

Laboratorio di Basi di Dati

Progetto

I progetti di fine corso si ispirano sempre ad esigenze reali. La specifica informale del problema data nei paragrafi seguenti è, come in ogni caso reale, incompleta e, in alcuni punti, ambigua o contraddittoria. Lo studente dovrà quindi raffinare e disambiguare le specifiche mediante l'interazione con il committente. In alcuni casi allo studente sarà richiesto di valutare diverse possibili alternative, per poi sceglierne una in maniera motivata. Le motivazioni di tutte le scelte interpretative, progettuali e implementative andranno sempre chiaramente documentate nel progetto e verranno discusse in sede di esame.

Nota: alcune delle funzionalità richieste dalla specifica potrebbero non essere realizzabili con singole query, ma richiedere l'uso di strumenti più avanzati messi a disposizione dal DBMS, come le procedure. In ogni caso, tali procedure avrebbero una o più query come parte principale. L'uso di queste caratteristiche avanzate aumenta notevolmente il valore di un progetto. Tuttavia, nel caso si decida di non utilizzarle nell'implementazione, è necessario comunque presentare lo pseudocodice corrispondente, e realizzare completamente le query relative.

Specifiche

La Biblioteca rappresenta un semplice gestore di catalogo bibliografico online. Il catalogo conterrà una serie di pubblicazioni, ciascuna delle quali caratterizzata da un titolo, una lista di autori, un editore, una serie di metadati (vedi dopo) e, opzionalmente, una descrizione testuale (sunto del contenuto o presentazione della pubblicazione), un indice (composto dai titoli dei vari capitoli/sezioni della pubblicazione, numerati e ordinati) e una serie di sorgenti tramite le quali poter accedere alla pubblicazione o a parti di essa. Ciascuna sorgente sarà caratterizzata da un tipo, una URI, da un formato e da una descrizione. Esempi di sorgenti valide potrebbero essere i seguenti:

- tipo="immagine", URI="http://server.net/cover.jpg", formato="image/jpeg", descrizione="copertina"
- tipo="download", URI="http://server.net/book.pdf", formato="application/pdf", descrizione="versione elettronica gratuita"
- tipo="acquisto", URI="http://www.amazon.it/xyz", formato="cartaceo, copertina rigida", descrizione="acquista online"

Per quanto riguarda invece i metadati, questi saranno costituiti (almeno) dalle seguenti informazioni: codice ISBN, numero di pagine, lingua, data di pubblicazione, lista delle ristampe (numero e data) e parole chiave identificative (zero o più).

Ciascuna pubblicazione potrà essere associata a un certo numero di recensioni testuali, le quali dovranno anche indicare l'utente che le ha inserite e la data/ora dell'inserimento. Le recensioni saranno moderate, quindi è necessario prevedere un flag che indichi se la recensione è stata approvata oppure no dai moderatori. Infine, gli utenti potranno anche assegnare il loro like alle pubblicazioni, che quindi dovranno essere associate anche alla lista dei like ricevuti, con utente e data.

Sarà inoltre necessario prevedere opportune strutture per immagazzinare i dati dell'utenza. Il sistema, infatti, prevede due tipologie di utenza: attiva e passiva. Gli utenti potranno registrarsi nel sito fornendo i loro dati anagrafici, un indirizzo email (che verrà usato anche come username) e, ovviamente, la password scelta. Il livello di utenza iniziale sarà sempre quello passivo.

Per tener traccia delle modifiche effettuate alla bibliografia dai vari utenti attivi, il sistema dovrà mantenere una "storia" di ciascuna scheda bibliografica. Ogni entry di tale storia conterrà un timestamp, il nome dell'utente e una descrizione della modifica. Potete scegliere voi il dettaglio di questa descrizione: di base, sono accettabili anche descrizioni del tipo "ha modificato la pubblicazione", "ha inserito la pubblicazione", "ha approvato una recensione dell'utente X", ecc.

Ci sono indubbiamente molti vincoli che possono essere applicati ai contenuti di questa base di dati. Ad esempio, non dovrebbe essere possibile per un utente inserire più di una recensione per la stessa pubblicazione. L'individuazione dei vincoli e la loro implementazione (con vincoli sulle tabelle, trigger o quantomeno definendo il codice e le query necessari a effettuarne il controllo) costituiscono un requisito importante per lo sviluppo di un progetto realistico, e ne verrà tenuto conto durante la valutazione finale.

Operazioni da realizzare

Di seguito sono illustrate schematicamente le operazioni previste sulla base di dati, ciascuna da realizzare tramite una query (o, se strettamente necessario, tramite più query, opzionalmente racchiuse in una stored procedure). Ovviamente, ogni ulteriore raffinamento o arricchimento di queste specifiche aumenterà il valore del progetto.

1. Modifica del livello di un utente (da attivo a passivo e viceversa).
2. Estrazione elenco delle ultime dieci pubblicazioni inserite.
3. Estrazione elenco delle pubblicazioni aggiornate di recente (ultimi 30 giorni).
4. Estrazione elenco degli utenti più "collaborativi" (cioè quelli che hanno inserito più pubblicazioni).
5. Estrazione elenco delle pubblicazioni inserite da un utente.
6. Estrazione catalogo, cioè elenco di tutte le pubblicazioni con titolo, autori, editore e anno di pubblicazione, ordinato per titolo.
7. Estrazione dati completi di una pubblicazione specifica dato il suo ID.
8. Ricerca di pubblicazioni per ISBN, titolo, autore, e parole chiave.

9. Inserimento di una recensione relativa a una pubblicazione.
10. Approvazione o di una recensione (da parte del moderatore).
11. Inserimento di un like relativo a una pubblicazione.
12. Calcolo numero dei like per una pubblicazione.
13. Estrazione elenco delle recensioni approvate per una pubblicazione.
14. Estrazione elenco delle recensioni in attesa di approvazione.
15. Estrazione log delle modifiche effettuate su una pubblicazione.
16. Estrazione elenco delle pubblicazioni per le quali è disponibile un download.
17. Estrazione della lista delle pubblicazioni in catalogo, ognuna con la data dell'ultima ristampa
18. Data una pubblicazione, restituire tutte le pubblicazioni del catalogo aventi gli stessi autori

E' possibile inserire procedure di gestione aggiuntive che si ritengono utili.

Tecnologie da utilizzare

Il DBMS da utilizzare per la realizzazione del progetto è MySQL. MariaDB è un'alternativa accettabile.

Opzionalmente, il database realizzato potrà essere dotato di un'interfaccia scritta con un linguaggio di programmazione a scelta (Java o PHP) tramite la quale invocare le query richieste (fornendone gli eventuali parametri) e visualizzarne i risultati. Il progetto può essere integrato con il progetto del corso Object-Oriented Software Design. Se interessati contattare i responsabili dei due corsi.

Svolgimento e Documentazione del Progetto

Le specifiche fornite potrebbero non risultare esaustive o completamente definite. Ogni funzionalità aggiunta o raffinata, anche tramite l'interazione con il committente, sarà adeguatamente valutata. Tutte le scelte progettuali vanno comunque discusse e motivate.

Il progetto dovrà essere svolto secondo le seguenti fasi:

1. Specifica dei requisiti e glossario dei termini.
2. Progettazione concettuale tramite il modello Entità-Relazione.
3. Formalizzazione di tutti i vincoli non esprimibili nel modello ER.
4. Ristrutturazione e ottimizzazione del modello ER.
5. Traduzione del modello ER nel corrispondente modello relazionale.
6. Implementazione effettiva del modello relazionale e di tutti i vincoli tramite SQL.
7. Implementazione delle query, procedure, funzioni, ecc. tramite SQL.

Tutte le fasi del progetto dovranno essere corredate da adeguata documentazione **in formato elettronico** che illustri quanto viene realizzato e le scelte intraprese. In particolare, in riferimento ai passi elencati in precedenza, dovranno essere necessariamente inclusi nella documentazione (1) la specifica dei requisiti e il glossario dei termini, (2) gli schemi ER risultanti dai passi (2) e (4), debitamente commentati, la formalizzazione di tutti i vincoli non esprimibili nel modello ER (3) e il modello relazionale della base di dati ottenuto al passo (5), in cui siano messe in evidenza le chiavi delle varie tabelle e le relazioni tra queste ultime.

Il database finale, risultato dei passi (6) e (7), dovrà essere consegnato nella forma di uno script SQL contenente la struttura del database (istruzioni CREATE, comprese eventuali procedure, funzioni e viste), dei dati di prova (istruzioni INSERT) e il codice delle query necessarie a realizzare le funzionalità richieste.

Suggerimento: potete usare le funzioni di esportazione presenti in phpMyAdmin e MySQL workbench oppure direttamente il comando mysqldump per esportare un dump del database contenente sia la struttura che i dati. Inoltre, per praticità, nella maggior parte dei casi potrete incorporare la anche definizione delle query nel dump del database, definendole come viste o stored procedures con nomi significativi come "Query_1", ecc.

Infine, per ciascuna delle funzionalità richieste dovrà essere fornito un esempio di esecuzione delle relative query sui dati di esempio forniti, mostrandone i risultati.

Le parti della specifica marcate come *opzionali*, se omesse, non renderanno il progetto insufficiente ma non gli permetteranno comunque di raggiungere il massimo dei voti. Nel caso si decida di realizzarle, non sarà necessario che siano perfette o complete, ma che dimostrino chiaramente il vostro impegno nell'affrontare una tematica avanzata.

Nel caso di gruppi di lavoro composti da più componenti, *il contributo effettivo offerto da ciascun componente* alla realizzazione finale deve essere descritto nella documentazione (indicando, ad esempio, chi si è dedicato prevalentemente alla programmazione server, chi ha realizzato il layout, chi ha programmato lato client, ecc.). In sede di esame, i responsabili potranno essere chiamati a riferire sugli aspetti loro delegati.

Valutazione del Progetto

La documentazione e il codice del progetto descritti nella sezione precedente andranno consegnati via moodle entro la data prevista per l'appello di esame. La discussione orale del progetto sarà concordata di conseguenza.

In sede di esame saranno discussi eventuali problemi riscontrati nel progetto e potrà essere richiesta una dimostrazione del funzionamento di alcune query e/o

dell'eventuale interfaccia realizzata (è quindi opportuno portare con sé un portatile con il progetto installato e funzionante).

Nel valutare il progetto consegnato saranno prese in considerazione le seguenti caratteristiche:

1. Rispetto delle specifiche e loro corretto affinamento, ove necessario.
2. Correttezza e conformità alle specifiche dei modelli realizzati.
3. Correttezza tecnica dell'implementazione fisica del database (struttura, vincoli).
4. Correttezza tecnica delle query realizzate e loro aderenza alle funzionalità richieste.
5. Consistenza tra i vari artefatti prodotti durante lo sviluppo del progetto, ad esempio tra requisiti, glossario dei termini, schema E/R, implementazione, etc.
6. Adeguatezza della documentazione.

A questa valutazione si aggiungerà quella generale derivata dalla discussione del progetto in sede d'esame.

Sarà tenuto in considerazione anche la auto valutazione e la valutazione degli altri membri del gruppo.