming**: Ricerche che avrebbero dovuto individuare forme flessive derivate da un tema lessicale (ad esempio, la ricerca della parola "run" doveva trovare anche "running", "ran", ecc.).
- **Formattazione incorporata**: Sono stati eseguiti test su documenti con formattazione incorporata, come file *MS Word* e *HTML*.

4.1 Caratteristiche non supportate da Autopsy 4.21

Durante i test, sono emerse alcune limitazioni di *Autopsy 4.21*. Le seguenti funzionalità non sono supportate nativamente:

- Gli operatori logici (*AND*, *OR*, *NOT*) non sono gestiti in modo nativo e devono essere simulati tramite ricerche separate.
- La ricerca basata su *stemming* non è supportata, il che limita la capacità di trovare parole flessive.
- La ricerca di stringhe in formato esadecimale non è prevista.

4.1.1 Risultati Osservati

La maggior parte delle stringhe presenti nei file attivi e cancellati è stata trovata correttamente, con poche eccezioni:

- Il numero di telefono **"(901)555-1111"** non è stato individuato dalla ricerca predefinita di numeri di telefono, ma è stato trovato correttamente tramite una ricerca manuale.
- La ricerca predefinita di **Social Security Numbers (SSN)** non è disponibile nella versione di *Autopsy* utilizzata. Tuttavia, inserendo manualmente le stringhe SSN nella ricerca di parole chiave, tutte le corrispondenze sono state trovate correttamente.
- La ricerca basata su **stemming** non è supportata da *Autopsy*. Pertanto, nel test case **FT-SS-09**, non tutte le parole derivate dal tema lessicale sono state trovate.
- La ricerca di stringhe in formato **esadecimale** non è supportata, causando il fallimento del test case **FT-SS-10**.
- Molte stringhe **UTF** nei test di ricerca in spazi non allocati non sono state trovate, evidenziando una limitazione nella gestione di questo tipo di dati.

Per una visione più dettagliata dei risultati, sono stati allegati screenshot della tabella degli **Expected Results**, scaricata dal sito del NIST, che è stata modificata con i seguenti colori:

- **In giallo**: Risultati corrispondenti a quelli attesi.
- **In rosso**: Risultati che non corrispondono a quelli attesi.
- **In verde**: Nuovi risultati inattesi trovati durante l'analisi.

MS Windows Data Set

Case/String	Encoding	Active	Deleted	Unalloc
FT-SS-01				
DireWolf	ascii	0896 0898 0900	0897 0899 0901	0902
FT-SS-02				
Wolf	ascii	0864 0866 0868	0865 0867 0869	0870
DireWolf	ascii	0896 0898 0900	0897 0899 0901	0902
WOLF	ascii	0880 0882 0884	0881 0883 0885	0886
WereWolf	ascii	0912 0914 0916	0913 0915 0917	0918
wolf	ascii	0848 0850 0852	0849 0851 0853	0854
FT-SS-03				
Wolf	ascii	0864 0866 0868	0865 0867 0869	0870
WOLF	ascii	0880 0882 0884	0881 0883 0885	0886
wolf	ascii	0848 0850 0852	0849 0851 0853	0854
FT-SS-04				
Panda and fox	ascii	2944 2946 2948	2945 2947 2949	2950
FT-SS-05				
DireWolf	ascii	0896 0898 0900	0897 0899 0901	0902
WereWolf	ascii	0912 0914 0916	0913 0915 0917	0918
FT-SS-06				
fox and not tiger				
	ascii	0784 0792 0800	0788 0796 0804	0808
	utf-16-be	0787 0795 0803	0791 0799 0807	
	utf-16-le	0786 0794 0802	0790 0798 0806	
	utf-8	0785 0793 0801	0789 0797 0805	
FT-SS-07-CJK-char				
中国				
	utf-16-be	1987 1995 2003	1991 1999 2007	2011
	utf-16-le	1986 1994 2002	1990 1998 2006	2010
	utf-8	1985 1993 2001	1989 1997 2005	2009
東京				
	utf-16-be	2883 2891 2899	2887 2895 2903	2907
	utf-16-le	2882 2890 2898	2886 2894 2902	2906
	utf-8	2881 2889 2897	2885 2893 2901	2905
FT-SS-07-CJK-hangul				
서울				

Figure 4.1: Expected Results - Parte 1

MS Windows Data Set

Case/String	Encoding	Active	Deleted	Unalloc
	utf-16-be	2371 2379 2387	2375 2383 2391	2395
	utf-16-le	2370 2378 2386	2374 2382 2390	2394
	utf-8	2369 2377 2385	2373 2381 2389	2393
FT-SS-07-CJK-kana				
みつびし				
	utf-16-be	2307 2315 2323	2311 2319 2327	2331
	utf-16-le	2306 2314 2322	2310 2318 2326	2330
	utf-8	2305 2313 2321	2309 2317 2325	2329
スバル				
	utf-16-be	2179 2187 2195	2183 2191 2199	2203
	utf-16-le	2178 2186 2194	2182 2190 2198	2202
	utf-8	2177 2185 2193	2181 2189 2197	2201
FT-SS-07-Cyrillic				
Собчак				
	utf-16-be	1187 1195 1203	1191 1199 1207	1211
	utf-16-le	1186 1194 1202	1190 1198 1206	1210
	utf-8	1185 1193 1201	1189 1197 1205	1209
FT-SS-07-Latin				
Schönheit				
	utf-16-be	2499 2507 2515	2503 2511 2519	2523
	utf-16-le	2498 2506 2514	2502 2510 2518	2522
	utf-8	2497 2505 2513	2501 2509 2517	2521
garçon				
	utf-16-be	2691 2699 2707	2695 2703 2711	2715
	utf-16-le	2690 2698 2706	2694 2702 2710	2714
	utf-8	2689 2697 2705	2693 2701 2709	2713
FT-SS-07-NoBOM				
QuarterHorse				
	ascii	2560 2568 2576	2564 2572 2580	2584
	utf-16-be	2563 2571 2579	2567 2575 2583	2587
	utf-16-le	2562 2570 2578	2566 2574 2582	2586
	utf-8	2561 2569 2577	2565 2573 2581	2585
Россия				
	utf-16-be	1059 1067 1075	1063 1071 1079	1083
	utf-16-le	1058 1066 1074	1062 1070 1078	1082
	utf-8	1057 1065 1073	1061 1069 1077	1081
ايران				
	utf-16-be	1539 1547 1555	1543 1551 1559	1563
	utf-16-le	1538 1546 1554	1542 1550 1558	1562
	utf-8	1537 1545 1553	1541 1549 1557	1561
中国				
	utf-16-be	1923 1931 1939	1927 1935 1943	1947
	utf-16-le	1922 1930 1938	1926 1934 1942	1946
	utf-8	1921 1929 1937	1925 1933 1941	1945
FT-SS-07-Norm				
Mäuse (NFD)				
	utf-16-be	3283 3291 3299	3287 3295 3303	3307
	utf-16-le	3282 3290 3298	3286 3294 3302	3306
	utf-8	3281 3289 3297	3285 3293 3301	3305
Mäuse (NFC)				
	utf-16-be	3091 3099 3107	3095 3103 3111	3115

Figure 4.2: Expected Results - Parte 2

MS Windows Data Set

Case/String	Encoding	Active	Deleted	Unalloc
	utf-16-le	3090 3098 3106	3094 3102 3110	3114
	utf-8	3089 3097 3105	3093 3101 3109	3113
infinity (No Ligature)				
	ascii	3344 3352 3360	3348 3356 3364	3368
	utf-16-be	3347 3355 3363	3351 3359 3367	3371
	utf-16-le	3346 3354 3362	3350 3358 3366	3370
	utf-8	3345 3353 3361	3349 3357 3365	3369
infinity (Ligature)				
	utf-16-be	3411 3419 3427	3415 3423 3431	3435
	utf-16-le	3410 3418 3426	3414 3422 3430	3434
	utf-8	3409 3417 3425	3413 3421 3429	3433
liberté (NFD)				
	utf-16-be	3219 3227 3235	3223 3231 3239	3243
	utf-16-le	3218 3226 3234	3222 3230 3238	3242
	utf-8	3217 3225 3233	3221 3229 3237	3241
liberté (NFC)				
	utf-16-be	3027 3035 3043	3031 3039 3047	3051
	utf-16-le	3026 3034 3042	3030 3038 3046	3050
	utf-8	3025 3033 3041	3029 3037 3045	3049
mañana (NFD)				
	utf-16-be	3155 3163 3171	3159 3167 3175	3179
	utf-16-le	3154 3162 3170	3158 3166 3174	3178
	utf-8	3153 3161 3169	3157 3165 3173	3177
mañana (NFC)				
	utf-16-be	2963 2971 2979	2967 2975 2983	2987
	utf-16-le	2962 2970 2978	2966 2974 2982	2986
	utf-8	2961 2969 2977	2965 2973 2981	2985
FT-SS-07-RTL				
مرحبا				
	utf-16-be	1475 1483 1491	1479 1487 1495	1499
	utf-16-le	1474 1482 1490	1478 1486 1494	1498
	utf-8	1473 1481 1489	1477 1485 1493	1497
FT-SS-08-Email				
berlin@deutchland.net iron.man@marvel.com	ascii	1280 1282 1284	1281 1283 1285	1286
	ascii	0944 0952 0960	0948 0956 0964	0968
	utf-16-be	0947 0955 0963	0951 0959 0967	0971
	utf-16-le	0946 0954 0962	0950 0958 0966	0970
	utf-8	0945 0953 0961	0949 0957 0965	0969
kgb@mosgov.red.square.ru	ascii	1264 1266 1268	1265 1267 1269	1270
potus@capitol.gov	ascii	1248 1250 1252	1249 1251 1253	1254
FT-SS-08-Phone				
(901)555-1111	ascii	1296 1298 1300	1297 1299 1301	1302
202.555.3270	ascii	1392 1394 1396	1393 1395 1397	1398
301.555-9009				
	ascii	1328 1336 1344	1332 1340 1348	1352

Figure 4.3: Expected Results - Parte 3

MS Windows Data Set

Case/String	Encoding	Active	Deleted	Unalloc
	utf-16-be	1331 1339 1347	1335 1343 1351	1355
	utf-16-le	1330 1338 1346	1334 1342 1350	1354
	utf-8	1329 1337 1345	1333 1341 1349	1353
800-555-1122	ascii	1312 1314 1316	1313 1315 1317	1318
FT-SS-08-SS				
123-45-6789	ascii	1008 1010 1012	1009 1011 1013	1014
987-65-4321	ascii	1024 1026 1028	1025 1027 1029	1030
999-55-1321	ascii	1040 1042 1044	1041 1043 1045	1046
FT-SS-09-Doc				
crossbow	utf-8	8007 9007		8515 9516
flintlock	utf-8	8001 9001		8508 9509
longbow	utf-8	8006 9006		8514 9514
nitroglycerin				
	UTF	9005		9515
	utf-8	8005		8513
peroxide				
	UTF	9004		9516
	utf-8	8004		8512
revolver				
	utf-16-be	9002		9510
	utf-8	8002		8510
rifle	utf-8	8000 9000		8508 9508
shotgun				
	utf-16-be	9003		9511
	utf-8	8003		8511
FT-SS-09-Frag				
California	utf-8	6000		
Washington	utf-8	6006		
FT-SS-09-Lost				
SecretKey				
	ascii			7000
	utf-16-be			7003
disconnected				
	ascii			7016
	utf-16-le			7018
FT-SS-09-MFT				
bear				
	ascii	7008	7012	
	utf-16-be	7011	7015	
	utf-16-le	7010	7014	
	utf-8	7009	7013	
FT-SS-09-Meta				
cañón	utf-8	2625 2633 2641	2629 2637 2645	2649
thunderbird	ascii	0928 0930 0932	0929 0931 0933	0934
FT-SS-09-Stem				
cities				
	ascii	0480 0488 0496	0484 0492 0500	0504
	utf-16-be	0483 0491 0499	0487 0495 0503	0507
	utf-16-le	0482 0490 0498	0486 0494 0502	0506
	utf-8	0481 0489 0497	0485 0493 0501	0505
city	ascii	0464 0466 0468	0465 0467 0469	0470
knife	ascii	0576 0578 0580	0577 0579 0581	0582

Figure 4.4: Expected Results - Parte 4

MS Windows Data Set

Case/String	Encoding	Active	Deleted	Unalloc
knives	ascii	0592 0594 0596	0593 0595 0597	0598
plan	ascii	0144 0146 0148	0145 0147 0149	0150
plans	ascii	0160 0162 0164	0161 0163 0165	0166
planned	ascii	0176 0178 0180	0177 0179 0181	0182
planner	ascii	0192 0194 0196	0193 0195 0197	0198
planning	ascii	0208 0210 0212	0209 0211 0213	0214
stealer	ascii	0736 0738 0740	0737 0739 0741	0742
stealing	ascii	0672 0674 0676	0673 0675 0677	0678
steals	ascii	0688 0690 0692	0689 0691 0693	0694
steal	Ascii	0656 0658 0660	0657 0659 0661	0662
FT-SS-10-Hex				
Panda	ascii	2944 2946 2948	2945 2947 2949	2950
FT-SS-10-Regex				
DireWolf	ascii	0896 0898 0900	0897 0899 0901	0902
WereWolf	ascii	0912 0914 0916	0913 0915 0917	0918

DONE

Figure 4.5: Expected Results - Parte 5

Chapter 5

Conclusioni

Il progetto di validazione condotto su *Autopsy 4.21* ha dimostrato che questo strumento forense è in grado di gestire efficacemente molte delle principali operazioni di ricerca di stringhe in un contesto di digital forensics. In particolare, la maggior parte delle stringhe nei file attivi e cancellati è stata individuata correttamente.

Tuttavia, sono emerse alcune limitazioni significative:

- L'assenza del supporto predefinito per la ricerca di numeri di *Social Security Number*.
- La mancanza di supporto per la ricerca basata su *stemming*, che ha limitato l'efficacia della ricerca di parole derivate.
- L'impossibilità di eseguire ricerche esadecimali.
- La difficoltà nel trovare stringhe Unicode in spazi non allocati.

Nonostante queste carenze, le funzionalità di ricerca basate su espressioni regolari e il supporto per più lingue si sono rivelati solidi.

Autopsy rappresenta quindi uno strumento flessibile e potente, che con alcuni miglioramenti, soprattutto per quanto riguarda la gestione delle stringhe Unicode normalizzate e le ricerche avanzate con operatori logici, potrebbe diventare ancora più efficace nelle indagini digitali future.