**Esercizio da svolgere**

Scrivere un programma Python che legge una lista di log dal file “anonimizza\_test1.json”[[1]](#footnote-1).

Ciascun elemento della lista di log è costituito dalle seguenti 9 informazioni:

- Data/Ora,

- Nome completo dell'utente

- Utente coinvolto

- Contesto dell'evento

- Componente

- Evento

- Descrizione

- Origine

- Indirizzo IP

Il programma deve anonimizzare la lista. In particolare:

- per ogni utente (campo “Nome completo dell'utente”) generare un identificatore unico (per esempio un numero progressivo)

- sostituire nel log l’identificatore unico al nome dell’utente

- generare una tabella che mantiene l’informazione di quale identificatore è stato associato a ciascun utente

- eliminare da tutti i log il campo “Utente coinvolto”, che normalmente contiene la stringa “-“, ma potrebbe contenere anch’esso il nome dell’utente

Il programma infine deve salvare la lista di log anonimizzata nello stesso formato di partenza (json) e la tabella che associa ciascun utente al suo identificatore unico (anche la tabella va salvata in formato json). Il file di log anonimizzato va salvato come lista di log, e ciascun log come una lista di informazioni.

N.B. Non è necessario sviluppare subito tutte le funzioni richieste. Procedere anche per gradi includendo inizialmente una parte delle funzioni. Per esempio, si potrebbe iniziare a creare solo la tabella che associa utenti e identificatori unici senza procedere ad anonimizzare i log.

Commenti

Soluzione astratta

1. *Leggere la lista dei log in formato json e memorizzarla in una struttura dati python*
2. *Estrarre la lista degli utenti*
3. *Assegnare un ID unico a ciascun utente (creare una tabella/funzione/codifica)*
4. *Sostituire in ciascun log l’ID unico al nome utente ed eliminare il campo “utente coinvolto”*
5. *Salvare su file la lista dei log anonimizzata e la tabella*

Estensione della soluzione.

Introduzione di sottoprogrammi tenendo conto che vogliamo successivamente:

* Generare UUID migliori
* Ridurre dipendenze dai path dei file (usare parametri da linea di comando)
* Poter anonimizzare un secondo file di log e per gli stessi utenti voler usare gli stessi ID. Come si può generalizzare la soluzione?
* Poer effettuare la de-anonimizzazione

1. La struttura interna del file .json non è specificata e va determinata esaminadolo visivamente. [↑](#footnote-ref-1)