

L4 – Google Takeout

Titolo

Ottenere ed esplorare l'archivio Google Takeout.

Scopo

Lo scopo dell'esercizio è ottenere la maggior quantità possibile di dati che Google mantiene su un account. Nell'ordine sarà presentato come ottenere i file, e successivamente come interpretare alcuni di questi dati di backup.

Per chi

L'esercizio è pensato per un pubblico ampio, e non sono richieste conoscenze preliminari. È necessario possedere un account Google personale e un computer.

Durata

Gran parte del tempo necessario allo svolgimento è da riservare al download dell'archivio.

Tempo stimato complessivo : 24h

Tempo di svolgimento ad archivio scaricato : 1h

Livello arcobaleno

Livello L4 – transparency

Sottolivello L1

Difficoltà

Media, si tratta di saper navigare tra i contenuti di una cartella e interpretare file per lo più testuali, ma richiede conoscenza e consapevolezza delle tracce della propria attività legata all'account Google personale.

Strumenti, s.o., pacchetti/app

È necessario possedere un computer e un software per decomprimere file.

Programmi specifici per interpretare file diversi dipendono dai file stessi, e possono variare per ogni utente.

I più comuni saranno trattati nella sezione "Istruzioni dettagliate".

Costi

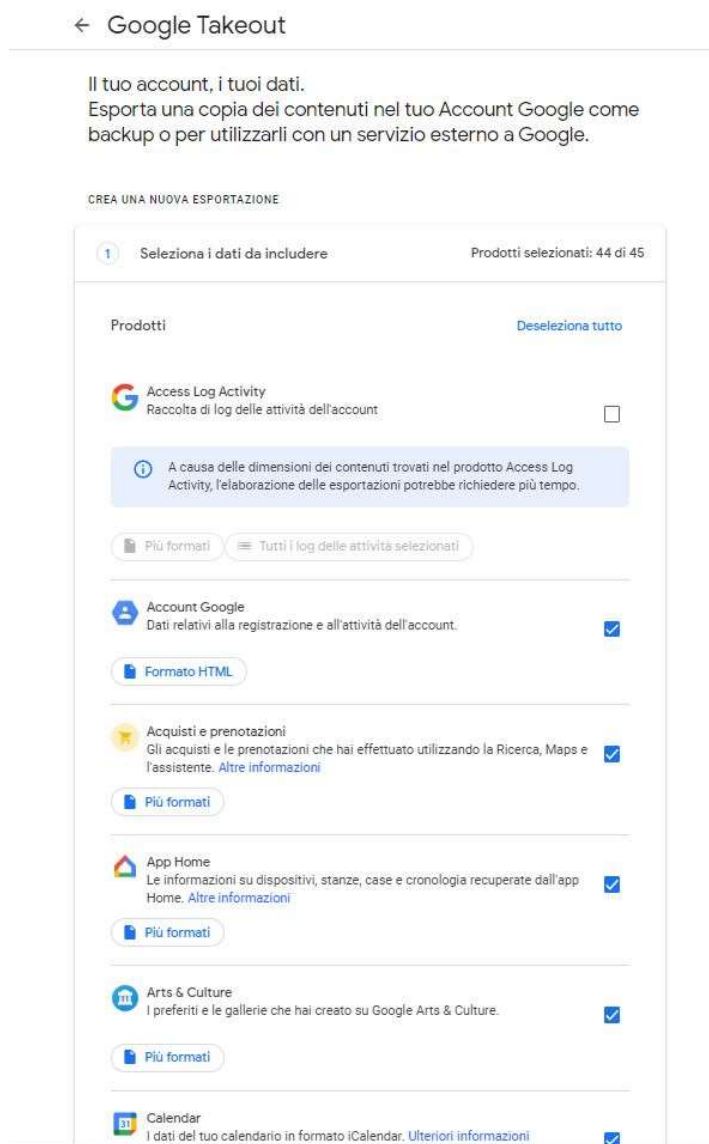
Il download dell'archivio è gratuito, i costi possono dipendere dai software scelti per l'analisi.

È possibile completare l'esercizio gratuitamente.

Istruzioni dettagliate

Il primo passo per ottenere l'archivio di dati personali tramite Google Takeout è collegarsi al sito web seguente : <https://takeout.google.com/settings/takeout>

Effettuato il login con il proprio account Google compare una lista di “prodotti”, selezionabili tramite checkbox a lato di ogni voce in elenco.



Di default dovrebbero essere selezionate tutte le voci meno “Access Log Activity”, che consiglio di selezionare, procedendo poi a scorrere la lista fino a selezionare tutto, in modo da ottenere il 100% dei dati personali mantenuti da Google.

Terminato il processo di selezione scorrere fino a fondo elenco e cliccare sul pulsante “Passaggio Successivo”.

Selezionare le voci come in elenco di seguito :

- Metodo di recapito -> Invia tramite email il link per il download (questa opzione evita problemi di spazio dei servizi cloud)
- Frequenza -> esporta un archivio (1 esportazione)
- Tipo di file e dimensioni -> preferenza personale tra .zip e .tgz
 - > 50GB se la scelta precedente è .tgz o è possibile aprire file .zip64, o 2GB

Cliccare “Crea Esportazione”

2

Scegli il tipo di file, la frequenza e la destinazione

Destinazione

Trasferimento a:

Invia tramite email il link per il download

Quando i file sono pronti, riceverai un'email con un link per il download. Avrai una settimana di tempo per scaricare i file.

Frequenza

☒ Esporta un archivio

1 esportazione

☐ Esporta ogni due mesi per un anno

6 esportazioni

Tipo di file e dimensioni

Tipo di file:

.zip

I file zip possono essere aperti su quasi tutti i computer.

Dimensioni file:

2 GB

Le esportazioni di dimensioni superiori a quelle indicate verranno suddivise in diversi file.

Crea esportazione

A questo punto è necessario attendere di ricevere tramite e-mail i link per il download degli archivi contenenti i dati richiesti.

Potrebbe essere necessario aspettare anche 24 ore.

Ricevuta la e-mail sarà possibile ottenere gli archivi tramite i link di download presenti sotto la dicitura “Gestisci Esportazioni”.

Analisi dell’archivio

Il primo passo è decomprimere le cartelle ottenute tramite i link di download.

Al loro interno sono presenti sottocartelle specifiche per ogni prodotto selezionato precedentemente in fase di richiesta. Se sono stati creati più link di download e, di conseguenza, più cartelle, consiglio di riorganizzare l’archivio in modo da avere tutti i contenuti in una cartella unica.

Il contenuto delle cartelle varia da account a account, così come il numero di sottocartelle presenti. Verranno qui analizzati alcuni tra i file più diffusi, e verrà fornita un’overview su come aprire e interpretare i vari formati di dati.

Per una panoramica più ordinata contenente anche in dettaglio le descrizioni dei formati file è possibile aprire il link “archive_browser” presente in archivio. È utile avere il link aperto mentre si esplora l’archivio, per avere più informazioni sul contenuto delle directory.

Per ogni prodotto è disponibile una schermata di informazioni generali, l’elenco di file esportati e la descrizione dei formati dei file.

Formati File più frequenti

- CSV: di default su Windows sono aperti con Excel, per avere una visualizzazione più ordinata è possibile seguire i seguenti passi
Dati -> Recupera e trasforma dati -> da testo/CSV -> Delimitatore = “virgola”
- HTML: sono aperti tramite browser
- JSON/GeoJSON: possono essere trattati come file .txt, aperti con text editor, notepad, atom, ...
- iCalendar: visualizzabili con app di gestione del calendario o Outlook, il contenuto è visibile anche se aperti in formato testo, ma meno ordinato
- KML/KMZ: sono file di Google Earth/Google Maps, sono di fatto file xml apribili con editor di testo
- mbox: file contenenti informazioni sulla posta elettronica, si possono visualizzare con Outlook o affini
- PNG/JPG/GIF/MP4: formati per foto-video standard e molto diffusi.

Directory Principali

Lo scopo di questo esercizio è di mostrare come ottenere i dati e come leggerli, ma, vista la quantità e la natura dei dati appena ottenuti, vale la pena offrire qualche spunto su come navigare l'archivio per essere un po' più consapevoli di cosa si ha davanti. Inoltre, data la variabilità del contenuto dell'archivio appena ottenuto, non è possibile offrire una guida completa che comprenda ogni possibile file e directory, né fornire immagini “guida” utili senza presentare dati personali dell'autore di questo esercizio.

Per affrontare la fase di interpretazione dei dati, è sempre bene tenere a mente che, perché l'esercizio sia utile e non un compito tecnico, è necessario interrogarsi a ogni step su cosa si sta trattando, consapevoli che il soggetto dell'analisi sono informazioni personali.

Alcuni quesiti utili possono essere :

- Come sono stati ottenuti questi dati?
- Ero a conoscenza della raccolta di questo tipo di dati?
- A che scopo questi dati sono raccolti e mantenuti?

Come primo passo si consiglia di aprire la cartella “Access Log Activity”, contenente file CSV che mostrano tutti gli accessi effettuati dall'account con relativa app richiedente accesso, IP, OS, timestamp... e dispositivi a cui l'account ha avuto accesso.

Se presente, la cartella “Chrome” contiene anch'essa informazioni molto interessanti sulle abitudini di ricerca, segnalibri, estensioni, ecc. Sono qui presenti anche le informazioni di autofill per form online, in cui è possibile trovare ad esempio numeri di cellulare o indirizzi di residenza, con tutti i dati forniti alla funzione di autofill.

In “Cronologia delle Posizioni” è presente una lista, divisa per anno, di ricerche e posizioni salvate, con coordinate, veicolo su cui si viaggiava/metodo di spostamento, luogo di interesse più vicino alla posizione rilevata e grado di probabilità/certezza sui dati raccolti.

Mail, Calendario e Contatti sono ottimi esempi per vedere l'età dei dati mantenuti, conversazioni salvate e media condivisi.

Risultati attesi

A download completato risulta possibile visualizzare tutti i dati che Google mantiene sull'account personale scelto in fase di login.

Obiettivi formativi

Da quantità e tipo di dati ottenuti e analizzati è certamente possibile aumentare la consapevolezza personale rispetto a come Google tratta i dati nei suoi servizi e app.

L'invito è quello di ragionare su come ogni nostra interazione con servizi e app di Google resti (in/consapevolmente) registrata, in alcuni casi anche per tempi significativamente più lunghi del necessario, oltre ad avere un'idea più chiara dell'esposizione delle proprie informazioni personali.

See also, prosegui con, propedeuticità

Il servizio di Takeout descritto è esclusivo a Google, è possibile ma non garantito che altre compagnie abbiano servizi simili su cui applicare un'analisi simile.

Restando all'interno dell'archivio ottenuto in questo esercizio, il consiglio è di approfondire ogni directory e file.

Autore

Daniele Scaccabarozzi