Capitolo 1

# Protezione delle fonti e di voi stessi

La maggior parte di noi non ne è molto consapevole, ma siamo tutti sotto sorveglianza. Le società di telecomunicazioni e i giganti delle tecnologie hanno accesso a un'enorme quantità di dati privati su chiunque utilizzi telefoni e Internet, dalla nostra esatta posizione fisica in qualsiasi momento al contenuto dei nostri messaggi e delle nostre e-mail, e possono condividere questi dati con gli investigatori. Anche quando i nostri dati privati non vengono inviati direttamente alle società che si occupano di tecnologie, i nostri dispositivi registrano comunque, localmente, ogni nostro movimento. Sapreste nominare ogni singola pagina web che avete visitato il mese scorso? Probabilmente il vostro browser web sì, così come i web tracker che seguono le vostre attività su Internet.

Oltre alla costante sorveglianza di base che tutti devono subire, coloro che, per lavoro, hanno accesso a dataset riservati spesso sono sottoposti a una sorveglianza aziendale o governativa ancora più rigorosa. I loro computer e telefoni di lavoro sono dotati di spyware preinstallato che monitora tutto ciò che fanno. Grandi database tengono traccia esattamente delle ricerche eseguite, dei termini di ricerca, di quali documenti aprono, scaricano o stampano e quando fanno tali operazioni.

È in questo ambiente che le persone comuni si ritrovano a diventare fonti. Nel corso del loro lavoro, sono testimoni di qualcosa di immorale o preoccupante. Potrebbero creare una cartella con documenti incriminanti, o salvare screenshot della chat aziendale, o fare qualche ricerca su database interni per saperne di più e assicurarsi che i loro sospetti siano fondati.

## In questo capitolo

Potrebbero inviarsi tramite e-mail alcuni documenti o copiare file su una chiavetta USB che collegano al computer di lavoro. Potrebbero mandare messaggi ad amici o familiari per chiedere consiglio mentre pensano al da farsi. Nella maggior parte dei casi le fonti non sono consapevoli dell'enorme traccia digitale che hanno già lasciato nel momento stesso in cui si sono rivolti a un giornalista o alle forze dell'ordine.
In questo capitolo imparerete a proteggere le fonti e i dataset che da esse ottenete. Esaminerò le considerazioni editoriali ed etiche legate alla redazione di documenti e al decidere quali informazioni pubblicare, nonché dove salvare i dataset in base alla loro riservatezza. Vi mostrerò come verificare che i dataset siano autentici, descrivendo come ho potuto procedere in passato per i dati hackerati da chi ha approfittato della pandemia di Covid-19 e per i registri delle chat di WikiLeaks. Verificare l'autenticità dei dataset è importante non solo per scrivere articoli accurati, ma anche per proteggere la vostra reputazione di giornalista. Infine, imparerete a utilizzare i gestori di password, a crittografare i dischi e a proteggervi dai documenti dannosi.

## Comunicazioni in sicurezza con le fonti

Poiché tutto ciò che facciamo lascia una traccia di dati, diventa complicato e difficile proteggere le fonti. Dopo aver pubblicato un report di successo sulla base di informazioni ottenute da un informatore anonimo, dovreste aspettarvi che l'obiettivo della vostra indagine avvii una propria indagine sull'identità della vostra fonte. L'equilibrio di poteri tra una fonte riservata e gli investigatori sulle sue tracce è estremamente asimmetrico. Se siete giornalisti o ricercatori e cercate di proteggere la vostra fonte, non sempre è sufficiente fare le cose "proprio per bene". Dal momento che gran parte della protezione della vostra fonte è fuori dal vostro controllo, è importante concentrarsi su quelle poche cose che potete controllare.
In qualità di giornalista o ricercatore, verificare che i dati che avete ottenuto siano autentici è una delle vostre principali responsabilità. Il modo più semplice per autenticare i documenti consiste nel chiedere all'azienda o al governo che li ha prodotti se sono effettivamente reali, ma questo percorso è irto di rischi per la vostra fonte. In alcuni casi, non volete fornire alcun dettaglio che possa rivelare l'identità della fonte. Ne parlerò approfonditamente più avanti in questo capitolo. Potreste anche non voler rivelare che siano trapelati determinati documenti, un argomento su cui scoprirete qualcosa di più in un paragrafo del capitolo dedicato all'oscuramento.
In questa parte del capitolo descriverò quali fonti corrono rischi e quali no, nonché le strategie per ridurre al minimo tali rischi. Discuterò anche delle differenze tra lavorare con fonti riservate che hanno accesso legittimo a informazioni riservate e hacker che infrangono la legge per ottenerle. È importante considerare attentamente l'impatto che le vostre scelte di investigazione potrebbero avere sulla vostra fonte, preferibilmente prima ancora di iniziare a parlarci.

### Lavorare con dati pubblici

Alcuni dataset non rappresentano alcun rischio per la fonte. Quando il governo pubblica una serie di documenti in risposta a una richiesta di registri pubblici o quando i documenti vengono resi pubblici nell'ambito di una causa legale, potete includere nel

vostro report tutti i dati che desiderate. Questi dati potrebbero contenere rivelazioni che qualche personaggio potente non vorrebbe divulgare, ma condividerle non metterà nessuno a rischio di ritorsioni, poiché i dati sono già pubblici.
Allo stesso modo, non dovete preoccuparvi della protezione della fonte per i dataset che potrebbero contenere dati riservati ma sono pubblici e ampiamente disponibili, come il dataset BlueLeaks che scaricherete nel Capitolo 2. Tutte le informazioni che potrete scoprire da quel dataset sono già state analizzate dagli investigatori dell'FBI che cercavano di determinare chi fosse l'hacker. In questi casi non è importante quante persone abbiano avuto accesso ai documenti. Non c'è alcuna possibilità di bruciare accidentalmente la vostra fonte fornendo troppi dettagli a un ufficio stampa governativo o aziendale quando chiederete se i dati che avete sono reali e se hanno una dichiarazione da fare in proposito. Poiché il dataset è già pubblico, qualsiasi sia il danno alla fonte, è già stato fatto.

## Proteggere le informazioni riservate

Se avete a che fare con un dataset proveniente da una fonte *riservata*, rivelare la sua identità potrebbe causarne il licenziamento, l'arresto o addirittura l'omicidio. Il passo più elementare che dovreste compiere per proteggere la vostra fonte è, semplicemente, non parlarne con nessuno che non collabori strettamente con voi nella vostra indagine. Non pubblicate sui social dettagli sulla vostra fonte che non intendete rendere pubblici, non parlatene agli amici alle feste e non parlatene nemmeno ai colleghi che non sono coinvolti nell'investigazione.
Se state interrogando un'azienda o un ente governativo in merito a un dataset trapelato che avete ottenuto, non fornite loro alcun dettaglio sulla vostra fonte riservata, anche se ve lo chiedono esplicitamente. Se venite arrestati e la polizia vi chiede di rivelare la vostra fonte, avete il diritto di rimanere in silenzio e potete certamente esercitarlo: non fornite alla polizia informazioni di cui non sia già in possesso. L'unico caso in cui siete obbligati a rivelare informazioni sulla vostra fonte è se un giudice ve lo ordina, e anche in quel caso potete opporvi.

## Ridurre al minimo le tracce digitali

Assicuratevi di lasciare la traccia digitale più piccola possibile quando comunicate con la vostra fonte. Per quanto possibile, evitate ogni comunicazione tramite e-mail, messaggi SMS, telefonate, messaggi nelle app dei social e così via. Non seguite la vostra fonte riservata sui social e assicuratevi che anche la fonte non vi segua.
Se dovete inviare messaggi o effettuare chiamate, utilizzate un'app di messaggistica crittografata, come Signal, di cui parlerò nel Capitolo 2, e assicuratevi che la vostra fonte elimini tutti i record delle sue chat con voi. Spesso dovrete registrare ciò che vi ha detto la vostra fonte per poterlo riferire, ma adottate misure per proteggere tali registrazioni, per esempio rimuovetele dalle app di messaggistica sul telefono e conservatele localmente sul vostro computer anziché in un servizio cloud. Se non avete più bisogno delle registrazioni delle vostre conversazioni dopo aver pubblicato il report, per esempio non intendete scrivere altri articoli sull'argomento, eliminatele.
Assicuratevi che la vostra fonte sappia di non dover cercare su Internet né voi né i report che avete pubblicato, in un modo che potrebbe essere considerato un legame sospetto. In

passato, contro le fonti è stata usata come prova la cronologia delle ricerche di Google. Per esempio, nel 2018, l'informatrice dipendente del Dipartimento del Tesoro Natalie Mayflower Sours Edwards è stata incriminata per aver presumibilmente fornito un dataset segreto al giornalista di BuzzFeed Jason Leopold. I documenti in questione hanno fatto trapelare i dettagli di transazioni finanziarie sospette riguardanti esponenti del Partito Repubblicano, membri importanti della campagna elettorale di Donald Trump del 2020, un agente russo legato al Cremlino e certi oligarchi russi. Durante l'indagine sulla fuga di notizie, l'FBI ha ottenuto un mandato di perquisizione per accedere alla cronologia delle sue ricerche su Internet, e nel suo atto d'accusa è riportata la ricerca di articoli basati sul contenuto della fuga di notizie a lei attribuita, poco dopo la loro divulgazione.

## Lavorare con hacker e informatori

I passaggi che dovete intraprendere per proteggere la vostra fonte variano notevolmente a seconda della sofisticazione tecnica della persona in questione. Non tutte le fonti sono *informatori*, persone dotate di un accesso interno a dataset o documenti e che, per motivi etici, divulgano prove di illeciti. A volte la vostra fonte potrebbe essere un *hacker attivista* ("hacktivisti") che intende mettere in ginocchio aziende o enti governativi che considerano inique.

A differenza della maggior parte degli informatori, gli hacker tendono a rendersi conto che sono sotto sorveglianza e che tutto ciò che fanno lascia una traccia digitale, quindi di solito adottano contromisure per nascondere le proprie tracce. È normale che gli informatori rivelino la propria identità ai giornalisti per motivi di verifica, anche se poi non vengono nominati pubblicamente, mentre gli hacker, in genere, rimangono completamente anonimi. Tuttavia, spesso gli hacker possono fornire informazioni tecniche utilizzabili per autenticare in modo indipendente un dataset, utilizzando l'intelligence open source. Come per qualsiasi fonte, non potete necessariamente fidarvi di ciò che vi dicono gli hacker, ma la loro esperienza può aiutarvi a verificare in modo indipendente che i dati che vi hanno fornito siano autentici. Per questi motivi, spesso il rischio per la vostra fonte è minore quando pubblicate documenti provenienti da hacker piuttosto che da informatori.

Quando comunicate con una fonte hacker, è importante restare fedeli al vostro ruolo di giornalista o ricercatore. Negli Stati Uniti, non state infrangendo alcuna legge se vi limitate a parlare con una fonte che è un hacker, ma *la vostra fonte quasi certamente sta infrangendo le leggi* hackerando aziende o governi e sottraendo dati. Non fate nulla che possa poi essere interpretato come una complicità con l'hacker. Per esempio, non chiedete loro di ottenere determinati dati per voi; lasciate che vi diano qualunque dato loro decidano di darvi. Se siete giornalisti e lavorate per una pubblicazione affermata, potreste cavarvela meglio contro le cause legali rispetto a un libero professionista. Anche se tutti dovrebbero essere protetti allo stesso modo dalla legge, le redazioni spesso dispongono di risorse come avvocati e fondi per la difesa legale. Quando temete che qualcosa che state facendo possa mettervi nei guai, consultate un avvocato.

A volte le fonti fingono di essere hacktivisti o informatori, ma in realtà sono hacker finanziati da uno Stato. Per esempio, gli hacker militari russi si sono spacciati per hacktivisti quando hanno violato la campagna presidenziale del Partito Democratico e di Hillary Clinton nel 2016, interferendo con le elezioni americane inviando dataset hackerati a WikiLeaks. Questo tipo di dataset potrebbe essere autentico e degno di nota, ma forse

non volete finire per diventare una pedina nella guerra dell'informazione di qualcun altro. Se non siete sicuri della credibilità della vostra fonte o se credete che essa potrebbe avere secondi fini, o se siete sicuri che siano disonesti con voi, è importante menzionare nel vostro report il vostro scetticismo a riguardo della vostra fonte e il motivo per cui avete dei dubbi. WikiLeaks ha fatto il contrario: ha insistito sul fatto che la sua fonte non era l'intelligence russa, mentre sapeva che era vero il contrario, e ha perfino diffuso la teoria complottista secondo la quale Seth Rich, un dipendente del Partito Democratico che era stato assassinato, fosse la vera fonte del gruppo, provocando problemi per anni contro i membri della sua famiglia.

## Archiviazione sicura dei dataset

Mentre vi preparate a ricevere un dataset da una fonte, valutate innanzitutto quanto ritenete che sia riservato quel dataset, poiché da questo dipende come dovrete proteggerlo e come continuerete a proteggere la vostra fonte. Ho già accennato al fatto che alcuni dataset sono completamente pubblici, mentre altri sono segreti di sicurezza nazionale, altamente classificati, e altri ancora si collocano a metà. Potreste incontrare dataset molto particolari, che non rientrano in una di queste categorie, ma in generale esistono tre diversi livelli di riservatezza: bassa, media e alta.

### Dataset a bassa riservatezza

Un dataset può avere una bassa riservatezza se soddisfa uno dei seguenti criteri.

- È già completamente pubblico, come nel caso di documenti in risposta a una richiesta di registri pubblici o di dataset pubblici che chiunque può scaricare da un collettivo per la trasparenza, come DDoSecrets (*Distributed Denial of Secrets*). Ne riparleremo nel Capitolo 2.
- Le forze dell'ordine o un'azienda concorrente hanno già ottenuto l'accesso al dataset, il che significa che il modo in cui lo archiviate non causerà ritorsioni contro la vostra fonte.
- Non contiene *informazioni di identificazione personale* (PII, *Personal Identifiable Information*) che descrivo in dettaglio più avanti nel capitolo.

Fondamentalmente, se non riuscite a pensare a nessun danno che potrebbe derivare se un certo dataset venisse condiviso più di quanto previsto, anche con le forze dell'ordine o gli investigatori, probabilmente si tratta di un dataset a bassa riservatezza.
È sicuro lavorare con dataset a bassa riservatezza anche nel *cloud*, ovvero con servizi di archiviazione come Google Drive, iCloud e Dropbox, servizi di hosting come AWS (*Amazon Web Services*) e qualsiasi altro servizio in cui chiunque, oltre a voi e alle persone con cui lavorate avrà accesso ai dati. Tuttavia, i servizi cloud sono tutti vulnerabili alle cause legali, quindi se state indagando su governi o aziende dotate di avvocati potenti, potrebbero ottenere mandati per i fornitori di servizi cloud, volti a ottenere i dati associati al vostro account. Inoltre, i dati archiviati nel cloud sono sicuri quanto il vostro stesso account. Assicuratevi di avere una password complessa e attivate funzionalità come l'autenticazione a due fattori per rendere il vostro account molto più resistente all'hacking.

## Dataset a media riservatezza

La maggior parte dei dataset che non hanno una bassa riservatezza ha una media riservatezza; non sono stati ancora resi pubblici, ma la loro sicurezza non richiede l'adozione di misure estreme. Per esempio, un dataset che descriverò più avanti in questo capitolo e che include le cartelle cliniche di centinaia di migliaia di pazienti, ha una media riservatezza. Questi dataset dovrebbero essere archiviati su dischi *crittografati* o protetti in modo tale che solo il proprietario possa sbloccarli per accedere ai dati. In questo modo, se il vostro laptop dovesse essere rubato, smarrito o sequestrato durante un'irruzione della polizia, nessuno potrà accedere ai vostri file. Se non avete già crittografato il vostro disco, lo farete negli Esercizi 1.1 e 1.2.

I dati a media riservatezza dovrebbero sempre rimanere sul disco del computer o su un disco rimovibile. Evitate di archiviarli in servizi cloud, a meno che non abbiate una buona ragione per farlo o non siate in grado di crittografarli in modo che il servizio cloud non abbia modo di decrittografarli. L'archiviazione di dataset su dischi locali crittografati riduce notevolmente il rischio che un estraneo possa accedervi senza autorizzazione.

Potete lavorare con dati a media riservatezza sul vostro computer da lavoro, purché proteggiate la vostra macchina. Ecco come fare.

- Assicuratevi che il disco del vostro computer sia crittografato.
- Adottate misure per proteggere fisicamente il computer. Assicuratevi che lo schermo si blocchi automaticamente dopo un breve periodo di inattività e che richieda una password per lo sblocco.
- Installate tempestivamente gli aggiornamenti software e fate attenzione ai programmi che installate e ai documenti che aprite sul vostro computer. Se eseguite accidentalmente del software dannoso o aprite un documento dannoso, qualcuno potrebbe hackerare il vostro computer e ottenere l'accesso ai vostri dataset.
- Archiviate il dataset su un disco USB esterno, che vi consente di archiviare più dati di quanti ne stanno sul vostro computer. Questo vi permette anche di viaggiare con il vostro laptop senza preoccuparvi di proteggere i dataset archiviati su di esso. Assicuratevi che anche tale disco esterno sia crittografato (ne parleremo più avanti nel capitolo).
- Non archiviate i file in cartelle del computer che vengono caricate automaticamente nel cloud. Per esempio, molti utenti Mac configurano il proprio computer per caricare la cartella *Documenti* su iCloud, il servizio di archiviazione cloud di Apple. Se il vostro computer è configurato in questo modo, non inserire in tale cartella i file relativi alle vostre indagini.

In generale, lavorate solo *localmente* con dati a media riservatezza, ovvero con quei file archiviati sul disco del vostro computer e che non siano già stati divulgati su alcun servizio online. In alcuni casi è ragionevole lavorare da remoto anche con dataset a media riservatezza. Se lavorate con altre persone, potrebbe essere necessario utilizzare una soluzione crittografata di condivisione dei file, in modo che il servizio che state utilizzando non possa decrittografare i file, e che solo voi e i vostri colleghi possiate farlo. Un'opzione semplice consiste nell'inviarsi file utilizzando l'app di messaggistica Signal. Se poi voi o l'azienda per la quale lavorate ospitate uno strumento sicuro per la ricerca di dataset, come Aleph (trattato nel Capitolo 5), è ragionevole copiare i dati utilizzando tale strumento.

Tutti i dataset con cui lavorerete in questo libro sono a bassa riservatezza, poiché sono già pubblici. Tuttavia, le tecniche che apprenderete si applicheranno anche a dataset a media riservatezza, poiché lavorerete con i dati localmente sul vostro computer. Anche se è accettabile lavorare con questi particolari dataset nel cloud, il fatto di imparare a lavorarci localmente vi fornirà le tecniche necessarie per impiegare dataset a riservatezza più elevata.

## Dataset ad alta riservatezza

I dataset ad alta riservatezza sono di gran lunga i più difficili da manipolare, e questo per una buona ragione. Lo Snowden Archive, per esempio, è ad alta riservatezza. Ho dedicato anni a scrivere report su questa enorme quantità di documenti governativi segreti forniti dall'informatore della NSA (*National Security Agency*) Edward Snowden, che ha rivelato il fatto che gli enti di spionaggio statunitensi e NATO stavano conducendo sorveglianze senza mandato e violazioni della privacy su una scala inimmaginabile. Non volevamo che l'FBI o la NSA vi potessero accedere, il che rendeva fuori questione l'utilizzo di servizi cloud, ma, cosa forse ancora più importante, non volevamo che vi accedessero nemmeno i servizi di intelligence stranieri. Partivamo dal presupposto che gli hacker governativi avessero la capacità di hackerare da remoto praticamente qualsiasi computer potessimo utilizzare, a meno che non prendessimo provvedimenti per assicurarci che tale computer non si connettesse mai a nessuna rete remota.

Scendere nei dettagli su come condurre indagini ad alta riservatezza va oltre gli scopi di questo libro e, comunque, non avrete bisogno di questo livello di competenze per affrontare i prossimi capitoli. Tuttavia, come riferimento per il futuro, questa parte del capitolo descrive come dovreste procedere se vi ritrovate a lavorare con una cache di documenti top-secret.

Se un dataset ha un'alta riservatezza, finché non siete in prossimità della pubblicazione del report, archiviatelo o fatevi accesso solo utilizzando computer dotato di *air gap*, ovvero senza mai alcuna connessione a Internet. Trasferite i file dal computer isolato con air gap solo quando sono già stati redatti e sono pronti per la pubblicazione. In breve: acquistate un computer nuovo, non collegatelo mai a Internet e usate solo quello. Oppure, se avete un vecchio computer adatto, formattate il suo disco, reinstallate il sistema operativo e utilizzate quel computer senza mai collegarlo a Internet. Questi passaggi vi aiuteranno ad assicurarvi di iniziare con un sistema "pulito", privo di tracker o malware. Per renderlo ancora più sicuro, aprite il case del computer e rimuovete fisicamente la schedina hardware del wireless.

Affronterete parecchi problemi per rendere possibile il trasferimento dei dati tra il vostro computer isolato con air gap e il vostro computer di lavoro: per esempio, l'installazione o l'aggiornamento del software sul vostro computer isolato con air gap richiederà il download dei file su un altro computer, una verifica minuziosa che si tratti solo di software legittimo, quindi il trasferimento sul computer isolato con air gap per l'installazione. Tuttavia, vale la pena di svolgere questi passaggi aggiuntivi, quando una loro violazione potrebbe avere conseguenze gravi.

È molto importante anche che il disco del vostro computer isolato con air gap e tutti i dischi USB che gli collegate siano crittografati con passphrase complesse. Considerate anche la sicurezza fisica del luogo in cui conservate il computer isolato con air gap e i dischi USB. Se possibile, conservateli in una cassaforte o in un caveau con una buona

serratura. Se ciò non fosse possibile, almeno teneteli in una stanza chiusa, della quale solo poche persone possiedono le chiavi. Spegnete sempre il computer isolato con air gap quando non lo usate, per complicare gli attacchi contro la crittografia del disco.
Quando lavorate sul computer isolato con air gap, fate attenzione ai dispositivi elettronici connessi a Internet nelle vicinanze, dotati di microfoni o fotocamere. Evitate di tenere conversazioni relative a documenti altamente riservati a portata d'orecchio dei microfoni e considerate se vi è la possibilità che eventuali fotocamere vicine (inclusi, naturalmente, gli smartphone) possano catturare fotografie del vostro schermo.

## Autenticazione dei dataset

Non potete credere a tutto ciò che leggete su Internet, e i documenti o i dataset "interessanti" che vi inviano persone anonime non fanno eccezione. La disinformazione è pervasiva. Nel report pubblicato è importante spiegare, almeno brevemente, che cosa vi rende tanto sicuri dei dati. Se non potete autenticarli ma desiderate comunque pubblicare il vostro report nel caso in cui i dati fossero reali o che qualcuno *possa* autenticarli, chiaritelo. In caso di dubbi, se proprio dovete sbagliare, fatelo per eccesso di trasparenza.
Il modo in cui si può verificare che un dataset sia autentico dipende assolutamente dalla natura dei dati. Bisogna affrontare il problema caso per caso. Il modo migliore per verificare un dataset consiste nell'utilizzare l'*intelligence open source* (OSINT) o le informazioni disponibili al pubblico che chiunque abbia sufficiente competenza può trovare. Ciò potrebbe significare esplorare gli account dei social media, consultare la Wayback Machine di Internet Archive (`https://web.archive.org`), ispezionare i metadati delle immagini o dei documenti pubblici, pagare servizi per i dati storici di registrazione dei nomi di dominio o consultare altri tipi di registri pubblici. Se il vostro dataset include un database tratto da un sito web, per esempio, potreste dover confrontare le informazioni contenute in quel database con le informazioni disponibili pubblicamente sul sito web, per verificare che corrispondano.
La trattazione di OSINT in questo libro si concentra su come l'ho utilizzato nelle mie ricerche. Se volete saperne di più, consultate *OSINT Techniques: Resources for Uncovering Online Information*, di Michael Bazzell, insieme agli strumenti complementari elencati su `https://inteltechniques.com/tools`. Bazzell descrive un gran numero di strumenti e tecniche per scoprire dettagli che potrebbero aiutarvi a verificare i dataset utilizzando OSINT.
In questa parte del capitolo condividerò due esempi di autenticazione dei dati tratti dalla mia esperienza: uno su un dataset del gruppo contrario alle vaccinazioni America's Frontline Doctors e un altro sui registri delle chat trapelati da un gruppo Twitter di WikiLeaks.

> **NOTA**
>
> Nel 2023, Twitter ha cambiato nome, diventando "X". In questo libro continuerò a fare riferimento agli account e ai post esistenti prima di questo rebranding come "account Twitter" e "tweet".

## Il dataset AFLDS

A fine 2021, nel bel mezzo della pandemia di Covid-19, un hacker anonimo mi ha inviato centinaia di migliaia di cartelle cliniche e prescrizioni di società di telemedicina che collaborano con AFLDS (*America's Frontline Doctors*). Si tratta di un gruppo di destra contrario ai vaccini che inganna le persone sulla sicurezza del vaccino Covid-19 e induce i pazienti a spendere milioni di dollari per farmaci come l'ivermectina e l'idrossiclorochina, che sono inefficaci nel prevenire o curare il virus. Il gruppo è stato inizialmente formato per aiutare la campagna di rielezione di Donald Trump nel 2020 e la leader del gruppo, Simone Gold, è stata arrestata per aver preso d'assalto il Campidoglio degli Stati Uniti il 6 gennaio 2021. Nel 2022, ha scontato due mesi di carcere per il suo ruolo nell'attacco.
La mia fonte mi ha detto di aver ottenuto i dati scrivendo un programma che ha inviato migliaia di richieste a un sito web gestito da una delle società di telemedicina, Cadence Health. Ogni richiesta ha restituito dati su un paziente differente. Per verificare l'affermazione, ho creato io stesso un account sul sito web di Cadence Health. Mi sembrava tutto legittimo. Le informazioni che avevo in mano su ciascuno dei 255.000 pazienti erano esattamente le stesse che mi è stato chiesto di fornire quando ho creato il mio account sul servizio e vari nomi e ID di categorie nel dataset corrispondevano a ciò che potevo vedere sul sito web. Ma come potevo essere sicuro che i dati dei pazienti stessi fossero reali, che queste persone non fossero, semplicemente, inventate?
Ho scritto un semplice script Python per scorrere i 72.000 pazienti (coloro che avevano pagato per ricevere un'assistenza sanitaria falsa) e ho inserito ciascuno dei loro indirizzi e-mail in un file di testo. Poi ho incrociato questi indirizzi e-mail con un dataset completamente distinto contenente le informazioni personali dei membri di Gab, un social network popolare tra utenti fascisti, attivisti anti-democratici e anti-vaccini. A inizio 2021, un hacktivista conosciuto con il nome di "JaXpArO and My Little Anonymous Revival Project" aveva violato Gab e si era impossessato di 65 GB di dati, inclusi circa 38.000 indirizzi e-mail dei suoi utenti. Pensando che potesse esserci una sovrapposizione tra gli utenti AFLDS e Gab, ho scritto un altro semplice programma Python che confrontava gli indirizzi e-mail di ciascun gruppo e mi mostrava tutti gli indirizzi presenti in entrambi gli elenchi. Ce n'erano diversi.
Armato di queste informazioni, ho iniziato a esplorare le timeline pubbliche di Gab relativamente agli utenti i cui indirizzi e-mail erano presenti in entrambi i dataset, alla ricerca di post su AFLDS. Utilizzando questa tecnica, ho trovato diversi pazienti di AFLDS che hanno pubblicato la loro esperienza su Gab, cosa che mi ha portato a credere che i dati fossero autentici. Per esempio, secondo le note sul consulto medico del dataset hackerato, un paziente ha creato un account sul sito di telemedicina e quattro giorni dopo ha ricevuto un teleconsulto. Circa un mese dopo, ha scritto a Gab dicendo: "Quelli di Frontline Doctors sono finalmente riusciti a fornirmi HCQ/Zinco" (HCQ è l'abbreviazione di idrossiclorochina).
Il Capitolo 13 è dedicato interamente alla mia indagine su AFLDS e descrive in modo più approfondito i dettagli tecnici del mio script Python. Dopo aver letto i prossimi capitoli, avrete acquisito le conoscenze di programmazione Python necessarie per capire come ha funzionato lo script.

## La chat del gruppo Twitter di WikiLeaks

A fine 2017, la giornalista Julia Ioffe ha pubblicato una rivelazione su "Atlantic": WikiLeaks era entrata nei messaggi su Twitter di Donald Trump Jr. Tra le altre cose, prima delle elezioni del 2016, WikiLeaks aveva suggerito a Trump Jr. che, anche se suo padre avesse perso le elezioni, non avrebbe dovuto arrendersi. "Ciao Don", ha scritto l'account Twitter verificato `@WikiLeaks`, "se tuo padre 'perderà' pensiamo che sia molto più interessante che NON lo ammetta [*sic*] e che si impegni a DENUNCIARE i media e i brogli che si sono verificati, come ha lasciato intendere che potrebbe fare".

A metà del 2015, un volontario di lunga data di WikiLeaks, conosciuto con lo pseudonimo di Hazelpress, ha avviato un gruppo Twitter privato con WikiLeaks e con i suoi principali sostenitori. Dopo aver visto il suo gruppo prendere una piega sempre più di destra, cospiratoria e immorale, e in particolare dopo aver appreso dei messaggi segreti di WikiLeaks con Trump Jr., Hazelpress ha deciso di denunciare il suo stesso gruppo di denuncia. Da allora si è fatta avanti pubblicamente come Mary-Emma Holly, un'artista che aveva dedicato anni a un'attività di ricercatrice legale volontaria per WikiLeaks.

Per portare avanti la fuga di notizie di WikiLeaks, Holly ha effettuato l'accesso al suo account Twitter, lo ha reso privato, ha smesso di seguire tutti e ha cancellato tutti i suoi tweet. Ha anche cancellato tutti i suoi messaggi, a eccezione di quelli del gruppo Twitter privato WikiLeaks e ha cambiato il suo nome-utente su Twitter. Utilizzando il browser web Firefox, è poi andata alla conversazione, che conteneva 11.000 messaggi ed era andata avanti per due anni e mezzo, e ha letto gli ultimi messaggi del gruppo. Ha fatto scorrere i messaggi, ha aspettato che Twitter caricasse altri messaggi, ha fatto scorrere di nuovo verso l'alto e ha continuato a farlo per *quattro ore*, finché non ha raggiunto il primo messaggio del gruppo. Poi ha utilizzato l'opzione *Salva pagina con nome* di Firefox per salvare una versione HTML della pagina web, nonché una cartella piena di risorse, come le immagini pubblicate nel gruppo.

Ora che aveva una copia locale e offline di tutti i messaggi contenuti nel gruppo, Holly l'ha fatta trapelare ai media. A inizio 2018, ha inviato un messaggio Signal al numero di telefono elencato nella pagina dei suggerimenti di *The Intercept*. A quel tempo, ero io a controllare su Signal i suggerimenti in arrivo. Utilizzando OnionShare, un software che ho sviluppato per questo scopo e che descrivo in dettaglio nel Capitolo 2, mi ha inviato un file crittografato e compresso, insieme alla password per decrittografarlo. Dopo averlo estratto, ho trovato un file HTML da 37 MB, così grande che il mio browser web non rispondeva quando ho provato ad aprirlo e che in seguito ho diviso in più file, per facilitarmi il lavoro, e una cartella con 82 MB di risorse.

Come potevo verificare l'autenticità di un file HTML così grande? Se avessi potuto, in qualche modo, accedere a quegli stessi dati direttamente dai server di Twitter, sarebbe stato possibile; solo un insider di Twitter sarebbe stato in grado di creare falsi messaggi da visualizzare sul sito web di Twitter, e anche questo sarebbe estremamente impegnativo. Quando l'ho spiegato a Holly (che all'epoca conoscevo ancora solo come Hazelpress), mi ha dato il suo nome-utente e la sua password Twitter. Aveva già cancellato tutte le altre informazioni da quell'account. Con il suo consenso, ho effettuato l'accesso a Twitter con le sue credenziali, sono andato ai suoi messaggi e ho trovato il gruppo Twitter in questione. A prima vista sembrava che contenesse gli stessi messaggi presenti nel file HTML e ho confermato che l'account verificato `@WikiLeaks` pubblicava spesso sul gruppo.

Seguire questi passaggi mi ha reso estremamente fiducioso sull'autenticità del dataset, ma ho deciso di fare un ulteriore passo avanti di verifica. Potevo scaricare io stesso una copia distinta del gruppo Twitter per confrontarla con la versione inviatami da Holly? Ho cercato in giro e ho trovato DMArchiver, un programma Python che poteva fare proprio questo. Utilizzando questo programma, insieme al nome-utente e alla password di Holly, ho scaricato una versione testuale di tutti i messaggi contenuti nel gruppo Twitter. Ci sono voluti solo pochi minuti per eseguire questo strumento, anziché le quattro ore di scrolling verso l'alto in un browser web.

> **NOTA**
>
> Dopo questa indagine, il programma DMArchiver ha smesso di funzionare a causa di alcune modifiche apportate a Twitter, e oggi il progetto è stato abbandonato. Tuttavia, se vi trovate ad affrontare una difficoltà simile in un'indagine futura, cercate uno strumento che possa aiutarvi. Potreste anche prendere in considerazione l'idea di svilupparne uno vostro, utilizzando le competenze di programmazione che apprenderete nei Capitoli 7 e 8.

L'output di DMArchiver, un file di testo da 1,7 MB, era molto più semplice da utilizzare rispetto all'enorme file HTML e includeva anche i timestamp esatti. Ecco un frammento della versione testuale:

```
[2015-11-19 13:46:39] <WikiLeaks> We believe it would be much better for GOP to win.
[2015-11-19 13:47:28] <WikiLeaks> Dems+Media+liberals woudl then form a block to reign ↵
    in their worst qualities.
[2015-11-19 13:48:22] <WikiLeaks> With Hillary in charge, GOP will be pushing ↵
    for her worst qualities., dems+media+neoliberals will be mute.
[2015-11-19 13:50:18] <WikiLeaks> She's a bright, well connected, sadistic sociopath.
```

Ho potuto visualizzare la versione HTML in un browser web, per vederla esattamente come appariva originariamente su Twitter, il che è stato utile anche per acquisire le schermate da includere nel nostro report finale, come mostrato nella Figura 1.1.


**Figura 1.1** Uno screenshot del file HTML trapelato.

Insieme alla talentuosa giornalista Cora Currier, ho iniziato il lungo processo di lettura di tutti gli 11.000 messaggi della chat, prestando la massima attenzione al 10% di essi, provenienti dall'account `@WikiLeaks` – presumibilmente controllato da Julian Assange, editore di WikiLeaks – e selezionando tutto nell'interesse pubblico. Abbiamo scoperto i seguenti dettagli.

- Assange ha espresso il desiderio che i Repubblicani vincessero le elezioni presidenziali del 2016.
- Assange e i suoi sostenitori erano intensamente impegnati nello screditare due donne svedesi che lo avevano accusato di stupro e molestie, oltre a screditare i loro avvocati. Assange e i suoi difensori hanno dedicato settimane a discutere su come sabotare gli articoli sul caso di stupro di cui le giornaliste femministe stavano scrivendo.
- Assange ha cercato di screditare la regista Laura Poitras per il modo in cui lo ha ritratto in *Risk*, il documentario del 2016 su WikiLeaks. Il film include una scena in cui Assange dice al suo avvocato che i suoi accusatori facevano parte di una "formazione politica femminista radicale assolutamente di bassa lega", e in un'altra scena dice: "Parte del problema in questo caso è che ci sono due donne, e il pubblico non riesce a tenerle separate. Se fosse una sola, si potrebbe dire: 'È una fuori di testa'. E penso che ormai sarebbe già successo".
- Assange ha usato un linguaggio transfobico e misogino quando ha parlato di Chelsea Manning, la sua fonte dal 2010, e delle sue amiche. Discuterò ulteriormente la relazione di Manning con WikiLeaks nel Capitolo 2.
- Dopo che il giornalista di Associated Press Raphael Satter ha scritto un articolo sui danni causati dalla pubblicazione di informazioni personali identificabili da WikiLeaks, Assange lo ha definito un "ratto", un "ebreo impegnato nella questione ((()))", riferendosi a un meme antisemita neonazista. Poi ha detto ai suoi sostenitori: "Infangatelo. Convincetelo a esibire i suoi pregiudizi".

Potete leggere il nostro report su questo dataset su `https://theintercept.com/2018/02/14/julian-assange-wikileaks-election-clinton-trump`. Dopo che *The Intercept* ha pubblicato questo articolo, Assange e i suoi sostenitori mi hanno preso di mira personalmente con insulti antisemiti, e Russia Today, la stazione televisiva statale, ha trasmesso un servizio su di me. Parlerò di WikiLeaks e della sua storia in modo più approfondito nel Capitolo 2. Le tecniche che potete utilizzare per autenticare i dataset variano notevolmente a seconda della situazione. A volte potete fare affidamento su OSINT, altre volte potete contare sull'aiuto della vostra fonte e talvolta dovrete trovare un metodo completamente differente.

## Oscuramento

Una volta autenticato il dataset, dovete considerare se e come desiderate oscurare, ovvero nascondere o eliminare, le informazioni riservate prima di pubblicare i risultati della vostra indagine. In alcuni casi potrebbe essere sicuro pubblicare i documenti originali senza alcun oscuramento, mentre in altri potreste scegliere di non pubblicare alcun documento. In questa parte del capitolo discuterò di come prendere queste decisioni e i motivi per cui potreste scegliere di oscurare o non oscurare le informazioni.

## Quali dati pubblicare

Nel decidere quanti dati pubblicare, considerate se il vostro metodo di segnalazione delle rivelazioni consentirà agli investigatori di risalire alla vostra fonte. Per esempio, se il dipartimento Risorse umane di un'azienda invia un'e-mail a tutti i suoi 10.000 dipendenti e uno di loro vi fa trapelare il messaggio, sarà molto difficile per l'azienda trovare il colpevole. Ma se solo 10 persone hanno accesso a un certo documento, o se i file log del database mostrano un elenco di 10 persone che vi hanno avuto accesso di recente, l'azienda ha un vero elenco di sospetti su cui lavorare.
Quante persone hanno accesso ai dati che avete ottenuto, quanto sono riservati, che cosa sta rischiando la vostra fonte e quanto si sente tranquilla sono tutti fattori che determinano i diversi metodi o le diverse quantità di dati da pubblicare. L'elenco seguente presenta le opzioni da considerare, ordinate dal rischio maggiore a quello minore per la fonte.

- Pubblicare i documenti o i dataset inalterati.
- Pubblicare i documenti dopo averli oscurati e privati dei metadati.
- Pubblicare i documenti dopo averli ricreati da zero digitandoli a mano in nuovi documenti separati e pubblicando quelli. Quando ricreate i documenti, rimuovete eventuali tracker nascosti e rendete impossibile capire dai documenti stessi se la vostra fonte li ha ottenuti fotografando lo schermo, copiandoli su una chiavetta USB, caricandoli su un sito web o utilizzando qualche altro metodo.
- Non pubblicare affatto i documenti, limitandovi a descriverli e citarli.
- Non citare nemmeno i documenti, descrivendo semplicemente le rivelazioni che contengono. Se gli investigatori non sanno quali documenti sono stati violati, ma sanno solo che una notizia accurata rivela in qualche modo informazioni riservate, avranno più difficoltà a fare progressi nelle loro indagini.

Il fatto di pubblicare i documenti è più trasparente per i vostri lettori; fornendo prove dirette renderete il vostro lavoro più credibile, ma ciò deve essere valutato rispetto alla protezione della vostra fonte. Prenderete queste decisioni caso per caso, ma tenete sempre presente i rischi che fate correre alla vostra fonte.

## Che cosa oscurare

Se avete considerato attentamente i rischi per la vostra fonte e avete deciso di pubblicare i documenti invece di limitarvi a descriverli, il passo successivo è decidere quali informazioni contenute in quei documenti oscurare, eventualmente, prima della pubblicazione. Ci sono tre ragioni per procedere all'oscuramento: continuare a proteggere la fonte, proteggere la privacy delle altre persone coinvolte o proteggere informazioni governative o aziendali che dovrebbero giustamente rimanere segrete.

### Proteggere la fonte

Se il vostro dataset include salvataggi di un sito web privato o di un database cui la vostra fonte ha fatto accesso, vi consigliamo di oscurare il nome-utente o qualsiasi altra informazione identificativa prima della pubblicazione. Inoltre, assicuratevi di non pubblicare accidentalmente metadati che potrebbero rivelare la vostra fonte. Questo libro non descrive i molti modi in cui ciò potrebbe accadere, ma ecco due esempi comuni:

spesso i documenti Word includono il nome dell'autore e spesso le foto includono le coordinate GPS e il tipo di fotocamera utilizzata.
Nel 2012, John McAfee, il controverso dirigente milionario della software-house, era in fuga. La polizia ha fatto irruzione nella sua casa in Belize (ex Honduras Britannico) e lui è fuggito dal Paese. In un post sul blog, ha scritto: "Al momento sono al sicuro e in compagnia di due intrepidi giornalisti [*sic*] di Vice Magazine... Non siamo in Belize, ma non siamo ancora del tutto fuori pericolo". Quel giorno, "Vice" pubblicò il suo articolo su McAfee, che includeva una fotografia. Secondo i metadati della foto, era stata scattata con un iPhone 4S e includeva le coordinate GPS di un'abitazione in Guatemala. Non rimuovendo i metadati dalla foto, "Vice" ha pubblicato accidentalmente la sua posizione esatta. Se solo "Vice" avesse fatto uno screenshot dell'immagine e l'avesse pubblicato, avrebbe cancellato i metadati e mantenuto segreta la posizione.
Nel 2017, quando il presidente Donald Trump definì "fake news" le accuse secondo le quali la Russia avrebbe interferito nelle elezioni americane, l'informatore della NSA Reality Winner inviò a *The Intercept* in forma anonima un documento top secret con la prova che la NSA aveva, in effetti, assistito a un attacco informatico dei russi contro i funzionari elettorali locali. *The Intercept* pubblicò il documento, e poco dopo Reality Winner venne arrestato. Il documento pubblicato includeva un tipo di metadati chiamato *printer dots*, punti gialli quasi invisibili che le stampanti aggiungono alla stampata e che includono il numero di serie della stampante e il timestamp della stampa. Sebbene non ci siano prove che gli investigatori li abbiano notati per condurre all'arresto di Reality Winner (era una delle sei persone che avevano stampato questo documento e l'unica che aveva scritto un'e-mail a *The Intercept*), i *printer dots* avrebbero potuto rivelare la fonte. *The Intercept* avrebbe potuto disinnescare questo problema ricreando il documento (ridigitandolo e ricreando la grafica) e pubblicando questa versione, invece della versione scansionata dell'originale.

## Proteggere le informazioni personali contenute nei dataset

Molti dataset includono nomi, indirizzi e-mail, nomi utente, numeri di telefono, indirizzi di casa, password e altre informazioni personali di persone non in vista. Molti documenti governativi e aziendali includono informazioni personali di semplici dipendenti che non aggiungerebbero nulla al vostro articolo ma potrebbero rendere queste persone bersaglio di molestie. Anche quando si ha a che fare con personaggi pubblici, nella maggior parte dei casi è comunque opportuno oscurare le loro informazioni personali, a meno che la loro pubblicazione sia fondamentale per il report. Per esempio, se il fulcro della vostra indagine fosse una lussuosa villa di proprietà di un miliardario, potrebbe essere ragionevole pubblicarne l'indirizzo. Ma se state scrivendo un articolo non correlato a quel miliardario, non c'è alcun motivo di specificare il suo indirizzo di casa.
Anche se ritenete che gli obiettivi della vostra indagine siano perfetti idioti, è meglio oscurare le loro informazioni personali se la loro inclusione non aggiunge nulla al vostro report. Anche gli idioti hanno diritto alla loro privacy e il fatto stesso di aver pubblicato inutilmente informazioni personali potrebbe essere utilizzato per screditare la vostra segnalazione, indipendentemente dalle rivelazioni in essa contenute.
L'eccezione a questa regola si ha quando il fatto di rivelare pubblicamente certe informazioni è una parte importante del vostro articolo e potrebbe tenere al sicuro altre persone. Per esempio, è etico nominare qualcuno che commette abusi sul posto di lavoro o in un certo settore di attività o denunciare qualcuno in quanto membro di un gruppo

che incita all'odio. Ma anche quando fate *coming out* pubblicamente relativamente a qualcuno, non pubblicate informazioni personali non necessarie, come il suo indirizzo di casa. Potreste essere accusati di molestie, il che potrebbe distrarre dall'illecito che state cercando di denunciare.

### Proteggere segreti legittimi

Occasionalmente, i governi e le aziende hanno ragioni legittime per mantenere certi segreti. Per esperienza, questa situazione è rara: il governo degli Stati Uniti ha un grave problema di sovraclassificazione. Questo è uno dei motivi per cui è importante chiedere alle parti coinvolte un commento prima di pubblicare il vostro articolo: un ente governativo o un'azienda potrebbe fornirvi un contesto, che potrebbe farvi decidere di non rendere pubblici i dati che avete raccolto. Per esempio, una volta ho preso parte alla decisione di oscurare i dettagli di un documento top-secret del governo statunitense relativo al programma di armi nucleari di un altro paese.

## Effettuare richieste di commento

Date sempre alle persone o alle aziende di cui state scrivendo la possibilità di raccontare la loro versione dei fatti. Anche se siete convinti che non risponderanno in modo sincero o non risponderanno affatto, dovreste comunque provare a contattarli, spiegare che cosa pubblicherete e dare loro la possibilità di difendersi. Se rispondono, citate la loro risposta nel report che pubblicate (e se sapete che non dicono la verità, spiegatelo insieme alla loro citazione). Se non rispondono o si rifiutano di commentare, includete anche questo dettaglio nel vostro report.

Per esempio, nel 2017, ho segnalato la fuga di registri delle chat di neonazisti, di cui parlerò nel Capitolo 14. Nel mio articolo, ho nominato un membro del gruppo di odio favorevole alla schiavitù *League of the South* che è stato arrestato durante la sanguinosa battaglia condotta da *Unite the Right* a Charlottesville, Virginia, per aver portato con sé una pistola nascosta. Aveva postato messaggi in una chat dicendo che aveva "conti in sospeso" con gli antifascisti locali, perché lo avevano fatto licenziare. Usando gli elenchi pubblici, sono risalito al suo numero di telefono. Ho impostato un numero di telefono virtuale utilizzando Google Voice, dal momento che non volevo dargli il mio numero privato, e l'ho chiamato. Gli ho lasciato messaggi, ma non mi ha mai risposto.

Se la vostra indagine prevede un contraddittorio, ovvero se le persone sulle quali state indagando non ne saranno contente, aspettate un po' prima di pubblicare il vostro report, per contattarle e dare loro una possibilità di replica. È educato concedere loro almeno 24 ore per rispondere, un tempo non sufficiente per sabotare il vostro articolo. Potrebbero far trapelare il vostro articolo a una pubblicazione loro favorevole, perché la pubblichi con un'accezione positiva, o annunciare ai loro follower che sta arrivando un pezzo problematico o tentare di utilizzare altri mezzi legali per impedirvi di pubblicarlo. Sono stato coinvolto in indagini in cui si sono verificati tutti questi scenari.

È probabile che non siate esperti di tutti gli aspetti di cui state parlando; quindi, spesso è opportuno consultare esperti esterni (professori universitari, autori, scienziati e così via) e includere le loro citazioni nei vostri report. Nel mio reportage ho intervistato professori di crittografia, esperti di disinformazione, medici e difensori dei diritti civili che lavorano per organizzazioni no-profit. Anche se siete esperti dell'argomento della

vostra indagine, fornire voci esterne spesso aggiunge valore al vostro articolo, cosa che vi aiuterà a sostenere in modo più convincente le vostre argomentazioni.
Se vi fidate degli esperti con cui state parlando, contattateli nelle prime fasi del processo di segnalazione. È anche prassi condividere con loro i documenti riservati, purché si impegnino a mantenerli segreti fino alla pubblicazione. Nel caso di documenti altamente riservati, potrebbe essere necessario organizzare di persona la visita di esperti esterni e la visualizzazione dei file solo sul computer isolato con air gap. A volte questi esperti possono indirizzarvi in direzioni di ricerca che non vi sarebbero venute in mente.
Ora che avete visto come proteggere le vostre fonti e autenticare le informazioni che esse vi forniscono, esaminiamo alcuni modi per proteggere il vostro computer e i vostri account online, per mantenere al sicuro i vostri dataset e altri record riservati.

## Gestori di password

La maggior parte delle persone non usa password univoche, il che significa che le riutilizza in più posti. Questa è una pessima idea, poiché qualsiasi password duplicata è sicura tanto quanto il luogo meno sicuro in cui l'avete utilizzata. Andate su `https://haveibeenpwned.com`, cercate il vostro indirizzo e-mail o numero di telefono e vedrete un elenco di violazioni dei dati che avete subito. Se la vostra password LinkedIn è stata rivelata in una violazione dei dati alcuni anni fa ma è la stessa che usate per il vostro account Gmail, per accedere al vostro laptop o per sbloccare il vostro disco USB crittografato pieno di dataset riservati, potreste essere nei guai.
La soluzione consiste nel rendere tutte le vostre password uniche e forti, il che significa semplicemente abbastanza lunghe e casuali da renderle impossibili da prevedere. Sfortunatamente, le password complesse sono difficili da memorizzare ed è impossibile per gli esseri umani memorizzare centinaia di password complesse e uniche. Eppure, ci viene richiesto di utilizzare centinaia di password nella nostra vita quotidiana.
Fortunatamente, possiamo fare in modo che i computer memorizzino la maggior parte delle nostre password per noi. I gestori di password impiegano un database crittografato di password che si sblocca con una master password, l'unica che dovrete memorizzare. I gestori di password spesso vi consentono di sincronizzare il vostro database di password sul cloud, il che va bene purché utilizziate una master password complessa. Anche se un hacker dovesse sottrarre il vostro database di password crittografate o se la società proprietaria del gestore di password consegnasse i vostri dati all'FBI o ad altre autorità, essi non saranno in grado di sbloccarlo senza la vostra master password. Un database di password crittografate è completamente inaccessibile a chiunque, senza la master password. Se la vostra master password è complessa, sarà letteralmente impossibile per loro indovinarla e le altre vostre password saranno al sicuro. La crittografia è davvero fantastica.

### La password di Twitter di Donald Trump

Ho appreso da un episodio dell'eccellente podcast *Darknet Diaries*, condotto da Jack Rhysider, che la password LinkedIn di Donald Trump è stata scoperta in una violazione dei dati nel 2012. La sua password, `yourefired`, era la sua esclamazione tipica in *The Apprentice*, il reality show da lui condotto. Mentre era candidato alla presidenza nel 2016, tre hacker olandesi, Victor, Edwin e Matt, che fanno

parte di un gruppo chiamato Guild of the Grumpy Old Hackers, hanno scoperto la sua password LinkedIn nel dataset di quella violazione. L'hanno provata sull'account Twitter `@realDonaldTrump` di Trump e… ha funzionato.

Potreste pensare: "Utilizzare un gestore di password non significa però mettere tutte le uova nello stesso paniere? Se viene violato, l'hacker avrà accesso a *tutto quanto*!" Questo è vero: è molto importante proteggere il vostro gestore di password, ma non usarne uno è come cercare di tenere centinaia di uova tutte in mano, senza usare un cestino e senza romperne nessuna. Alla fine, perdereste molte "uova". Inoltre, avete sempre la possibilità di utilizzare più gestori di password (più cestini) per ambiti differenti: se anche uno dovesse essere violato, gli altri rimarranno al sicuro.
Sono disponibili diversi buoni gestori di password e, se ne conoscete già uno che vi piace, usatelo assolutamente. Eccone tre che vi consiglio.

- *Bitwarden* Questo gestore è gratuito e open source e sincronizza le password tra i vostri computer e il telefono. Dispone di estensioni del browser per inserire automaticamente le password quando fate accesso ai siti web. È una buona scelta come gestore di password per esigenze di tutti i giorni. Potete scaricarlo da: `https://bitwarden.com`.
- *1Password* Come Bitwarden, anche 1Password sincronizza le password tra il computer e il telefono e dispone di un'estensione per il browser. È anch'esso una buona scelta come gestore di password quotidiano. Costa qualcosa, ma 1Password offre licenze gratuite ai giornalisti. Potete scaricarlo da `https://1password.com`, ma potete cercare in `https://1password.com/for-journalism` ulteriori informazioni sul programma di licenze gratuite.
- *KeePassXC* Questo software è ottimo per situazioni di alta sicurezza. A differenza di Bitwarden e 1Password, KeePassXC non sincronizza il database delle password crittografate sul cloud, il che lo rende meno comodo ma potenzialmente più sicuro. Funziona bene su computer isolato con air gap. Potete scaricarlo da: `https://keepassxc.org`.

Se desiderate utilizzare Bitwarden, 1Password o un gestore di password simile che si sincronizza tra più dispositivi, seguite le istruzioni di installazione sul loro sito web per installare il programma sul computer, sul telefono e come estensione del browser web. Se utilizzate un gestore di password solo locale, come KeePassXC, installatelo sul vostro computer.
Quando configurate per la prima volta il vostro gestore di password, è estremamente importante non dimenticare la master password. A differenza della maggior parte delle password dei siti web, una master password non può essere reimpostata. Se la dimenticate, rimarrete per sempre senza accesso al vostro gestore di password e perderete tutte le vostre password. Scrivete la master password su un biglietto, finché non l'avrete memorizzata, poi distruggete tale biglietto.
Le migliori master password sono le *passphrase*, una sequenza di parole scelte a caso da un dizionario. Sono anche più facili da ricordare rispetto alle password completamente casuali. Un esempio di una buona passphrase è qualcosa come `movie-flanked-census6-casino-change`. Non ha alcun significato, ma con la pratica non è troppo difficile da memorizzare.
Una volta configurato l'account del gestore delle password, aggiungete le altre password al gestore. Iniziate aggiungendo le password che usate di più: magari la password della

vostra casella e-mail o le password degli account dei social media. Se avete riutilizzato le stesse password per più account, cogliete l'occasione per *cambiarle* e migliorarle. Ogni volta che create una nuova password, utilizzate il generatore del vostro gestore di password, uno strumento incluso per aiutarvi a creare password complesse. In genere, i generatori di password dispongono di impostazioni che vi consentono di scegliere se generare una password o una passphrase, se impiegare numeri o caratteri speciali, quanto dovrebbe essere lunga la password e così via.
Bitwarden, per esempio, può creare sia password sia passphrase. La Figura 1.2 mostra il generatore di password di Bitwarden, che è configurato per creare una passphrase con cinque parole, separate da trattini, in maiuscolo e comprendenti un numero.

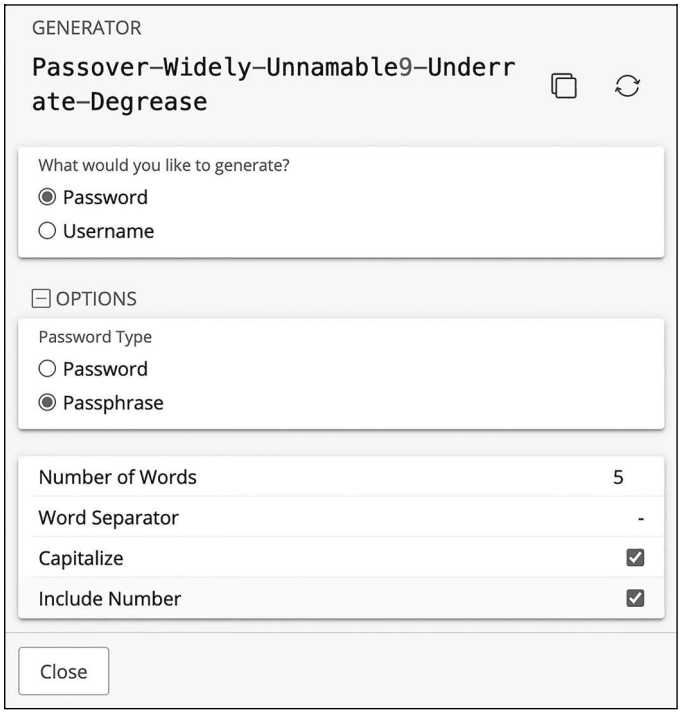


**Figura 1.2** Il generatore di password di Bitwarden.

Bitwarden può anche creare password complesse, come `Frz6ioX4o@cCY`. Tutte le vostre password dovrebbero essere passphrase complesse o password come questa.
Tutti i generatori di password previsti da 1Password, KeePassXC e altri includono funzionalità simili. Mentre Bitwarden vi consente di aprire lo strumento generatore di password in modo indipendente, altri gestori di password richiedono di aggiungere un nuovo elemento nel database delle password o di modificarne uno esistente per accedere al generatore.
Quando dovete inventare una nuova password, non è davvero importante se scegliete di utilizzare una password o una passphrase, purché sia forte e univoca. Tuttavia, le passphrase tendono a essere più facili da memorizzare e da inserire. Per questo motivo, tendo a utilizzare le password per accedere ai siti web (il mio gestore delle password le inserisce

per me) e le passphrase per tutto ciò che potrei aver bisogno di memorizzare o inserire, come la passphrase di crittografia del disco o la passphrase di accesso al mio computer. Ogni volta che create un nuovo account o fate accesso a un vostro account, aggiungete la password al vostro gestore di password.

## Crittografia del disco

La crittografia del disco vi consente di proteggere i vostri dati da chiunque possa avere accesso fisico al vostro telefono, laptop o disco USB. Impedisce a chiunque di accedere ai dati su un dispositivo se lo doveste perdere, se qualcuno ve lo rubasse, se vi venisse confiscato a un posto di frontiera o a un posto di blocco o se la vostra casa o il vostro ufficio dovessero essere perquisiti. Per esempio, se il disco interno del vostro laptop non è crittografato, chiunque ottenga un accesso fisico al computer potrà aprire il laptop, rimuovere il disco e collegarlo al proprio computer, accedendo a tutti i dati senza dover conoscere nessuna delle vostre password. Ma se il vostro disco è crittografato, tutti questi dati risulteranno completamente inaccessibili a chiunque non disponga della chiave. Grazie alla crittografia del disco, dovranno prima sbloccare il disco, in genere utilizzando una password, un PIN o dati biometrici come un'impronta digitale o la scansione del volto. Negli esercizi di questo capitolo imparerete a crittografare il vostro disco interno e il vostro disco USB da 1 TB.

Sebbene la crittografia del disco sia un aspetto importante della protezione dei dati, non protegge dagli attacchi remoti. Per esempio, se il vostro laptop è crittografato ma qualcuno vi induce con l'inganno ad aprire un documento Word dannoso che attacca il vostro computer, la crittografia del disco non gli impedirà di accedere ai vostri file. La crittografia del disco, inoltre, non sarà di grande aiuto se gli hacker riuscissero ad accedere al vostro dispositivo mentre è sbloccato, per esempio se vi allontanate dal vostro laptop mentre siete in un bar senza bloccare lo schermo o se gli hacker possono sbloccare facilmente il vostro telefono o indurvi a farlo utilizzando la protezione biometrica. Per esempio, dopo avervi arrestato, un poliziotto potrebbe agitarvi il telefono davanti al viso per sbloccarlo con una scansione del volto.

Ovviamente voi non impiegherete la crittografia del disco per commettere illeciti, ma la storia di Ross Ulbricht, creatore del sito web del market darknet Silk Road, è un buon esempio di come essa possa proteggervi. Nel 2013, Ulbricht stava usando il suo laptop crittografato presso la biblioteca pubblica di San Francisco quando due agenti dell'FBI sotto copertura lo distrassero fingendo di essere amanti nel corso di una lite. Dopo essersi assicurati che il suo schermo fosse sbloccato, lo hanno arrestato, poi hanno copiato file importanti dal suo computer. Se il suo schermo fosse stato bloccato e avesse impiegato una password complessa, la crittografia del disco avrebbe potuto impedire del tutto l'accesso ai suoi dati. Ulbricht è stato accusato di riciclaggio di denaro, pirateria informatica, traffico di droga e altri crimini.

La crittografia del disco interno del laptop è una misura di sicurezza di base, che tutti dovrebbero adottare. È facile e veloce da configurare, non richiede alcun impegno aggiuntivo ripetitivo e protegge la vostra privacy in caso di smarrimento del dispositivo. È un po' come indossare la cintura di sicurezza: non c'è davvero alcun buon motivo per non farlo. La crittografia del disco interno del laptop è particolarmente importante se vi troverete a lavorare con dati riservati.

# Esercizio 1.1 – Crittografare il disco interno

Questo esercizio vi mostra come crittografare il disco interno del vostro computer, indipendentemente dal fatto che abbiate un computer Windows, Mac o Linux. Procedete con la parte appropriata per il vostro sistema operativo.

## Windows

Le varie versioni di Windows e modelli di PC supportano tipi differenti di crittografia del disco. Le edizioni Pro di Windows includono BitLocker, la tecnologia di crittografia del disco Microsoft, mentre le edizioni Home includono la crittografia del dispositivo, che è fondamentalmente sempre BitLocker ma con funzionalità limitate. Tuttavia, queste funzionalità funzionano solo se il vostro PC è abbastanza recente. Se al momento dell'acquisto il vostro computer era dotato almeno di Windows 10, dovrebbe supportare la crittografia del disco, mentre se era dotato di una versione precedente di Windows, potrebbe non supportarla. Alla fine di questa parte del capitolo esaminerò le opzioni che avete a disposizione in quest'ultimo caso.

### BitLocker

Per scoprire se il vostro computer include BitLocker, fate clic su *Start* (l'icona di Windows in basso a sinistra sul computer), cercate *bitlocker* e fate clic su *Gestisci BitLocker*. Se la vostra versione di Windows impiega BitLocker, la finestra dovrebbe mostrare se BitLocker è abilitato sull'unità di sistema *C:* e dovreste avere la possibilità di abilitarlo. Se è così, fatelo subito.

Quando abilitate BitLocker, vi verrà chiesto di salvare una chiave di ripristino sul vostro account Microsoft, su un file di un'unità USB non crittografata o su un documento stampato. Salvare la chiave di ripristino sul vostro account Microsoft è l'opzione più semplice, ma significa che Microsoft o chiunque abbia accesso al vostro account Microsoft potrà accedere alla chiave necessaria per sbloccare il vostro disco. Se preferite non concedere a Microsoft questo accesso, stampate la chiave di ripristino. Dovreste salvare la chiave anche nel vostro gestore di password. Se il vostro computer si dovesse guastare, avrete bisogno della chiave di ripristino per accedere a tutti i dati contenuti nel vostro disco crittografato.

### Crittografia del dispositivo

Se la vostra versione di Windows non include BitLocker, provate ad applicare la crittografia del dispositivo. Fate clic su *Start,* quindi andate su *Impostazioni > Aggiornamento e sicurezza* (o *Privacy e sicurezza*, a seconda della versione di Windows). Quindi andate alla scheda *Crittografia dispositivo* per verificare se la crittografia è abilitata; in caso contrario, abilitatela.

Se non vedete la scheda *Crittografia dispositivo*, purtroppo il vostro PC non supporta la crittografia del dispositivo. Avete alcune opzioni. L'opzione più semplice consiste nell'eseguire l'aggiornamento alla versione Pro di Windows, che, in genere, costa un centinaio di euro, e poi utilizzare BitLocker. In alternativa, potete utilizzare VeraCrypt.

## VeraCrypt

VeraCrypt è un software gratuito e open source di crittografia del disco. Per iniziare, scaricate VeraCrypt da `https://veracrypt.fr`, installatelo sul computer e apritelo.
Fate clic su *Crea un volume* per lanciare la procedura guidata di creazione del volume VeraCrypt. VeraCrypt vi consente di scegliere tra tre tipi di volumi crittografati. Selezionate *Crittografa la partizione o tutto il disco di sistema* e fate clic su *Avanti*.
Nella pagina *Tipo di codifica di sistema*, scegliete *Normale* e fate clic su *Avanti*. Nella pagina *Area da codificare*, scegliere *Codifica la partizione di sistema Windows* e fate clic su *Avanti*. Nella pagina *Numero di sistemi operativi*, scegliete *Boot singolo* e fate clic su *Avanti* (se però disponete di più sistemi operativi sul vostro computer, scegliete *Avvio multiplo*). Nella pagina *Opzioni di codifica*, lasciate le impostazioni di default e fate clic su *Avanti*.
La pagina successiva è *Password*. Dovete inventare una passphrase complessa da inserire ogni volta che avviate Windows. Se la passphrase è debole, anche la crittografia del disco sarà debole. Vi consiglio di generare una passphrase complessa e di salvarla nel vostro gestore di password: in questo modo, se la doveste dimenticare al prossimo riavvio del computer, potrete cercarla nel gestore di password del vostro telefono. Immettete la passphrase due volte e fate clic su *Avanti*.
La pagina successiva si chiama *Raccolta di dati casuali*. VeraCrypt include una funzionalità, mediante la quale dovete muovere il mouse sulla finestra in modo casuale, in modo che il programma possa raccogliere informazioni il più possibile casuali dai movimenti del mouse, per rendere la crittografia più sicura. Muovete il mouse finché la barra nella parte inferiore della finestra non diventa verde, poi fate clic su *Avanti*. Fare nuovamente clic su *Avanti* nella pagina *Chiavi generate*.
La pagina *Disco di ripristino* vi chiede di creare un disco che potete utilizzare nel caso in cui il vostro disco venisse danneggiato e qualora riscontriate problemi nell'avvio di Windows. La creazione di un disco di ripristino non rientra negli scopi di questo libro, quindi selezionate pure *Salta verifica del Disco di Recupero* e fate clic su *Avanti*. Nella pagina *Disco di ripristino creato*, fare nuovamente clic su *Avanti*.
Nella pagina *Modo pulizia*, selezionate *Nessuno (il più veloce)* come modalità di cancellazione e fate clic su *Avanti*. Nella pagina *Pre-test di codifica del sistema*, fate clic su *Prova* per verificare che la crittografia del disco funzioni correttamente sul vostro computer: questo riavvierà il computer e dovrete inserire la passphrase VeraCrypt per l'avvio.
Al riavvio del computer, dovrebbe essere lanciato il bootloader di VeraCrypt e dovrete inserire la sua passphrase per procedere. Per *PIM*, basta premere *Invio*. Se tutto va bene, il boot avrà successo, Windows si avvierà e si aprirà nuovamente VeraCrypt nella pagina *Pre-test completato*. Fate clic su *Codifica* per iniziare a crittografare il vostro disco interno con VeraCrypt. D'ora in poi, dovrete inserire la vostra passphrase VeraCrypt ogni volta che avvierete il computer, ma anche tutti i vostri dati saranno protetti con questa passphrase.

## Mac OS

La tecnologia di crittografia del disco di Apple si chiama FileVault. Se utilizzate macOS Ventura o versioni successive, aprite l'app *Impostazioni di Sistema*, fate clic su *Privacy e sicurezza* a sinistra e scorrete il contenuto della finestra verso il basso fino alla sezione *FileVault*. (Se, invece, utilizzate una versione di macOS precedente a Ventura, aprite l'app

*Preferenze di Sistema*, fate clic su *Sicurezza e Privacy* e assicuratevi di essere nella scheda *FileVault*.) Se FileVault è disattivato, attivatelo.
La password che sblocca il disco del vostro Mac è la password che usate per accedere al vostro account. Assicuratevi che la password del vostro Mac sia complessa; se la vostra password è debole, anche la crittografia del disco sarà debole.
Se abilitate FileVault, dovrete salvare una chiave di ripristino. Salvate la chiave nel vostro gestore di password. Se dimenticate la password del vostro Mac, avrete bisogno della chiave di ripristino per accedere a tutti i vostri dati. Se utilizzate un gestore di password locale che non si sincronizza con il cloud, come KeePassXC, archiviate una copia della chiave di ripristino anche altrove, per esempio su un biglietto che conserverete in un luogo sicuro.

## Linux

Linux utilizza la tecnologia chiamata LUKS per la crittografia del disco. Potete controllare se il disco interno è crittografato con programma *Disks* (*Dischi*). Nella maggior parte delle versioni di Linux, per aprire questo programma dovete premere il tasto "Windows", digitare *disks* e premere Invio. Il programma mostra tutti i dischi connessi al computer e consente di formattarli. Se il vostro disco interno ha una partizione *Unlocked* con crittografia *LUKS*, significa che la crittografia del disco è abilitata.
Nel mio caso, rappresentato nella Figura 1.3, il mio disco interno è un'unità SSD Samsung da 500 GB, come indicato a sinistra nella figura. Il mio disco contiene quattro partizioni e l'ultima (*Partition 4*) è da 499 GB ed è crittografata con LUKS. Il vostro disco avrà un aspetto differente, ma saprete se è crittografato se la partizione principale riporta *LUKS*.
Sfortunatamente, non potete semplicemente attivare o disattivare la crittografia LUKS. Se il vostro disco non è crittografato, l'unico modo per crittografarlo consiste nel reinstallare Linux, assicurandovi di crittografare il disco. Quando installate Linux, uno dei primi passi nel processo di installazione sarà il partizionamento del disco; assicuratevi di abilitare la crittografia in questo passaggio. Se avete intenzione di reinstallare Linux, eseguite sempre prima il backup dei vostri dati. Dopo aver scelto la passphrase di crittografia, salvatene una copia nel vostro gestore di password; ne avrete bisogno ogni volta che avvierete il computer.

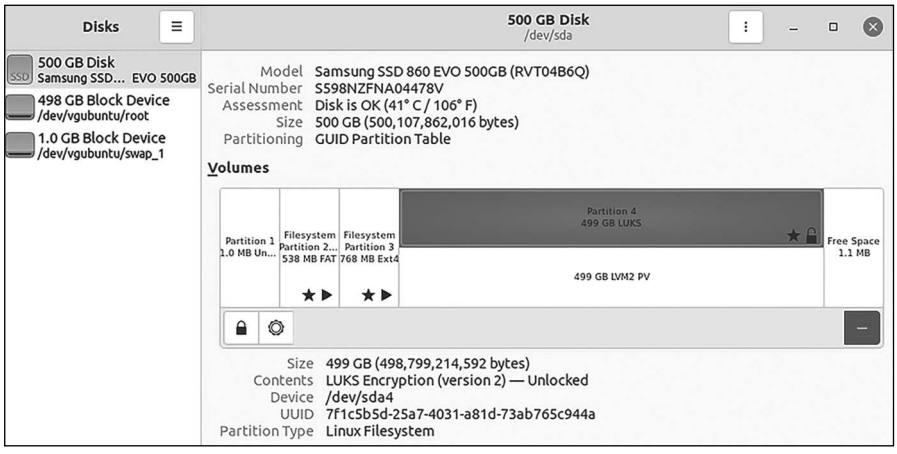


**Figura 1.3** Gestione dei dischi e delle partizioni con Disks in Linux.

# Esercizio 1.2 – Crittografare un disco USB

Il disco interno del vostro computer da solo probabilmente non è abbastanza grande per archiviare tutti i dataset con cui dovrete lavorare. Come ho detto nell'introduzione del libro, per completare gli esercizi di questo libro e lavorare con dataset di enormi dimensioni, è necessario un disco USB da *almeno 1 TB*. Per crittografare il disco USB è necessario anche formattarlo, eliminando così tutti i dati già presenti su di esso. Questo esercizio vi mostra come farlo per qualunque sistema operativo stiate utilizzando.

Prima di iniziare, esaminiamo alcune informazioni generali su come funzionano i dispositivi di archiviazione di massa (come i dischi rigidi, le schede SD e così via). I dispositivi di archiviazione sono generalmente suddivisi in una o più *partizioni*, chiamate anche *volumi*, e ciascuna partizione utilizza un formato chiamato *filesystem*. Potete considerare le partizioni come armadi che utilizzano sistemi di scaffalatura (filesystem) differenti per organizzare i dati. Sistemi operativi differenti utilizzano filesystem differenti. Windows, in genere, utilizza il filesystem NTFS, macOS normalmente utilizza APFS e Linux utilizza soprattutto ext4. Esistono anche filesystem che possono essere utilizzati da tutti e tre i sistemi operativi, come ExFAT.

Quando inizializzate un dispositivo di archiviazione, eliminate tutte le partizioni su di esso in modo che il disco contenga solo spazio non allocato. Poi potrete creare una nuova partizione (con i dischi USB, in genere creerete una singola partizione che occupa tutto lo spazio non allocato) e formattarla utilizzando il filesystem corrispondente al vostro sistema operativo.

Che stiate lavorando in Windows, macOS o Linux, iniziate collegando il vostro disco USB al computer. Aprite il vostro gestore di password e salvate una nuova passphrase complessa, creata utilizzando il generatore del vostro gestore di password. Assegnate alla password un nome simile *Crittografia del disco USB per i dataset*.

Per iniziare a crittografare il disco, procedete con la parte appropriata per il vostro sistema operativo.

## Windows

Gli utenti Windows con BitLocker dovrebbero impiegare il seguente paragrafo; se invece non avete BitLocker, procedete alla sezione VeraCrypt.

### BitLocker

Se avete un computer Windows con BitLocker, utilizzatelo per crittografare la vostra USB. Innanzitutto, assicuratevi di formattare il disco USB in formato NTFS. Per fare ciò, fate clic su *Start,* digitate *Gestione disco* e si aprirà l'app *Gestione disco* di Windows, rappresentata nella Figura 1.4, che elenca tutti i dischi connessi al PC e consente di formattarli. La parte inferiore della finestra mostra ciascun disco connesso al computer e le loro partizioni. Nella figura, che rappresenta la mia macchina, disco 0 è il mio disco rigido interno (come potete vedere, una delle partizioni è *C:*) e il disco 1 è un disco USB (una di queste partizioni è *D:*). Sul mio computer, il disco 1 ha una singola partizione da 32 GB e circa 86 GB di spazio non allocato.

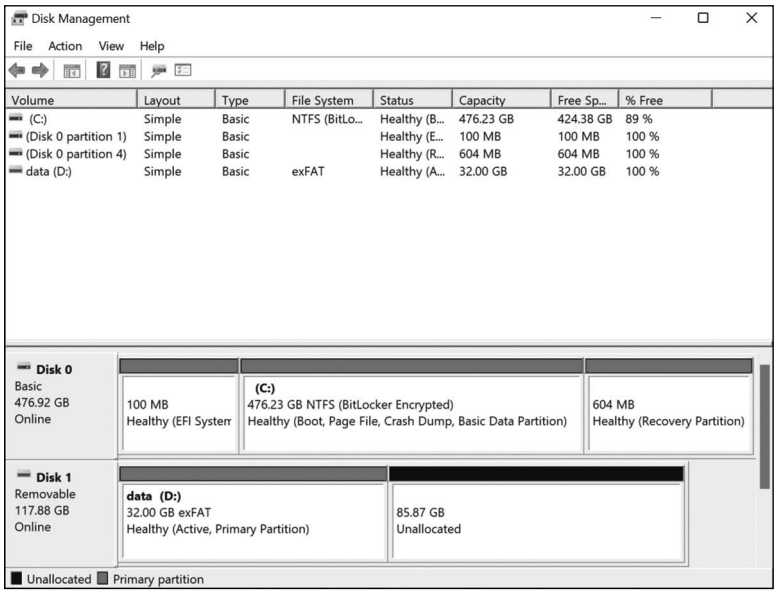


**Figura 1.4** L'app Disk Management (Gestione disco) di Windows.

Trovate il disco USB che dovete formattare. Fate clic destro su ogni partizione e scegliete *Elimina volume* fino a quando non avrete eliminato dal disco tutte le partizioni. Quindi fate clic destro sullo spazio non allocato nel disco e scegliete *Nuovo volume semplice*. Si aprirà una procedura guidata per aiutarvi a creare il volume. Scegliete tutto lo spazio su disco e formattatelo come NTFS. La procedura guidata vi chiederà un'*etichetta di volume*, che è solo un nome per la vostra partizione; nella Figura 1.4, l'etichetta per *D:* è *data*. Vi consiglio di chiamare questo disco *dataset*.

Una volta formattato il disco, fate clic su *Start,* cercate *bitlocker* e aprite *Gestisci BitLocker*. Ora dovreste vedere il vostro disco USB e avere la possibilità di attivare BitLocker. Quando abilitate BitLocker sul vostro disco USB, dovrebbe apparire una finestra che vi chiede come desiderate sbloccare questa unità. Scegliete *Usa una password per sbloccare l'unità,* quindi copiate e incollate la passphrase di crittografia del disco USB dal gestore delle password nel campo della password. Dovrete incollarla anche nel campo per reinserire la password. Quando abilitate BitLocker, vi verrà richiesto di salvare una chiave di ripristino in un file. Tuttavia, poiché state salvando la passphrase in un gestore di password, non avete bisogno della chiave di ripristino e potete eliminare il file.

## VeraCrypt

Se usate Windows Home e non avete BitLocker sul vostro computer, usate VeraCrypt per crittografare il vostro disco USB.

Se non avete VeraCrypt, scaricatelo da `https://veracrypt.fr`, installatelo sul computer e apritelo. Fate clic su *Crea un volume* per aprire la procedura guidata di creazione del volume VeraCrypt. Nella prima pagina della procedura guidata, VeraCrypt vi consente di scegliere tra tre tipi di volumi crittografati. Selezionate *Codifica una partizione/disco non di sistema* e fate clic su *Avanti*.

Nella pagina *Tipo di volume*, VeraCrypt vi chiede se desiderate creare un volume standard oppure nascosto. Selezionate l'opzione *Volume VeraCrypt standard* e fate clic su *Avanti*.
Nella pagina *Percorso del volume*, fate clic su *Seleziona unità,* scegliete il disco USB che desiderate crittografare e fate clic su *Avanti*.
Nella pagina *Modalità creazione volume*, selezionate l'opzione *Crea e formatta volume criptato*, poi fate clic su *Avanti*.
Nella pagina *Opzioni di codifica*, accettate le impostazioni di default e fate clic su *Avanti*.
Non potete fare nulla nella pagina *Dimensioni del volume*, poiché state crittografando un'intera partizione anziché creare un container di file crittografati, quindi fate semplicemente clic su *Avanti*.
Nella pagina *Password del volume*, copiate e incollate la passphrase di crittografia del disco USB dal gestore delle password nel campo *Password*, poi incollatela nuovamente nel campo *Conferma*. Quindi fate clic su *Avanti*.
Nella pagina *File grandi*, VeraCrypt vi chiede se nel vostro volume intendete archiviare file più grandi di 4 GB. Selezionare *Sì* e fate clic su *Avanti*.
Nella pagina *Formattazione del volume*, nel menu a discesa *File system*, selezionate *exFAT* e selezionate nella casella accanto, *Formattazione rapida*. VeraCrypt include una funzionalità, mediante la quale dovete muovere il mouse sulla finestra in modo casuale, in modo che il programma possa raccogliere informazioni il più possibile casuali dai movimenti del mouse, per rendere la crittografia più sicura. Muovete il mouse finché la barra nella parte inferiore della finestra non diventa verde, poi fate clic su *Formatta*.
Dovrebbe apparire una finestra che vi avverte che tutti i dati sulla vostra unità USB verranno cancellati e vi chiede se siete sicuri di procedere. Fate clic su *Sì*, quindi attendete che VeraCrypt crei una partizione crittografata sul disco USB. Se nella pagina precedente avete selezionato *Formattazione rapida*, l'operazione dovrebbe richiedere solo pochi secondi.
Nella pagina *Volume creato*, fate clic su *Esci* per uscire dalla procedura guidata e tornare alla finestra principale di VeraCrypt.
Dopo aver crittografato un disco USB con VeraCrypt, dovete utilizzare VeraCrypt per *montarlo*, ovvero renderlo disponibile sul vostro computer con una sua lettera di unità. Nella finestra principale di VeraCrypt, selezionate una lettera di unità disponibile (per esempio *F:*), fate clic su *Seleziona unità,* selezionate il vostro disco USB crittografato con VeraCrypt e fate clic su *OK*, quindi su *Monta*. Dopo aver fornito la passphrase di crittografia per sbloccarlo, VeraCrypt monterà il vostro disco USB crittografato in modo che possiate usarlo. Ora tutti i file salvati su questa unità verranno crittografati sul disco. Prima di scollegare il disco USB, smontatelo selezionando la lettera dell'unità in VeraCrypt e facendo clic su *Smonta*.

**NOTA**

VeraCrypt è utile anche se volete accedere allo stesso disco crittografato su più sistemi operativi, per esempio se dovete utilizzarlo sia su un PC Windows sia su un Mac. Tuttavia, ai fini di questo libro, solo gli utenti Windows che non dispongono di BitLocker dovrebbero utilizzare VeraCrypt. In generale, avrete meno problemi se rimarrete fedeli al software di crittografia del disco già integrato nel vostro sistema operativo.

## Mac OS

Aprite l'app *Utility Disco*, che potete trovare nella cartella *Applicazioni/Utility*. Questa app elenca tutti i dischi connessi al vostro computer e vi consente di formattarli.
In *Utility Disco*, selezionate il disco USB che avete connesso e fate clic sul pulsante *Inizializza*. Assegnate al disco il nome *dataset* e scegliete il formato *APFS (crittografato)*. Di conseguenza, vi verrà richiesta la password per sbloccare il disco crittografato. Copiate e incollate dal vostro gestore di password a *Utility Disco* la passphrase di crittografia del disco USB che avete creato all'inizio di questo esercizio. *Utility Disco* vi chiederà anche un suggerimento per la password, ma poiché state salvando questa passphrase nel vostro gestore di password e non volete preoccuparvi di memorizzarla, potete lasciare vuoto il suggerimento per la password.

## Linux

Aprite l'app *Disks* (*Dischi*) come avete fatto per l'Esercizio 1.1. Selezionate il vostro disco USB nell'elenco dei dischi, a sinistra, poi fate clic sul pulsante del menu (*Opzioni unità*) e scegliete il comando *Format Disk* (*Formatta disco*). L'operazione eliminerà tutti i dati presenti sull'unità USB.
Fate clic sul pulsante + per aggiungere una nuova partizione e impostate le dimensioni della partizione sull'opzione più grande. Assegnate all'unità il nome *dataset*, scegliete *Internal Disk for Use with Linux Systems Only* (*Disco interno per un uso esclusivo con sistemi Linux (Ext4)*) e selezionate la casella *Password Protect Volume (LUKS)* (*Volume protetto da password (LUKS)*). Vi verrà chiesto di inserire una password. Copiate e incollate dal vostro gestore di password a *Disks* (*Dischi*) la passphrase di crittografia del disco USB creata all'inizio di questo esercizio.

# Protezione dai documenti dannosi

Prima di iniziare a lavorare con qualsiasi dataset sul disco USB che avete appena crittografato, dovreste imparare a proteggervi da documenti potenzialmente dannosi che tali dataset potrebbero contenere.
Probabilmente qualcuno vi avrà già detto di evitare di aprire allegati di posta elettronica provenienti da fonti sconosciute. Questo è certamente un consiglio valido, ma sfortunatamente per i ricercatori, i giornalisti, gli attivisti e molte altre persone è impossibile da seguire. In questi ambiti lavorativi, spesso è *vostro preciso compito* aprire i documenti che vi inviano perfetti sconosciuti, inclusi i dataset trapelati o violati.
Aprire documenti di cui non vi fidate è pericoloso, perché potrebbe consentire ad altri di hackerare il vostro computer. I PDF e i documenti Microsoft Office o LibreOffice sono incredibilmente complessi. Possono essere impostati per caricare automaticamente un'immagine da un server remoto, monitorando quando un documento viene aperto e da quale indirizzo IP. Possono contenere codice JavaScript o macro che, a seconda di come è configurato il software, potrebbero lanciarsi automaticamente quando i documenti vengono aperti, potenzialmente assumendo il controllo del vostro computer. E come tutti i software, i programmi che usate per aprire i documenti, come Microsoft Office

e Adobe Reader, contengono bug che a volte possono essere sfruttati per assumere il controllo del vostro computer.
Questo, per esempio, è esattamente ciò che hanno fatto i servizi segreti militari russi durante le elezioni americane del 2016. In primo luogo, la Direzione Principale dello Stato Maggiore Generale delle Forze Armate della Federazione Russa ha hackerato un sistema elettorale statunitense VR Systems e ha ottenuto la lista dei funzionari elettorali relativi agli stati chiave. Poi ha inviato 122 messaggi e-mail ai clienti di VR Systems dall'indirizzo e-mail `vrelections@gmail.com`, con l'allegato `New EViD User Guides.docm`. Se qualcuno dei funzionari elettorali che hanno ricevuto questa e-mail avesse aperto l'allegato utilizzando una versione vulnerabile di Microsoft Word per Windows, il malware avrebbe creato nel loro computer una backdoor per gli hacker russi. Non sappiamo con certezza se qualcuno degli operatori abbia aperto l'allegato dannoso.
L'invio di e-mail dannose a obiettivi ben precisi in questo modo nell'ambito di un'operazione di hacking è chiamato *spearphishing*. La Figura 1.5 mostra un messaggio e-mail di spearphishing che aveva come obiettivo un funzionario elettorale nella Carolina del Nord, che *The Intercept* ha ottenuto utilizzando una richiesta di registri pubblici.

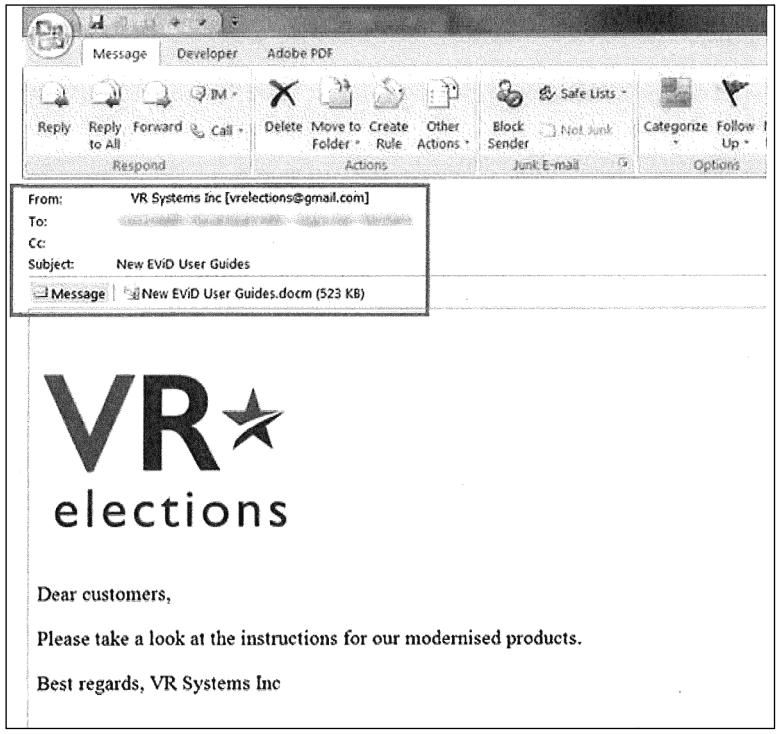


**Figure 1.5** Un'e-mail di spearphishing rivolta ai funzionari elettorali.

Nel 2017, Reality Winner ha fatto trapelare a *The Intercept* un documento riservato che descriveva questo attacco di spearphishing. Grazie alle sue denunce, il pubblico è molto più informato sugli attacchi condotti dalla Russia alle elezioni americane del 2016. In effetti, certi stati americani, come la Carolina del Nord, hanno appreso di essere stati attaccati da hacker russi solo leggendo *The Intercept*. Nel 2022, due ex funzionari elettorali

hanno dichiarato a *60 Minutes* che la divulgazione di Reality Winner ha contribuito a proteggere le elezioni di medio termine del 2018 da analoghi tentativi di hacking.
Per rendere più sicura l'apertura di documenti non attendibili, ho sviluppato un'app open source chiamata Dangerzone. Quando aprite un documento non attendibile in Dangerzone, l'app lo converte in un file PDF *certamente sicuro*, di cui potete fidarvi. Utilizzando una tecnologia chiamata *container Linux*, che sono come computer Linux veloci, piccoli e autonomi in esecuzione all'interno di un normale computer, converte, se necessario, il documento originale in un file PDF, lo divide in più pagine distinte e converte ogni pagina in pixel. Poi, in un altro container Linux, converte nuovamente i dati dei pixel in un file PDF. Potete anche chiedere a Dangerzone di utilizzare una tecnologia OCR (riconoscimento ottico dei caratteri) per esaminare l'immagine ed estrarne i caratteri, per aggiungere nuovamente il testo al PDF in modo da renderlo nuovamente ricercabile. Dangerzone, essenzialmente, "stampa" un documento e poi ne esegue una scansione, eliminando qualsiasi elemento dannoso in esso presente e rimuovendo anche i metadati presenti nel documento originale. Se apriste il documento dannoso `New EViD User Guides.docm` utilizzando Dangerzone, il software creerebbe un nuovo documento `New EViD User Guides-safe.pdf` che potreste aprire tranquillamente e senza rischi. Come ulteriore vantaggio, non è necessario essere connessi a Internet per utilizzare Dangerzone, che quindi funziona molto bene anche su computer isolato con air gap.
Parleremo ampiamente di Dangerzone e dei container Linux nel Capitolo 5, che spiega come rendere ricercabili i dataset. Nel frattempo, l'Esercizio 1.3 vi mostrerà come iniziare.

## Esercizio 1.3 – Installare e utilizzare Dangerzone

In questo esercizio installerete Dangerzone e lo utilizzerete per convertire i documenti in versioni sicure. La Figura 1.6 mostra uno screenshot di Dangerzone in azione: in questo caso, la conversione del documento non attendibile `D&D 5e - Players Handbook.pdf` in una versione sicura chiamata `D&D 5e - Players Handbook-safe.pdf`, sottoposta a OCR e ricercabile.


**Figura 1.6** Dangerzone in azione.

Scaricate e installate Dangerzone da `https://dangerzone.rocks`. L'app si basa su container Linux. Se lavorate su un computer Windows o macOS, il modo più semplice per far funzionare i container consiste nell'utilizzare un software chiamato Docker Desktop, che vi verrà chiesto di installare la prima volta che aprite Dangerzone. Per ora non dovete fare nulla con Docker Desktop; vi basta installarlo e aprirlo. Imparerete a utilizzarlo nel Capitolo 5.

Ora che Dangerzone è installato, provatelo. Aprite in Dangerzone un file PDF, un documento di Microsoft Office, un documento di LibreOffice o un'immagine sul vostro computer e convertitelo in un file PDF sicuro. Se qualcuno allega un documento a un'e-mail e non vi fidate, scaricatene una copia, aprite Dangerzone e fate clic su *Select Suspicious Documents.* Quindi cercate il documento che avete scaricato e utilizzate Dangerzone per convertirlo in una versione sicura.

## Macchine virtuali

Un'altra opzione, un po' più complicata, consiste nel configurare una *macchina virtuale* (*VM*). Le macchine virtuali sono qualcosa di più dei container Linux. Isolano il software in esecuzione all'interno della macchina virtuale più di quanto possano fare i container Linux e possono essere eseguite su qualsiasi sistema operativo. Se scegliete questa opzione, assicuratevi di disabilitare l'accesso a Internet nella vostra macchina virtuale prima di aprire i documenti. In questo modo, se il documento è dannoso, non potrà rivelare agli hacker che è stato aperto.

Fornire istruzioni dettagliate sull'utilizzo delle macchine virtuali non rientra negli scopi di questo libro. Tuttavia, se volete provarle, il modo più semplice per iniziare consiste nell'utilizzare il software di virtualizzazione gratuito e open source VirtualBox (`https://www.virtualbox.org`). VirtualBox funziona su Mac basati su Intel, Linux e computer Windows. Al momento, esiste una versione beta di VirtualBox che supporta i Mac dotati di Apple Silicon, ma presenta problemi. Se avete un Mac dotato di Apple Silicon, una buona scelta è UTM (`https://mac.getutm.app`). È gratuito e open source e potete trovare istruzioni dettagliate per l'installazione su sistemi operativi differenti su `https://docs.getutm.app/guides/guides`. Se invece desiderate qualcosa di un po' più semplice, vi consiglio di provare Parallels (`https://www.parallels.com`) o VMware Fusion (`https://www.vmware.com/products/fusion.html`); nessuno dei due, però, è gratuito.

Dangerzone funziona benissimo con PDF e documenti Word, ma non altrettanto bene con i fogli di lavoro. Indipendentemente dal tipo di file che aprite in Dangerzone, vi ritroverete sempre con un file PDF sicuro e i fogli di lavoro non sono pensati per essere letti in quel formato.

Se Dangerzone non riuscisse a svolgere un buon lavoro con un documento che intendete leggere, potete aprirlo in un altro modo, contenendo i rischi. Se ritenete che il documento non sia riservato, caricatelo su Google Drive e apritelo da lì, utilizzando l'interfaccia web di Google. In questo modo, tecnicamente è Google che apre il documento dannoso sui suoi computer, non voi.

## Riepilogo

In questo capitolo avete imparato a pensare alla protezione della vostra fonte in un mondo caratterizzato da una diffusa sorveglianza digitale. Avete anche imparato ad archiviare in modo sicuro i dataset, in base alla loro riservatezza; a verificare che i vostri dataset siano autentici; a oscurare le informazioni contenute nei documenti prima di pubblicare il report finale. Avete iniziato a utilizzare un gestore di password per mantenere le vostre password al sicuro, avete crittografato il vostro disco interno e configurato il disco USB che dedicherete ai *dataset* crittografati. Infine, vi siete esercitati a trasformare i documenti potenzialmente dannosi in documenti certamente sicuri utilizzando Dangerzone.
Nel prossimo capitolo utilizzerete il disco dei vostri *dataset* scaricando il vostro primo dataset sottoposto ad hacking.