

Lista Attività A.A. 2019/2020 – MongoDB

Effettuare a scelta una delle seguenti attività e inviare via mail la consegna al docente (riccardo.martoglia@unimore.it)

1. Modellazione e interrogazione su dati reali

Partendo da uno dei dataset reali Airbnb presenti alla pagina <http://insideairbnb.com/get-the-data.html>, convertirne in JSON la struttura modellandola in modo opportuno (manualmente su alcuni esempi o attraverso uno script python) (è possibile considerare anche solo una porzione dei campi originari). Quindi, inserire i dati in una collezione e interrogarli creando 5 interrogazioni di crescente difficoltà, incluse interrogazioni di aggregazione.

Consegna: PDF commentato con dati di esempio, codice interrogazioni ed eventualmente codice Python per conversione (includere dati e codici anche in un file .txt per facilitarne il testing)

2. Indexing e query optimization

Approfondire l'argomento studiando il capitolo 8 del libro "MongoDB in action" disponibile sul sito dell'insegnamento. Riassumere quindi alcune delle tecniche imparate e mostrarne un'applicazione pratica su 6 nuove interrogazioni sul dataset Amazon utilizzato nelle esercitazioni (decidere gli indici adeguati ottimizzandone l'esecuzione e commentare le scelte effettuate).

Consegna: PDF commentato con breve relazione e codice interrogazioni / creazione degli indici (includere dati e codice anche in un file .txt per facilitarne il testing)

3. MongoDB Usecase e Python

Studiare lo usecase "Inventory management" (<https://docs.mongodb.com/ecosystem/use-cases/>) nella documentazione di MongoDB. Implementare l'acquisto a carrello (con relativo aggiornamento delle quantità e del campo "bought_together" dei prodotti acquistati) su alcuni dati di esempio, adeguatamente modificati, del dataset Amazon utilizzato nelle esercitazioni.

Consegna: PDF commentato con dati di esempio e codice python (includere dati e codice anche in un file .txt per facilitarne il testing)

NOTE

Per le attività 2 e 3 è possibile lavorare su una porzione a piacere del dataset, facendo riferimento al link pubblicato nelle slide di laboratorio e/o estraendo i dati direttamente dal server. Per comodità è anche possibile scaricare la porzione di catalogo sui video game direttamente dal seguente link:

www.isgroup.unimore.it/files/meta_Video_Games.json.gz

Per quanto riguarda il codice Python, è possibile (e gradito) produrre e consegnare un notebook jupyter .ipynb (<https://jupyter.org/>) invece di codice .py e relativi commenti su PDF.

Scadenza per premio partecipazione: **7 novembre 2019**