SpotifyDB

Nicola Baesso – Gabriel Rovesti

(max 15 pagine)

Abstract

Spotify è un’applicazione di streaming musicale on-demand che permette in qualsiasi momento di riprodurre brani, album e vari tipi di contenuti non musicali (i.e. podcast). A quest’applicazione sono registrati degli utenti, che possono usufruire gratuitamente i contenuti della piattaforma oppure previo pagamento mensile con alcuni bonus (i.e. nessuna pubblicità, saltare liberamente l’esecuzione dei brani, ecc.). In generale, gli utenti possono creare delle proprie playlist ì, le quali possono essere private o collaborative; a loro volta, possono o meno pubblicare contenuti musicali e non (in qualità di artista e/o presentatore).

Il servizio offre anche la possibilità di poter pubblicare i contenuti (siano essi musicali o meno) prodotti in qualità di artisti i quali, naturalmente, possono essere solisti oppure far parte di un gruppo. In merito ai contenuti offerti, nella piattaforma contiene principalmente contenuti musicali di vario genere e non musicali. In merito ai brani offerti, questi possono essere rilasciati come singoli oppure far parte di album e possono posizionarsi, a seconda del loro ascolto e successo, in varie posizioni in classifica, differenziata per paese o regione geografica.

A fronte di questa descrizione, l’azienda omonima richiede lo sviluppo del database per garantire la gestione ottimale del servizio.

Analisi dei requisiti

Si vuole rappresentare la gestione della piattaforma Spotify tramite una base di dati. *A fini dimostrativi, vengono mostrate le funzionalità della stessa tramite dati verosimili autogenerati.*

Un **artista** è caratterizzato dal nome d’arte, dalla data di iscrizione, dalla nazionalità e da una descrizione dell’artista stesso. Due artisti potrebbero condividere lo stesso nome d’arte, è quindi necessario associare ad ogni artista un identificatore per permetterne la distinzione. Un artista può essere un gruppo o un presentatore, se non rientra in queste due categorie è da intendersi come un artista solista.

Di un gruppo, oltre ai campi già citati, è rilevante conoscerne il numero di componenti. Per un presentatore non sono rilevanti ulteriori informazioni.

Un **utente** è caratterizzato dall’username, autogenerato di 11 cifre e che funge da identificatore, dal nome, dal cognome, dalla data di nascita, dal sesso che può assumere tre valori (uomo, donna, non binario), dall’indirizzo e-mail e dalla password. La base di dati non deve consentire l’inserimento di un utente con etá (ricavabile dalla data di nascita) minore di 13 anni. Un utente si distingue categoricamente tra utenti con piano gratuito (da qui in avanti, utente gratuito) e utenti con piano premium (da qui in avanti, utente premium). In merito agli abbonamenti, si vuole conoscere il tipo, il prezzo e il numero dell’account associato. Qualora l’utente premium abbia un abbonamento di tipo Famiglia, associa più utenti allo stesso abbonamento (massimo 6, nel caso la piattaforma decida di aumentarli, la base di dati ne permette l’espansione). Ciascun utente è in grado di crearsi delle **playlist**, ciascuna identificata da un nome e da un certo numero di brani.

Un **brano** é caratterizzato da un nome (unico all’interno dell’artista che lo ha composto), uno o piú generi di riferimento, una durata specifica e se può essere riproducibile o meno (cioè se un brano non é disponibile in un certo mercato discografico o non é più riproducibile per motivi legali/di diritto d’autore. Per la base di dati un brano è non riproducibile se almeno una delle condizioni fornite è verificata). Inoltre, i brani possono avere delle collaborazioni (*featurings*) e si intende fare una distinzione tra l’artista principale e gli artisti che vi collaborano.

Essi possono essere rilasciati come singoli (non facenti parte di album), oppure essere parte di un **album**, di cui sono rilevanti il genere (o i generi), una durata e il nome che lo identifica. Gli artisti (distinti tra artisti singoli, presentatori di podcast oppure facenti parte di un gruppo, di cui si intende salvare il numero di componenti), identificati da un nome d’arte e dalla data di iscrizione della piattaforma e con una loro nazionalità, possono essere sotto contratto con al più un’**etichetta discografica,** ciascuna identificata da un nome e con una nazione di riferimento. Non sono rilevanti informazioni come la data di fondazione poiché dati che non portano valore al servizio offerto dall’azienda. Se un artista non possiede una casa discografica di riferimento, avrà un’etichetta definita come “indipendente”.

Sia che gli artisti siano indipendenti, sia che siano sotto contratto presso un’etichetta discografica, si intende memorizzare quando avviene il **rilascio** di un brano oppure di un album, ciascuno identificato da una data (unica all’interno della casa discografica e dell’artista che lo produce).

Quando i brani vengono rilasciati, essi possono entrare a far parte di una **classifica**, identificata semplicemente dal nome ottenendo, nel caso risulti un interesse da parte del pubblico, un posizionamento certificato (in una classifica possono entrare più canzoni di uno stesso album e uno stesso brano può apparire in più classifiche).

In merito agli artisti registrati, come citato, la piattaforma registra i dati anche dei presentatori di **podcast**, i quali possono far parte di generi diversi, hanno un certo numero di puntate, un nome identificativo e almeno un presentatore.

Glossario dei termini

|  |  |  |
| --- | --- | --- |
| **Termine** | **Descrizione** | **Collegamenti** |
| Artista | L’artista singolo/il gruppo/il presentatore di podcast memorizzato nella piattaforma | Brano, Etichetta Discografica |
| Utente | Iscritto alla piattaforma, che possiede o meno un abbonamento | Playlist, Collaborativa (utenti diversi possiedono varie playlist collaborative) |
| Playlist | Liste di canzoni create dalla piattaforma o dagli utenti stessi e fruibili a tutti | Utente, Brano |
| Classifica | Lista quantitativa che segna il successo in un certo momento di ascolti di brani/album | Brano |
| Brano | Prodotto del lavoro di un artista (singolo/gruppo), rilasciato sul mercato e curato da un’etichetta | Rilascio, Artista (sia principale che eventuali artisti collaboratori), Playlist, Classifica, Album |
| Album | Disco che viene rilasciato da parte di un’artista indipendente o da una casa discografica. Contiene un insieme di brani | Rilascio |
| Abbonamento | Contratto sottoscritto dall’utente premium su base mensile che fornisce bonus di utilizzo sugli utenti free | Premium |
| Premium | L’utente che paga un abbonamento e usufruisce di bonus sulla piattaforma | Abbonamento |
| Rilascio | Lancio nel mercato discografico di un brano/album da parte di un’artista indipendente o sotto una casa discografica | Album, Brano, Etichetta Discografica |
| Etichetta discografica | Aziende che si occupano della distribuzione/promozione e produzione di brani/album per artisti singoli/gruppi | Rilascio, Artista |
| Artista indipendente | Un artista che non ha casa discografica e produce per conto suo brani/album | Attributo di Artista |
| Podcast | Contenuti non musicali di vario genere (trasmissioni radio diffuse via Internet) che possiedono ospite/i e presentatore/i | Presentatore |

Operazioni tipiche

Le operazioni descritte si distinguono tra quelle svolte dall’artista stesso (se è indipendente o produttore di altri contenuti), le pubblicazioni fatte dalla casa discografica, le operazioni svolte dagli utenti finali e le azioni svolte dai gestori della piattaforma.(1)

|  |  |
| --- | --- |
| **Tipo di operazione** | **Frequenza** |
| Inserimento di nuove tracce/album | 60000 al giorno |
| Registrazione di nuovi utenti | 40000 al giorno |
| Inserimento di nuovi podcast | 140 al giorno |
| Fatturazione per i clienti premium | 1000000 al mese |
| Aggiornamento delle classifiche | 2200 al giorno |

(1) I dati riflettono quelli della piattaforma di ispirazione (dati reali) o frutto di calcoli concreti sulla piattaforma. Fonte: <https://www.businessofapps.com/data/spotify-statistics/>

Progettazione concettuale

Entità

|  |  |  |
| --- | --- | --- |
| **Artista** | | |
| NomeArte | varchar(50) primary key | Il nome con cui è conosciuto l’artista |
| DataIscrizione | date primary key | La data d’iscrizione dell’artista alla piattaforma |
| Descrizione | varchar(200) | Descrizione dell’artista |
| Nazionalità | Char(2) | La nazionalità dell’artista, espresso con il codice a 2 cifre |

Artista si può specializzare in Gruppo o in Presentatore.

|  |  |  |
| --- | --- | --- |
| **Gruppo** | | |
| NumeroComponenti | int | Il numero di componenti di cui è formato il gruppo |
|  |  |  |
| **Presentatore** | | |

|  |  |  |
| --- | --- | --- |
| **Podcast** | | |
| NomePodcast | varchar(50) primary key | Il nome del podcast |
| Genere (1,N) | varchar(200) | I generi che meglio descrivono un podcast |
| NumeroPuntate | Int | Numero di puntate del podcast |

|  |  |  |
| --- | --- | --- |
| **EtichettaDiscografica** | | |
| NomeEtichetta | varchar(50) primary key | Il nome dell’etichetta discografica |
| Nazione | Char(2) | La nazione dove l’etichetta ha sede, espresso con il codice a 2 cifre |

|  |  |  |
| --- | --- | --- |
| **Rilascio** | | |
| Data | date primary key | La data del rilascio |

|  |  |  |
| --- | --- | --- |
| **Album** | | |
| NomeAlbum | varchar(50) primary key | Il nome dell’album |
| Durata | Int | La durata dell’album in secondi |
| Genere (1,N) | varchar(200) | I generi presenti nell’album |

|  |  |  |
| --- | --- | --- |
| **Brano** | | |
| NomeBrano | varchar(50) primary key | Il nome con cui è conosciuto l’artista |
| Genere (1,N) | varchar(200) | I generi che meglio classificano il brano |
| Durata | int | La durata del brano in secondi |
| Riproducibile | bool | Flag che indica se il brano è riproducibile o meno |

|  |  |  |
| --- | --- | --- |
| **Classifica** | | |
| NomeClassifica | varchar(50) primary key | Il nome della classifica |

|  |  |  |
| --- | --- | --- |
| **Playlist** | | |
| NomePlaylist | varchar(50) primary key | Il nome della playlist |
| NumeroBrani | int | Il numero di brani presenti nella playlist |

Playlist si specializza in Privata o Collaborativa

|  |
| --- |
| **Privata** |

|  |
| --- |
| **Collaborativa** |

|  |  |  |
| --- | --- | --- |
| **Utente** | | |
| Username | char(11) primary key | Lo username, autogenerato con 11 cifre (da 0 a 9) |
| Nome | varchar(30) | Il nome dell’utente |
| Cognome | varchar(30) | Il cognome dell’utente |
| DataNascita | date | La data di nascita dell’utente |
| Sesso | char(2) | Il sesso dell’utente tra Uomo (M), Donna (F), Non Binario (NB) |
| Email | varchar(50) | L’indirizzo email associato all’utente |
| Password | varchar(50) | La password dell’utente |

Utente si specializza in Gratuito o Premium

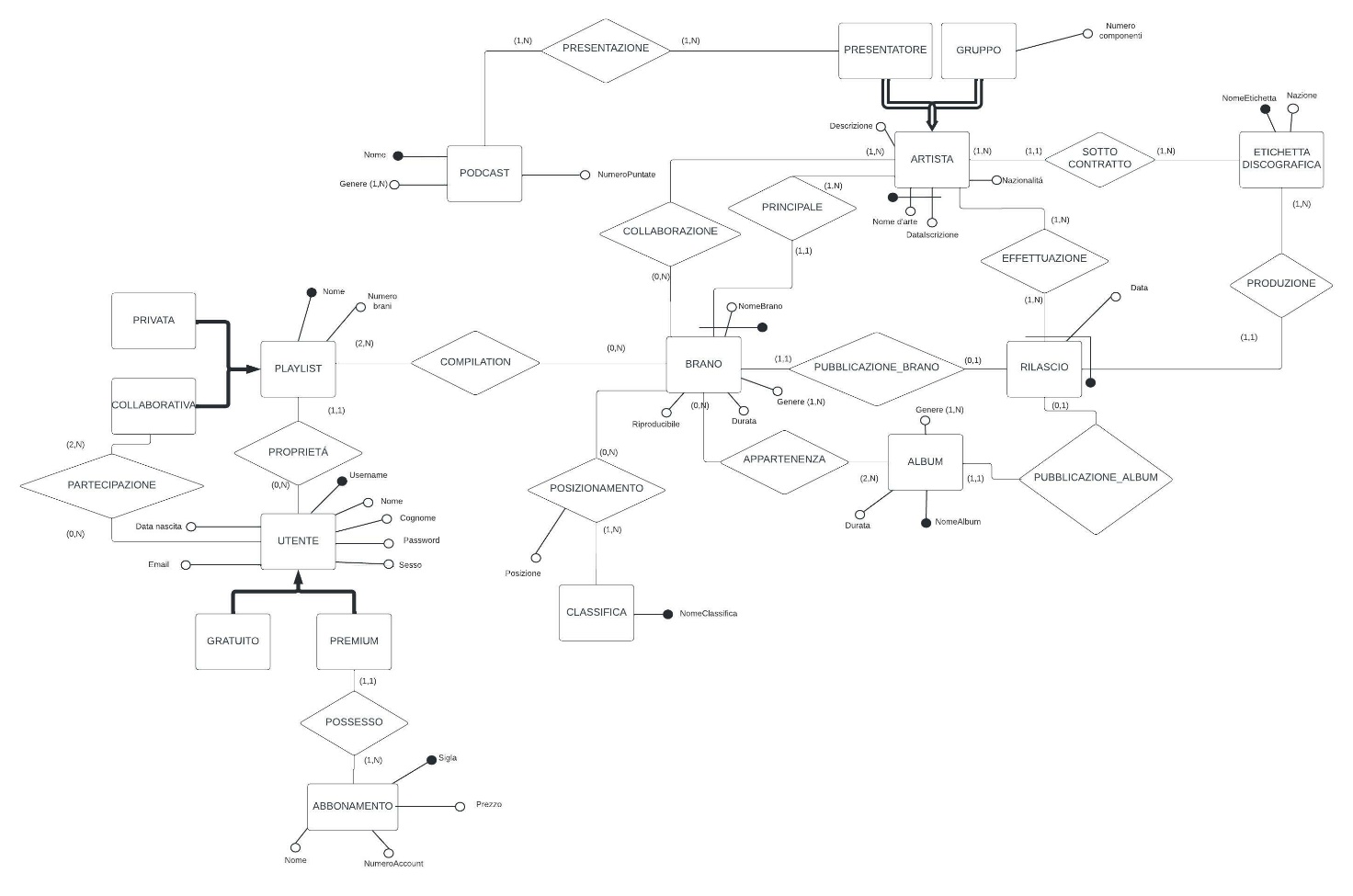
|  |
| --- |
| **Gratuito** |

|  |
| --- |
| **Premium** |

|  |  |  |
| --- | --- | --- |
| **Abbonamento** | | |
| Sigla | char(30) primary key | La sigla con cui viene identificato l’abbonamento |
| Tipo | varchar(20) | Il tipo di abbonamento |
| NumeroAccount | int | Il numero massimo di account che possono usufruire dell’abbonamento |
| Prezzo | double | Il prezzo dell’abbonamento |

Relazioni e cardinalità

|  |  |  |
| --- | --- | --- |
| **Relazione** | **Entità coinvolte** | **Descrizione** |
| Partecipazione | Collaborativa (2,N)  Utente (0,N) | Una playlist collaborativa ha almeno due utenti che collaborano.  Un utente puó collaborare in piú playlist collaborative. |
| Possesso | Premium (1,1)  Abbonamento (1,N) | Un utente premium ha un solo abbonamento.  Un abbonamento è associato ad almeno una persona. |
| Posizionamento | Brano (0,N)  Classifica (1,N) | Un brano puó apparire in piú classifiche.  Una classifica ha almeno un brano al suo interno. |
| Proprietà | Playlist (1,1)  Utente (0,N) | Una playlist è legata ad un solo utente.  Un utente puó avere piú playlist. |
| Compilation | Playlist (2,N)  Brano (0,N) | Una playlist ha almeno due brani.  Un brano puó far parte di piú playlist. |
| Appartenenza | Brano (0,N)  Album (2,N) | Un brano puó appartenere a piú album.  Un album ha almeno due brani. |
| Pubblicazione\_Brano | Brano (1,1)  Rilascio (0,1) | Un brano ha un solo rilascio.  Un rilascio puó riguardare un brano. |
| Pubblicazione\_Album | Album (1,1)  Rilascio (0,1) | Un album ha solo un rilascio.  Un rilascio puó riguardare un album. |
| Produzione | Rilascio (1,1)  EtichettaDiscografica (1,N) | Un rilascio è legato ad una sola etichetta discografica.  Un’etichetta discografica ha almeno un rilascio. |
| Effettuazione | Artista (1,N)  Rilascio (1,1) | Un artista effettua almeno un rilascio.  Un rilascio è legato ad un solo artista. |
| Collaborazione | Artista (1,N)  Brano (0,N) | Un artista collabora in almeno un brano.  Un brano può avere degli artisti collaboratori. |
| Principale | Artista (1,N)  Brano (1,1) | Un artista è l’artista principale di almeno un brano.  Un brano ha un solo artista principale. |
| Presentazione | Podcast (1,N)  Presentatore (1,N) | Un podcast ha almeno un presentatore.  Un presentatore presenta almeno un podcast. |
| Sotto Contratto | Artista (1,1)  EtichettaDiscografica (1,N) | Un artista ha una sola etichetta discografica.  Un’etichetta discografica ha almeno un’artista sotto contratto. |

Schema ER

Vincoli non esprimibili in ER

* Il rilascio riguarda esclusivamente un album oppure un brano (e non entrambi)
* Se l’abbonamento è di tipo famiglia, può essere collegato ad altre persone, altrimenti non deve essere collegato a più di una persona

Progettazione logica

Analisi delle ridondanze e ristrutturazione

Eliminazione delle generalizzazioni

Scelta degli identificatori primari

Diagramma ER ristrutturato

Schema relazionale e vincoli di integrità referenziale

Query SQL e indice

Codice C++

g++ main.cpp -L dependencies/lib -lpq -o db -std=c++11