



Corso di Laurea in Ingegneria Informatica
Università di Pisa
Anno Accademico 2023/2024

Progetto di Basi Di Dati

Studente: **Federico Massini**

Indice

1	Introduzione	1
2	Progettazione Concettuale	3
2.1	Dizionario delle Entità	3
2.2	Dizionario delle Relationship	6
2.3	Approfondimento Dizionari	9
2.3.1	Analisi Entità Figlie	9
2.3.2	Analisi Cardinalità Relazioni	9
3	Ristrutturazione schema E-R	10
3.1	Gestione delle Generalizzazioni	10
3.1.1	Operatore	10
3.1.2	Lingua	10
3.1.3	Fattura	10
3.1.4	Formato	11
3.2	Gestione attributi composti	11
3.2.1	Intestatario	11
3.3	Altre attività	11
3.4	Tavola dei volumi	12
4	Operazioni sui dati	15
4.1	Generi di un Film	15
4.1.1	Porzione della tavola dei volumi interessata	16
4.1.2	Tavola degli Accessi	16
4.1.3	Accessi Giornalieri	16
4.2	Nuova Connessione	17
4.2.1	Porzione della tavola dei volumi interessata	18
4.2.2	Tavola degli Accessi	18
4.2.3	Accessi Giornalieri	18
4.3	Inserimento Fattura	19
4.3.1	Porzione della tavola dei volumi interessata	19
4.3.2	Tavola degli Accessi	20
4.3.3	Accessi Giornalieri	20
4.4	Lista Film disponibili per Lingua	21
4.4.1	Porzione della tavola dei volumi interessata	21
4.4.2	Tavola degli Accessi	22
4.4.3	Accessi Giornalieri	22
4.5	Film con Registi Premiati	23
4.5.1	Porzione della tavola dei volumi interessata	23
4.5.2	Tavola degli Accessi	24
4.5.3	Accessi Giornalieri	24
4.6	Film vietati in base al Formato	25
4.6.1	Porzione della tavola dei volumi interessata	25

4.6.2	Tavola degli Accessi	26
4.6.3	Accessi Giornalieri	26
4.7	Valutazione Film	27
4.7.1	Porzione della tavola dei volumi interessata	28
4.7.2	Tavola degli Accessi	29
4.7.3	Accessi Giornalieri	30
4.8	Raccomandazione Contenuti	31
4.8.1	Porzione della tavola dei volumi interessata	32
4.8.2	Tavola degli Accessi	32
4.8.3	Accessi Giornalieri	33
4.9	Analisi ridondanze	34
4.9.1	PremiatoP	34
4.9.2	Valutazione	38
4.10	Aggiornamento dizionario Entità	41
4.11	Aggiornamento dizionario Relazioni	42
5	Area Analytics	43
5.1	Classifiche	43
5.1.1	Porzione della tavola dei volumi interessata	44
5.1.2	Tavola degli Accessi	45
5.1.3	Accessi Giornalieri	47
6	Progettazione Logica	48
6.1	Traduzione in Schema Logico	48
6.2	Vincoli di Integrità	49
6.2.1	Vincoli di Integrità Referenziale	49
6.2.2	Vincoli Intra-Relazionali	51
6.2.3	Vincoli Inter-Relazionali	52
6.3	Dipendenze Funzionali	53
6.4	Normalizzazione	55
6.5	Trattazione vincoli	56
6.6	Events	57
6.7	Trattazione Ridondanze	57
6.8	Popolamento	57

1 Introduzione

Il database che si desidera progettare su DBMS Oracle MySQL ha lo scopo di memorizzare i dati relativi a un sistema che gestisce FilmSphere, un servizio dedicato alla memorizzazione e allo streaming di contenuti video online. Il database sarà inoltre dotato di operazioni e funzionalità di data analytics. Il portale tiene traccia di ogni informazione d'interesse riguardante un utente, dai dati personali ai dati relativi alle visualizzazioni, permettendo di raccomandare ad esso contenuti in base alla sua "history" di visualizzazione. FilmSphere memorizza anche informazioni relative alle caratteristiche di video e audio, come la codifica e le dimensioni, fornendo informazioni anche sulle modalità di connessione tra utente e piattaforma geolocalizzando l'utente ad ogni suo accesso.

Ad ogni area indicata nelle specifiche corrisponde un colore nel diagramma E-R¹, qui di seguito sono riportate le informazioni.



Figura 1: Legenda presa dal diagramma E-R

Generalmente valgono questi colori sia per entità sia per relationship; ci sono però delle eccezioni per le quali ho pensato risultasse meglio rappresentare il collegamento tra due aree diverse "mescolando i colori". Qui di seguito sono riportate le figure (sono raffigurate come relationship in quanto la maggior parte delle eccezioni sono per l'appunto relazioni, nello schema è però presente anche un'entità di questo tipo).

¹Struttura del modello Entity-Relationship. L'E-R è un modello concettuale proposto nel 1976 che si è ormai affermato nelle metodologie di progetto.

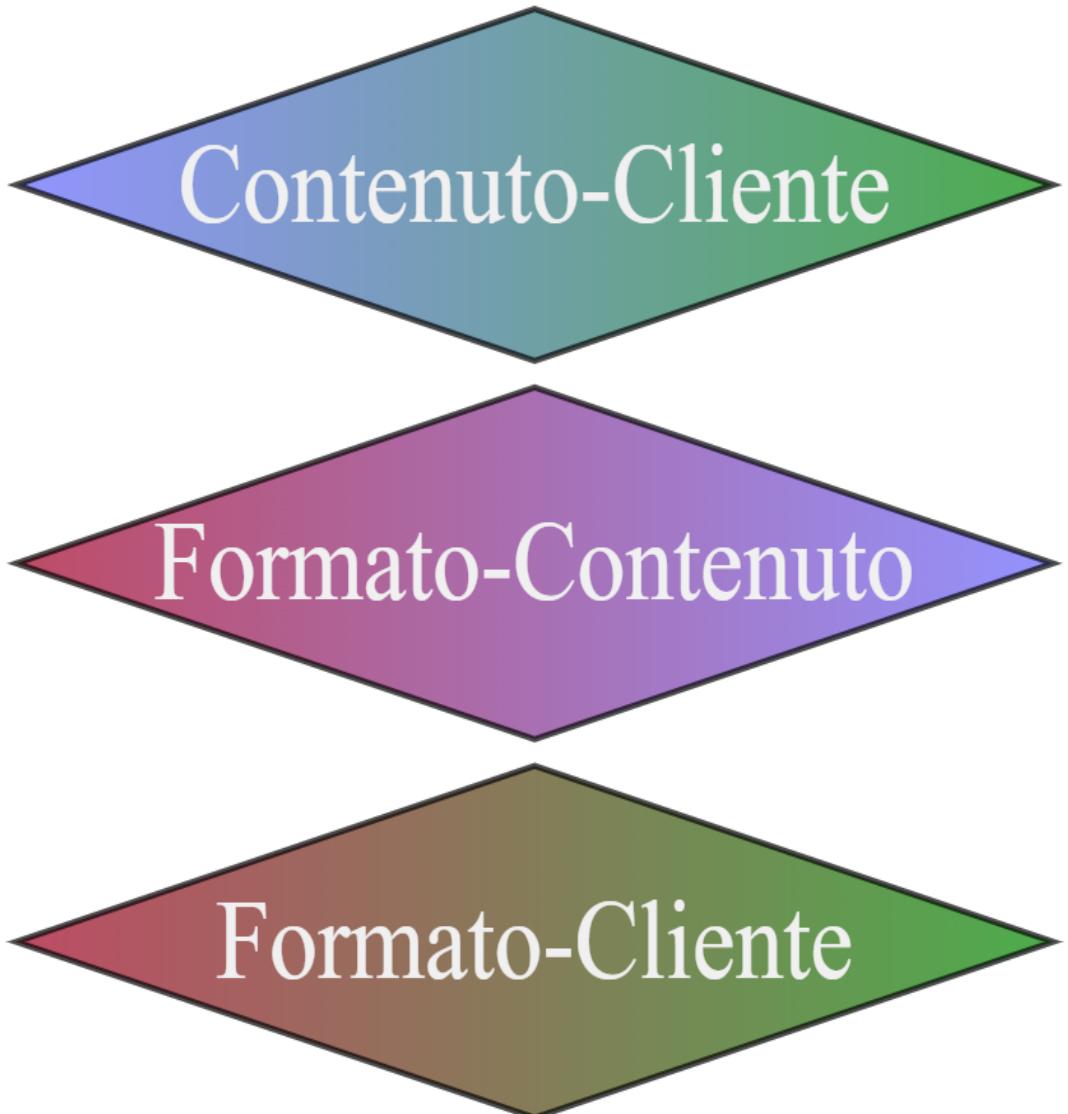


Figura 2: Legenda presa dal diagramma E-R

Nei prossimi paragrafi verranno illustrate più nello specifico varie caratteristiche del progetto.

2 Progettazione Concettuale

In questo capitolo verranno illustrati il dizionario delle Entità e il dizionario delle Relationship del diagramma E-R prima della ristrutturazione.

2.1 Dizionario delle Entità

Il dizionario delle Entità verrà descritto attraverso una tabella strutturata nel seguente modo:

Nome Entità	Attributi	Sinonimi	Descrizione
-------------	-----------	----------	-------------

Ad ogni inizio pagina verrà ripetuta l'intestazione della tabella per migliorare la visibilità.

Legenda:

Chiavi: Le chiavi verranno indicate con il testo in grassetto.

* : Con il simbolo * verranno indicate le entità padri di una generalizzazione, le entità figlie saranno menzionate nella descrizione e verranno trattate nel paragrafo 2.3.

(): Gli attributi composti verranno inseriti tra parentesi tonde.

Nome Entità	Attributi	Sinonimi	Descrizione
Film	ID , Durata, Descrizione, Titolo	Pellicola	Contenuto disponibile nella piattaforma Streaming
Operatore*	ID , Nome, Cognome, Popolarità	Organico	Addetti ai lavori facenti parte della troupe di un Film, comprendente due figure artistiche: Attore e Regista
Premio	Nome , Categoria , Peso	Vittoria	Riconoscimento conferito a Film e Operatori
Critico	ID , Nome, Cognome		Persone le cui recensioni influiscono sulla valutazione dei film
Lingua*	NomeLingua	Idioma	Lingue disponibili per l'audio e per i sottotitoli di un film

Nome Entità	Attributi	Sinonimi	Descrizione
Genere	ID , Descrizione	Categoria	Insieme di caratteristiche fondamentali che contraddistinguono una categoria di film
Paese	Nome , Longitudine, Latitudine	Area Geografica	Elenco di stati con le proprie coordinate
Utente	ID , Nome, Cognome, Email, Password, DataNascita	Cliente	Utenti registrati sulla piattaforma streaming tramite credenziali
Fattura*	Codice , DataEmissione, Quota		Documento che il venditore trasmette al compratore nel momento in cui avviene l'acquisto dell'abbonamento, si tiene traccia delle fatture già pagate e quelle non ancora pagate
Carta	PAN , CVV, DataScadenza, (Intestatario)		Carte di Credito utilizzate per effettuare pagamenti relativi ad abbonamenti
Connessione	IP , InizioConnessione , FineConnessione , Hardware	Collegamento	Indica informazioni relative alla connessione Utente-Piattaforma
Abbonamento	Tipo , Tariffa, Durata, MaxOre, MaxGB, EtaMin		Iscrizione sottoscritta da un utente per accedere a determinate modalità di visualizzazione dei contenuti

Nome Entità	Attributi	Sinonimi	Descrizione
Funzionalità	Nome, Descrizione		Insieme di funzionalità di un determinato abbonamento
Formato*	Codice, DataAggiornamento, Bitrate, DimensioneFile		Sistema di codifica dei dati di un Film riguardante formato audio e formato video
Codec	FileCodec, Specifiche		Software che si occupa della codifica e/o decodifica di un segnale
StatoRestrittivo	Stato, RestrizioneF, RestrizioneA		Insieme di stati che vietano determinati tipi di formato e/o abbonamento

2.2 Dizionario delle Relationship

Il dizionario delle Relationship verrà descritto attraverso una tabella strutturata nel seguente modo:

Relationship	Cardinalità	Attributi	Descrizione
--------------	-------------	-----------	-------------

Ad ogni inizio pagina verrà ripetuta l'intestazione della tabella per migliorare la visibilità.

In *Cardinalità* vengono esplicite le partecipazioni quantitative delle entità coinvolte, per alcune sarà opportuno giustificare i numeri inseriti, questo accorgimento verrà fatto nel paragrafo 2.3.

Relationship	Cardinalità	Attributi	Descrizione
Regia	Film (1,N) Regista (1,N)		Direzione della realizzazione di un film da parte di un Regista
Interpretazione	Film (0,N) Attore (1,N)		Presenza di un Attore in un Film
PremiazioneF	Film (0,N) Premio (0,N)	Data	Vittoria di un premio da parte di un Film
PremiazioneA	Attore (0,N) Premio (0,N)	Data	Vittoria di un premio da parte di un Attore
PremiazioneR	Regista (0,N) Premio (0,N)	Data	Vittoria di un premio da parte di un Regista
RecensioneC	Film (0,N) Critico (1,N)	Voto, Testo	Recensione di un Critico relativa alla visione di un Film in valutazione numerica e spiegazione testuale
RecensioneU	Film (0,N) Utente (0,N)	Voto, Testo	Recensione di un Utente relativa alla visione di un Film in valutazione numerica e spiegazione testuale

Relationship	Cardinalità	Attributi	Descrizione
Tipologia	Film (1,N) Genere (1,N)		Associazione di un film a uno o più generi
Linguaggio	Film (0,N) Lingua(1,N)		Lingue disponibili per doppiaggio e/o sottotitolaggio di un Film
Produzione	Film (1,1) Paese (0,N)	Anno	Informazioni riguardanti Paesi e Anno di Produzione di un Film
Specifiche	Film (1,N) Formato (1,N)	DataRilascio	Specifiche di un Film per quanto riguarda il formato audio e video
Aggiornamento	Formato (1,N) Codec (1,N)	Data	Associazione che permette di conoscere la Data di un determinato aggiornamento del Codec
Esclusione	Formato (0,N) StatoRestrittivo (0,N)		Stato che vieta un determinato tipo di formato
Restrizione	StatoRestrittivo (0,N) Abbonamento (0,N)		Stato che vieta un determinato tipo di Abbonamento
Offerta	Abbonamento (1,N) Funzionalità (1,N)		Pacchetto di Funzionalità che un Abbonamento offre
Sottoscrizione	Abbonamento (0,N) Utente (0,1)	DataInizio, DataFine	Sistema di memorizzazione delle sottoscrizioni di Abbonamenti da parte degli utenti
Possesso	Utente (0,N) Carta(1,1)		Registrazione di Carte di Credito da parte degli Utenti

Relationship	Cardinalità	Attributi	Descrizione
Fatturazione	Utente (0,N) Fattura(1,1)		Invio di una Fattura con informazioni relative all'acquisto di un Utente
Pagamento	Carta (0,N) Pagata(1,1)		Attestazione di un Pagamento di una Fattura tramite Carta di Credito
Instaurazione	Utente (1,N) Connessione (1,1)		Instaurazione di Connessioni con le quali un Utente accede a FilmSphere
Geolocalizzazione	Paese(0,N) Connessione(1,1)		Geolocalizzazione di una determinata Connessione
Visualizzazione	Film (0,N) Connessione (1,1)		Informazioni relative alla Visualizzazione di un determinato Film associata alla Connessione

2.3 Approfondimento Dizionari

In questo paragrafo verranno sciolti alcuni dubbi che potrebbero risultare dai dizionari visti nei paragrafi 2.1 e 2.2.

2.3.1 Analisi Entità Figlie

Operatore ha come entità figlie *Attore* e *Regista* per specificare il ruolo del soggetto nel Film dato. Entrambe le entità figlie non hanno attributi propri.

Lingua ha come entità figlie *Doppiaggio* e *Sottotitoli* per specificare se essa è disponibile solo in Audio, solo in Sottotitoli o entrambe. Entrambe le entità figlie non hanno attributi propri.

Fattura ha come entità figlie *Pagata* e *Non Pagata* per specificare se essa è stata pagata oppure no. L'entità Non Pagata non ha attributi propri, mentre Pagata ha come attributo proprio *DataPagamento* per tenere traccia di quando è stata pagata la Fattura.

Formato ha come entità figlie *FormatoAudio* e *FormatoVideo* per specificare il tipo di formato. FormatoAudio ha come attributo proprio *QualitàA*, mentre FormatoVideo ha come attributi propri *Risoluzione*, *QualitàV*, *RapportoAspetto*, *Lunghezza*.

2.3.2 Analisi Cardinalità Relazioni

Nella relazione **Interpretazione** ho scelto di utilizzare una cardinalità $\text{Film}(0,N)$ perché un film può non avere attori (esempio un documentario e/o un cartone animato).

Nelle relazioni riguardanti **Premio** ho deciso di usare cardinalità $\text{Premio}(0,N)$ in quanto un Premio può non esistere per una determinata categoria.

Nelle relazioni **RecensioneC** e **RecensioneU** ho scelto di utilizzare rispettivamente cardinalità di $\text{Critico}(1,N)$ e $\text{Utente}(0,N)$ perché un Critico per essere tale deve avere recensito almeno un film, un utente può anche non aver mai effettuato recensioni.

Nella relazione **Linguaggio** ho deciso di usare cardinalità $\text{Film}(0,N)$ in quanto un film può non avere alcuna lingua (esempio un film muto).

Nella relazione **Instaurazione** ho deciso di usare cardinalità $\text{Utente}(1,N)$ in quanto un Utente si è connesso almeno una volta al portale per essere registrato.

Nella relazione **Visualizzazione** ho scelto di utilizzare cardinalità $\text{Connessione}(1,1)$ assumendo che al termine della visione di un film si abbia una disconnessione dal portale.

3 Ristrutturazione schema E-R

In questo capitolo spiegherò come ho gestito la ristrutturazione dello schema, ovvero l'attività che consente di ottimizzare e semplificare il diagramma per la successiva traduzione nel modello logico.

3.1 Gestione delle Generalizzazioni

Le generalizzazioni non possono essere rappresentate direttamente nel modello relazionale², è quindi dovere del programmatore eliminarle sostituendole con entità e associazioni.

3.1.1 Operatore

La prima generalizzazione da trattare riguarda l'entità **Operatore**. È una generalizzazione Totale (ogni occorrenza di Operatore è occorrenza di almeno una tra le entità figlie) e Sovrapposta (ogni occorrenza dell'entità Operatore non è al più una delle entità figlie, possono esistere attori che lavorano anche come registi). Le entità figlie *Regista* e *Attore* sono state accorpate al genitore. A tale entità viene aggiunto un attributo **ruolo** che serve a distinguere il tipo di un'occorrenza. Si noti che verranno usati due ID diversi nel caso di operatori attori-registi.

3.1.2 Lingua

La seconda generalizzazione da trattare riguarda l'entità **Lingua**. È una generalizzazione Totale e Sovraposta. Le entità figlie *Doppiaggio* e *Sottotitoli* sono state accorpate al genitore. A tale entità viene aggiunto un attributo **tipologia** per distinguere una lingua disponibile per il sottotitolaggio o per il doppiaggio. Si noti che la chiave primaria è ora formata dai due attributi dell'entità e non più dal solo attributo **NomeLingua**.

3.1.3 Fattura

La terza generalizzazione da trattare riguarda l'entità **Fattura**. È una generalizzazione Totale e Esclusiva (ogni occorrenza di Fattura è al più una delle figlie). Le entità figlie *Pagata* e *NonPagata* sono state accorpate al genitore. A tale entità vengono aggiunti gli attributi propri delle ex entità figlie: **Scadenza** e **DataPagamento**. Si noti che non è stato aggiunto l'attributo booleano **Pagata** in quanto sarebbe ridondante, ovvero derivabile dal nuovo attributo aggiunto *DataPagamento* (Se DataPagamento ha valore **NULL** allora non è stata pagata, se ha valore diverso da **NULL** allora è stata pagata). È stata modificata anche la cardinalità della relazione Pagamento da (1,1) a (0,1).

²Modello logico proposto da E.F. Codd nel 1970, favorisce l'indipendenza dei dati e si basa sul concetto matematico di relazione. È il modello che verrà usato per la progettazione logica.

3.1.4 Formato

La quarta generalizzazione da trattare riguarda l'entità **Formato**. È una generalizzazione Totale e Esclusiva (ogni occorrenza di Fattura è al più una delle figlie). Le entità figlie *FormatoVideo* e *FormatoAudio* sono state accorpate al genitore. A tale entità vengono aggiunti gli attributi propri delle ex entità figlie. È stato aggiunto anche un attributo *TipoFormato* per determinare se esso è un formato Audio o Video.

3.2 Gestione attributi composti

Gli attributi composti, come le generalizzazioni non possono essere rappresentate direttamente nel modello relazionale.

3.2.1 Intestatario

L'attributo Intestatario è formato da Nome e Cognome ed indica i dati anagrafici segnati su una Carta di Credito. Per eliminare l'attributo composto e mantenere le informazioni sono stati creati gli attributi atomici *NomeIntestatario* e *CognomeIntestatario*.

3.3 Altre attività

Ho deciso di non effettuare alcun partizionamento/accorpamento di entità e associazioni in quanto ritengo non ci siano opportunità di ottimizzare lo schema attraverso esse.

Per quanto riguarda l'analisi delle ridondanze si guardi il paragrafo 4.9.

Nei paragrafi 4.10 e 4.11 verranno aggiornati i dizionari di entità e relationship dopo la ristrutturazione completa.

3.4 Tavola dei volumi

La tavola dei volumi è uno strumento utilizzato per specificare il numero stimato di istanze per ogni entità (E) e associazione (R) dello schema E-R. I valori utilizzati sono ovviamente approssimati ma indicativi e sono frutto di ipotesi fatte.

La tavola verrà descritta attraverso una tabella strutturata nel seguente modo:

Concetto	Tipo	Volume	Motivazione
----------	------	--------	-------------

Verranno prima descritte le entità e poi le relazioni prendendo come ordine quello relativo ai dizionari visti nei paragrafi 2.1 e 2.2. Ad ogni inizio pagina verrà ripetuta l'intestazione della tabella per migliorare la visibilità.

Concetto	Tipo	Volume	Motivazione
Film	E	1000	Ipotesi
Operatore	E	1500	Si suppone ci siano 1000 attori e 500 registi
Premio	E	40	Ipotesi
Critico	E	200	Ipotesi
Lingua	E	40	Ipotesi
Genere	E	16	Come riportato su Wikipedia esistono 16 generi cinematografici senza contare i sottogenitori
Paese	E	195	Stati riconosciuti sovrani
Utente	E	100.000	Ipotesi
Fattura	E	200.000	Si suppone una media di 2 fatture per utente
Carta	E	120.000	Si suppone una media di 1.2 carte registrate per utente
Connessione	E	5.000.000	Si suppone che un utente si colleghi in media 50 volte all'anno e che le connessioni siano azzerate al termine di ogni anno

Concetto	Tipo	Volume	Motivazione
Abbonamento	E	5	Specifiche progetto
Funzionalità	E	20	Ipotesi
Formato	E	60	Ipotesi
Codec	E	50	Ipotesi
StatoRestrittivo	E	58	Si suppone che circa il 30% degli stati abbia restrizioni
Regia	R	1200	Ipotizzo una media di 1.2 registi per film
Interpretazione	R	10.000	Ipotizzo una media di 10 attori per film
PremiazioneF	R	1000	Si tiene traccia dei premi degli ultimi 50 anni, il 50% è dedicato ai film
PremiazioneA	R	600	Il 30% dei premi degli ultimi 50 anni è dedicato agli attori
PremiazioneR	R	400	Il 20% dei premi degli ultimi 50 anni è dedicato ai registi
RecensioneC	R	4.000	Un critico recensisce in media 20 film
RecensioneU	R	100.000	Un Film viene recensito in media da 100 utenti
Tipologia	R	3000	Ipotizzo una media di 3 generi per film
Linguaggio	R	25.000	Un film ha in media 10 lingue per il doppiaggio e 15 per il sottotitolaggio
Produzione	R	1000	Cardinalità(1,1) con film

Concetto	Tipo	Volume	Motivazione
Specifiche	R	3000	Si suppone che un film abbia in media 3 formati
Aggiornamento	R	300	Si suppone che ogni formato abbia in media 5 file codec
Esclusione	R	23	Si suppone che il 40% degli stati restrittivi vietino un determinato tipo di formato
Restrizione	R	46	Si suppone che l'80% degli stati restrittivi vietino un determinato tipo di abbonamento ³
Offerta	R	30	Si suppone che un abbonamento abbia in media 6 funzionalità
Sottoscrizione	R	60.000	Si suppone che il 60% degli utenti abbia sottoscritto un abbonamento
Possesso	R	120.000	Cardinalità(1,1) con Carta
Fatturazione	R	200.000	Cardinalità(1,1) con Fattura
Pagamento	R	150.000	Si suppone che il 75% delle Fatture siano state pagate
Instaurazione	R	5.000.000	Cardinalità (1,1) con Connessione
Geolocalizzazione	R	5.000.000	Cardinalità (1,1) con Connessione
Visualizzazione	R	5.000.000	Cardinalità (1,1) con Connessione

³Si noti che gli stati restrittivi possono vietare sia formati che abbonamenti, è per questo che le percentuali sommate non fanno il 100%.

4 Operazioni sui dati

In questo capitolo verranno individuate ed analizzate le operazioni da effettuare sui dati che contribuiscono in maniera apprezzabile a determinare le prestazioni del database durante il normale carico applicativo. Per ognuna di esse, viene fornita una Descrizione, il tipo di dati di Input e Output e la Frequenza Giornaliera di esecuzione. Verrà inoltre riportata la frazione del diagramma E-R interessata e sarà individuata la porzione della tavola dei volumi coinvolta; sulla base di essa verrà stilata la tavola degli accessi.⁴

4.1 Generi di un Film

Descrizione: Questa operazione viene utilizzata da un utente quando vuole sapere i determinati generi di un film di cui conosce l'ID.

Input: ID Film.

Output: Uno o più generi.

Frequenza Stimata: 10.000 volte al giorno

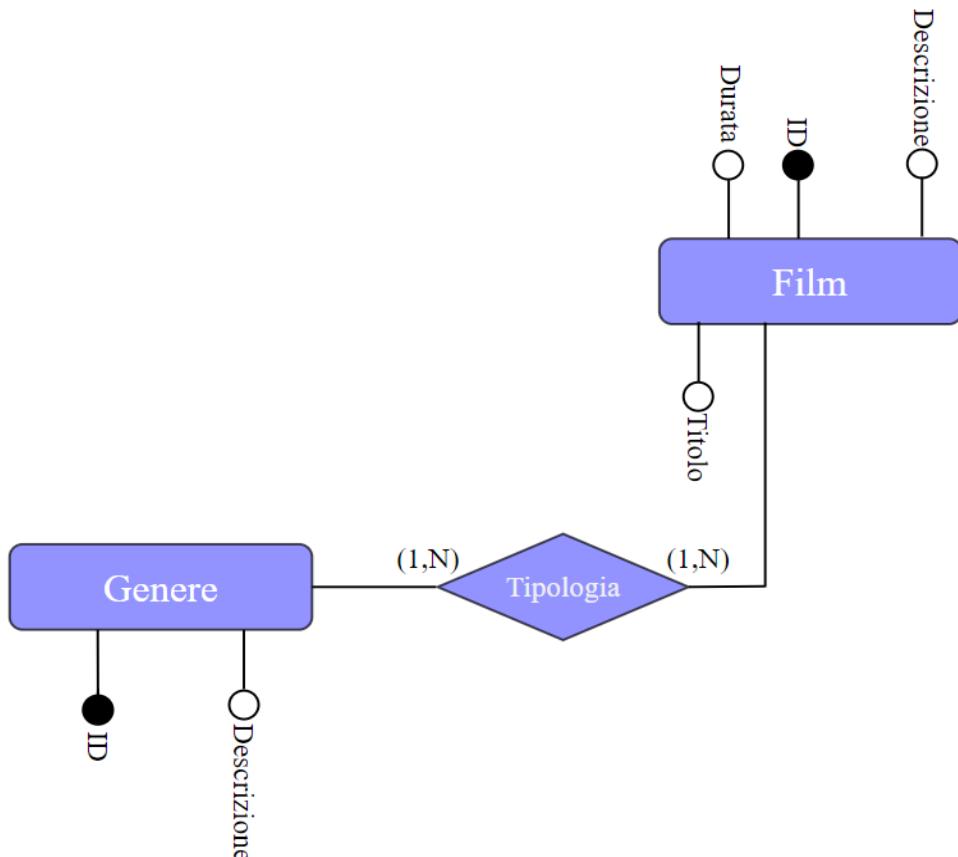


Figura 3: Porzione del diagramma E-R interessata

⁴Si tenga conto che il tipo di accesso in scrittura vale doppio rispetto all'accesso in sola lettura

4.1.1 Porzione della tavola dei volumi interessata

Concetto	Tipo	Volume
Film	E	1000
Genere	E	16
Tipologia	R	3000

4.1.2 Tavola degli Accessi

ID	Nome	TipoConcetto	TipoAccesso	Accessi	Descrizione
1	Tipologia	R	L	3000	Leggo tutta la Relazione per trovare i Generi del Film
2	Genere	E	L	3	Un Film ha in media 3 Generi e accedo all'entità per leggerne la descrizione in cui sarà scritto il nome del genere.

4.1.3 Accessi Giornalieri

$$10.000 * (3000+3) = 30.030.000$$

4.2 Nuova Connessione

Descrizione: Questa operazione permette di aggiungere una nuova connessione al Database a partire dall'indirizzo IP, l'inizio e la fine di essa, l'ID dell'Utente e il Film visualizzato.

Input: IP, InizioConnessione, FineConnessione, Hardware, ID Utente.

Output: Booleano di controllo.

Frequenza Stimata: 40.000 volte al giorno

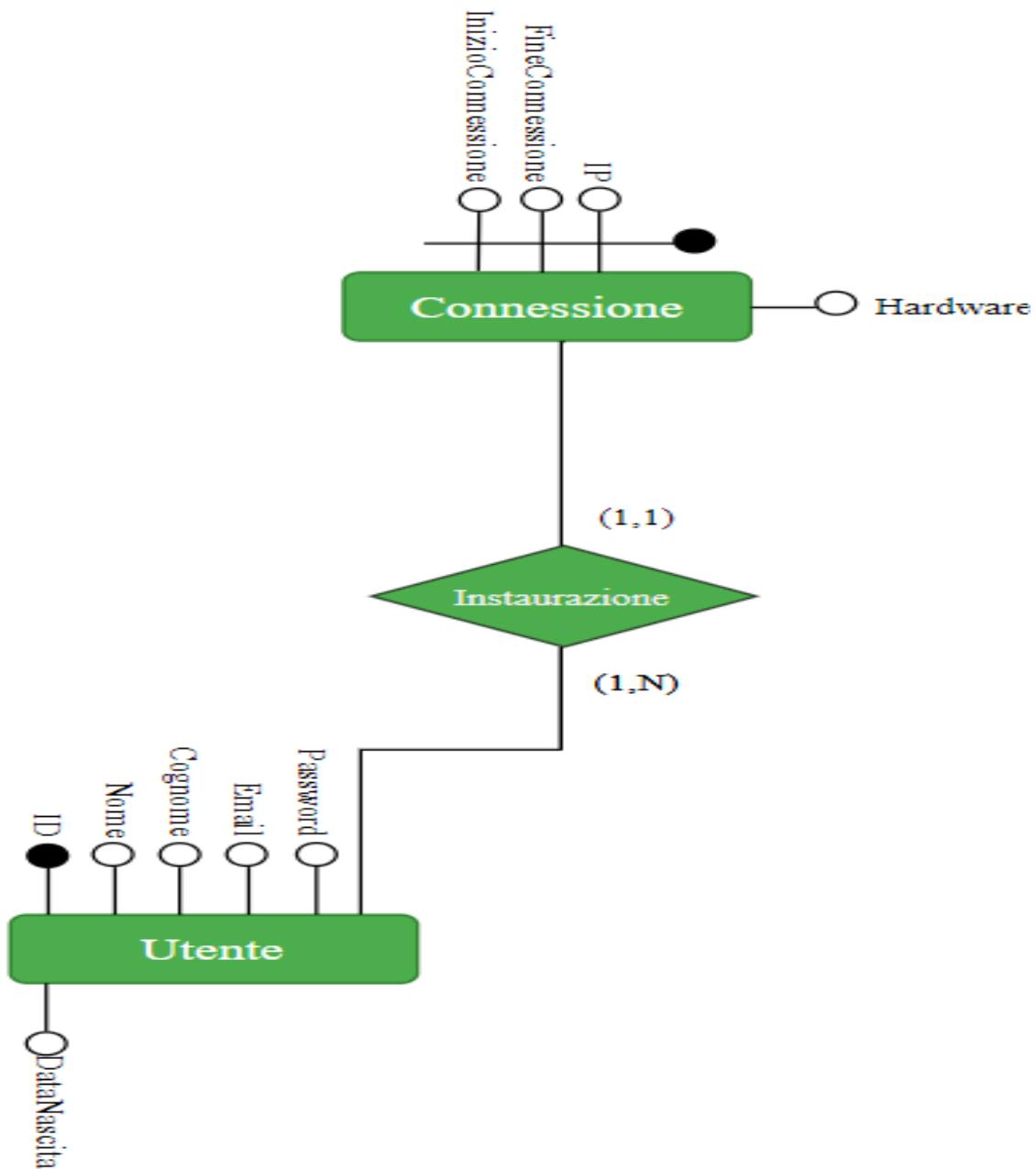


Figura 4: Porzione del diagramma E-R interessata

4.2.1 Porzione della tavola dei volumi interessata

Concetto	Tipo	Volume
Utente	E	100.000
Connessione	E	5.000.000
Instaurazione	R	5.000.000

4.2.2 Tavola degli Accessi

ID	Nome	TipoConcetto	TipoAccesso	Accessi	Descrizione
1	Connessione	E	S	1	Inserisco una nuova Connessione con i relativi attributi
2	Instaurazione	R	S	1	Associo la nuova Connessione all'ID dell'utente

4.2.3 Accessi Giornalieri

$$40.000 * (1 * 2 + 1 * 2) = 160.000$$

4.3 Inserimento Fattura

Descrizione: Tramite questa operazione si registra una nuova Fattura sul Database con la relativa Quota e Scadenza indicate dagli attributi del tipo di Abbonamento sottoscritto dall'Utente.

Input: ID Utente

Output: Booleano di controllo.

Frequenza Stimata: 100 volte al giorno

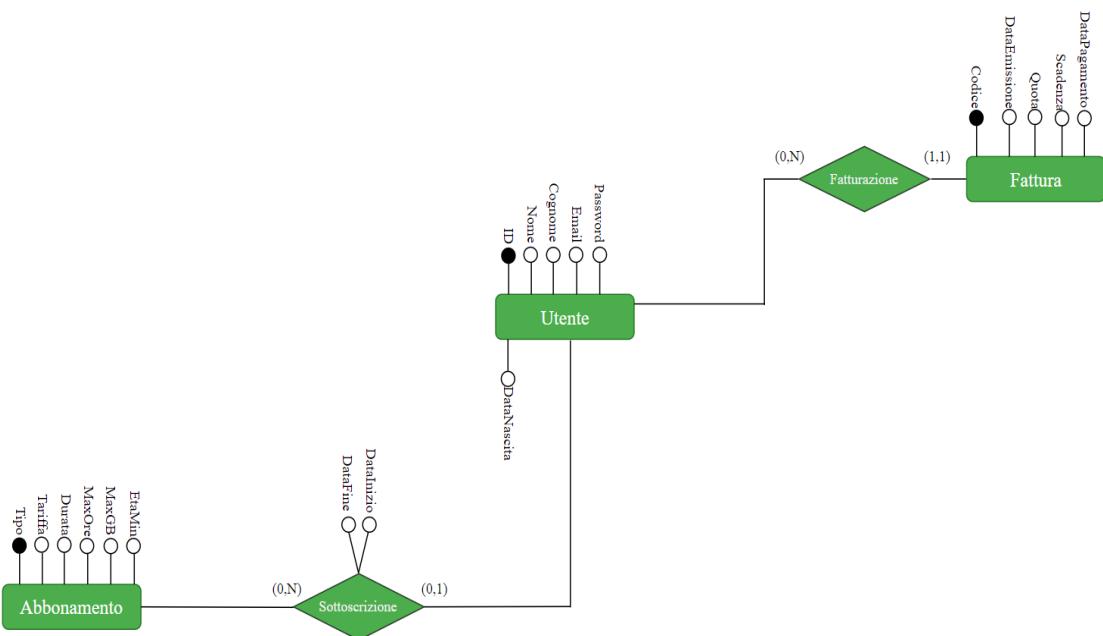


Figura 5: Porzione del diagramma E-R interessata

4.3.1 Porzione della tavola dei volumi interessata

Concetto	Tipo	Volume
Fattura	E	200.000
Utente	E	100.000
Abbonamento	E	5
Fatturazione	R	200.000
Sottoscrizione	R	60.000

4.3.2 Tavola degli Accessi

ID	Nome	TipoConcetto	TipoAccesso	Accessi	Descrizione
1	Sottoscrizione	R	L	60.000	Leggo interamente la relazione Sottoscrizione per sapere il Tipo di Abbonamento sottoscritto dall'Utente
2	Abbonamento	E	L	1	Leggo l'abbonamento per salvarmi Tariffa e Durata
3	Fatturazione	R	S	1	Associo la nuova Fattura all'Utente
4	Fattura	E	S	1	Scrivo la nuova Fattura con il relativo Codice e gli attributi.

4.3.3 Accessi Giornalieri

$$100 * (60.000 + 1 + 1 * 2 + 1 * 2) = 6.000.500$$

4.4 Lista Film disponibili per Lingua

Descrizione: Questa operazione viene utilizzata da un utente per avere una lista dei film disponibili⁵ per una lingua desiderata.

Input: Nome Lingua, Tipologia.

Output: Lista Film.

Frequenza Stimata: 200 volte al giorno

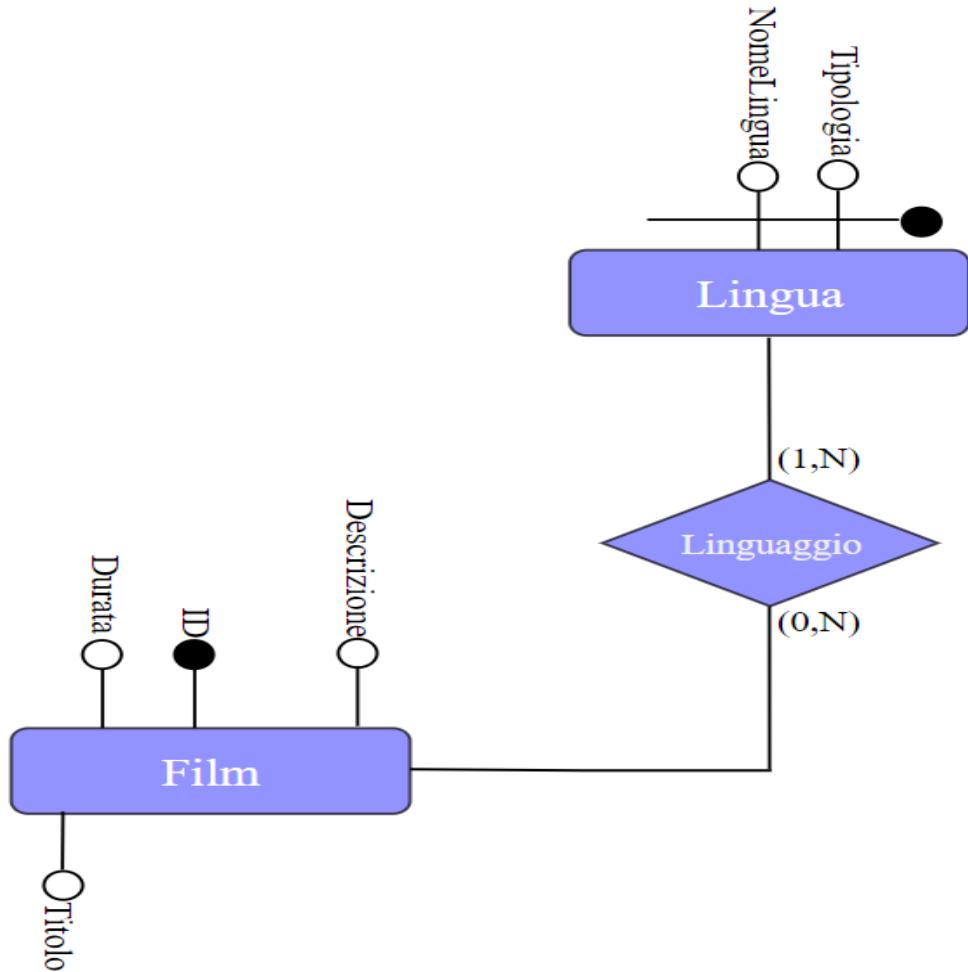


Figura 6: Porzione del diagramma E-R interessata

4.4.1 Porzione della tavola dei volumi interessata

Concetto	Tipo	Volume
Film	E	1000
Lingua	E	40
Linguaggio	R	25.000

⁵Con questa operazione non si tiene conto dell'eventuale divieto di uno o più film in base al formato e/o al tipo di abbonamento

4.4.2 Tavola degli Accessi

ID	Nome	TipoConcetto	TipoAccesso	Accessi	Descrizione
1	Linguaggio	R	L	25.000	Leggo per intero la relazione
2	Film	E	L	300	Valore ipotetico basandosi sul fatto che alcune lingue sono a disposizione in tutti i film e altre meno.

4.4.3 Accessi Giornalieri

$$200 * (25.000 + 300) = 5.060.000$$

4.5 Film con Registi Premiati

Descrizione: Questa operazione viene utilizzata da un utente per avere una lista dei film diretti da uno o più registi che hanno ricevuto almeno un premio.

Input: -

Output: ID Film.

Frequenza stimata: 500 volte al giorno

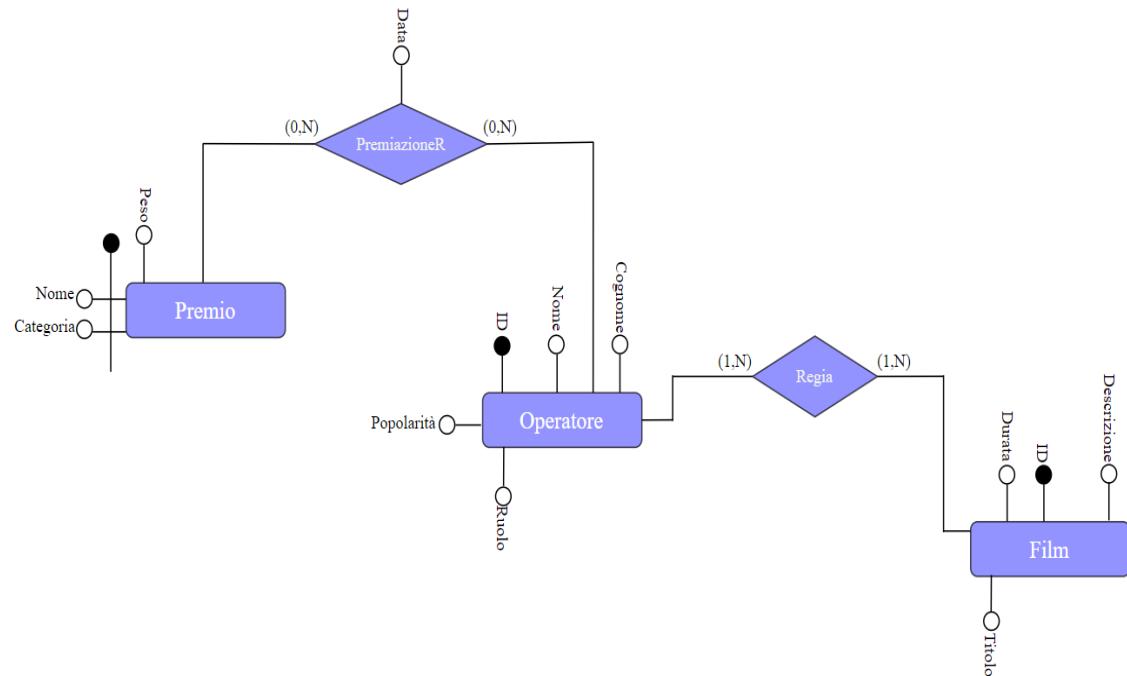


Figura 7: Porzione del diagramma E-R interessata

4.5.1 Porzione della tavola dei volumi interessata

Concetto	Tipo	Volume
Film	E	1000
Operatore	E	1500
Premio	E	40
Regia	R	1200
PremiazioneR	R	400

4.5.2 Tavola degli Accessi

ID	Nome	TipoConcetto	TipoAccesso	Accessi	Descrizione
1	Operatore	E	L	1500	Accedo all'entità per leggere il Ruolo.
2	PremiazioneR	R	L	200.000	Leggo l'intera relazione per ogni regista. 400*500
3	Regia	R	L	48.000	Leggo la relazione interamente per ogni regista premiato (40 registi, valore ipotetico). 40*1200
4	Film	E	L	80	Leggo l'entità 2 volte per ogni regista premiato in quanto un regista dirige in media 2 film. 2*40

4.5.3 Accessi Giornalieri

$$500 * (1500 + 200.000 + 48.000 + 80) = 124.790.000$$

4.6 Film vietati in base al Formato

Descrizione: Questa operazione permette ad un utente di conoscere i Film parzialmente vietati o interamente vietati perché presenti in determinati formati, vietati nel Paese scelto.

Input: Paese presente nella lista di Stati Restrittivi

Output: ID Film Parzialmente Vietati o Interamente Vietati.

Frequenza stimata: 100 volte al giorno

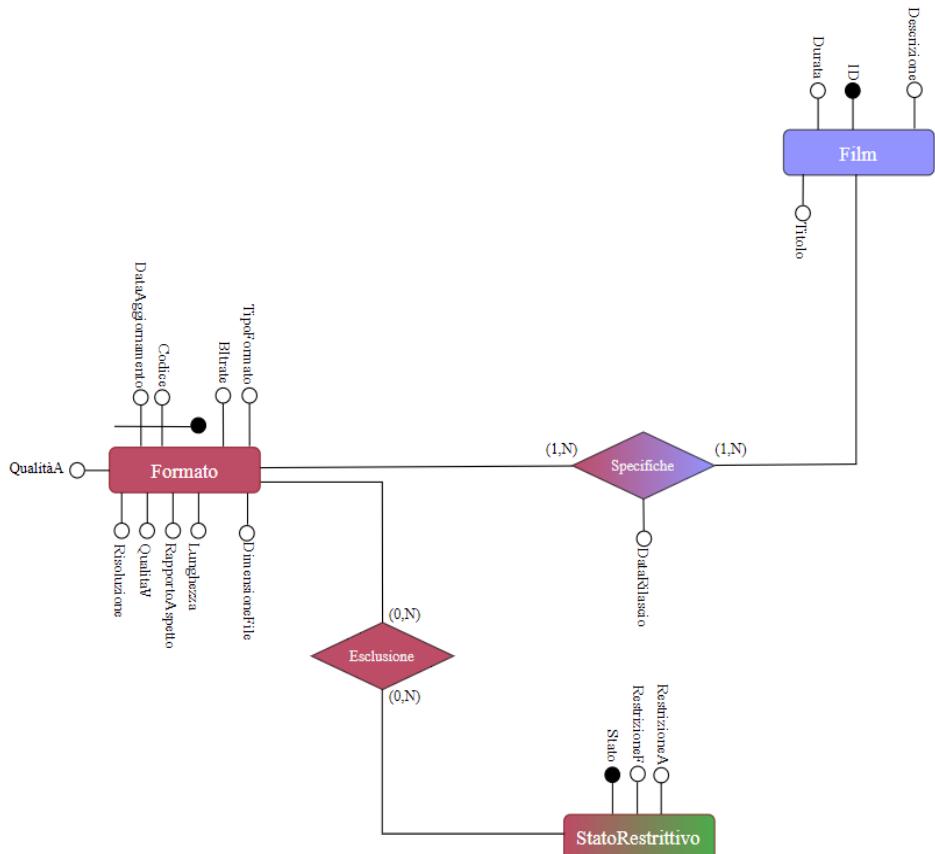


Figura 8: Porzione del diagramma E-R interessata

4.6.1 Porzione della tavola dei volumi interessata

Concetto	Tipo	Volume
Film	E	1000
Formato	E	60
StatoRestrittivo	E	58
Esclusione	R	23
Specifiche	R	3000

4.6.2 Tavola degli Accessi

ID	Nome	TipoConcetto	TipoAccesso	Accessi	Descrizione
1	StatoRestrittivo	E	L	1	Leggo per scoprire se il Paese dato in Input è restrittivo su formato meno.
2	Esclusione	R	L	23	Leggo interamente la Relazione per sapere il codice del Formato vietato.
3	Specifiche	R	L	3000	Leggo la relazione interamente per ogni formato vietato (Si stima che uno Stato vieti al massimo 1 formato)

4.6.3 Accessi Giornalieri

$$100 * (1 + 23 + 3000) = 302.400$$

4.7 Valutazione Film

Descrizione: Questa operazione permette di conoscere la valutazione⁶ di un determinato Film.

Input: ID Film

Output: Un numero reale da 0 a 10.

Frequenza stimata: 5.000 volte al giorno

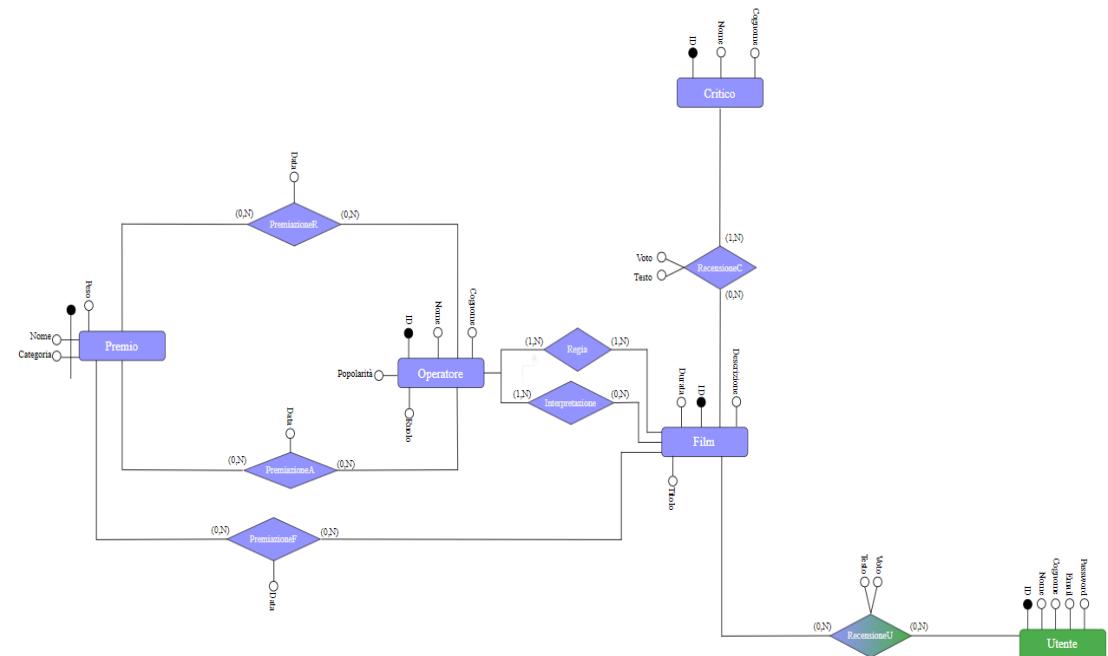


Figura 9: Porzione del diagramma E-R interessata

La Valutazione viene calcolata tramite i seguenti parametri:

Recensione Utente: Ogni utente può recensire un Film con un Voto da 1 a 10 e ogni recensione andrà a comporre la media ValutazioneUtenti che avrà un peso pari a $\frac{1}{10}$ sulla valutazione finale del film.

Recensione Critico: Ogni recensione da parte di un critico comporrà la media ValutazioneCritici che avrà un peso pari a $\frac{5}{10}$ sulla valutazione finale del film.

Premiazioni: Per quanto riguarda le premiazioni degli artisti avranno entrambi un peso pari a $\frac{1}{10}$ sulla valutazione finale, mentre le premiazioni dei film avranno un peso pari a $\frac{2}{10}$ e verranno usati i seguenti criteri:

⁶Un numero reale da 0 a 10

Regista	
Peso Totale ⁷	ValutazioneRegisti
0	4
Da 1 a 15	7
Più di 15	10
Attore	
Peso Totale	ValutazioneAttori
0	4
Da 1 a 8	6
Da 9 a 15	8
Più di 15	10
Film	
Peso Totale	ValutazionePropria
0	4
Da 1 a 10	7
Più di 10	10

Qui di seguito è fornita la formula per calcolare la Valutazione.

$$\frac{VU + 5 * VC + VR + VA + 2 * VP}{10}$$

Dove n è il numero di recensioni degli utenti, m è il numero di recensioni dei critici e per tutte le Valutazioni è stata usata una sigla intuitiva (VU ValutazioneUtenti, VC ValutazioneCritici, etc).

4.7.1 Porzione della tavola dei volumi interessata

Concetto	Tipo	Volume
Film	E	1000
Utente	E	100.000
Critico	E	200
Operatore	E	1500
Premio	E	40
RecensioneU	R	100.000
RecensioneC	R	4.000
Regia	R	1.200
Interpretazione	R	10.000
PremiazioneR	R	400
PremiazioneA	R	600
PremiazioneF	R	1000

⁷Ogni premio ha un Peso in base all'importanza, ed è un numero intero compreso tra 1 e 5. Il Peso Totale si calcola sommando il peso di tutti i Premi ottenuti dai Registi del Film.

4.7.2 Tavola degli Accessi

ID	Nome	TipoConcetto	TipoAccesso	Accessi	Descrizione
1	RecensioneU	R	L	100.000	Leggo per intero la Relazione per tenere traccia del Voto di ogni Recensione.
2	RecensioneC	R	L	4.000	Leggo per intero la Relazione per tenere traccia del Voto di ogni Recensione.
3	Regia	R	L	1.200	Leggo per intero la Relazione per trovare i Registi del Film
4	Interpretazione	R	L	10.000	Leggo per intero la Relazione per trovare gli Attori del Film
5	PremiazioneR	R	L	480	Leggo per intero la Relazione per ogni Regista del Film (1.2 registi in media per Film). $400 * 1.2$

ID	Nome	TipoConcetto	TipoAccesso	Accessi	Descrizione
6	PremiazioneA	R	L	6.000	Leggo per intero la Relazione per ogni Attore del Film (10 attori in media per Film). 600*10
7	PremiazioneF	R	L	1000	Leggo per intero la Relazione
8	Premio	E	L	12	Leggo l'attributo Peso, per ogni Premio del Film o Artista. Valore Stimato.

4.7.3 Accessi Giornalieri

$$5.000 * (100.000 + 4.000 + 1.200 + 10.000 + 480 + 6.000 + 1.000 + 12) = \\ 613.460.000$$

4.8 Raccomandazione Contenuti

Descrizione: L'operazione in questione consente di conoscere i 5 film con "compatibilità" più alta per ogni Utente. La compatibilità è calcolata tenendo conto degli ultimi 5 film guardati dall'utente; di essi si controlla il Genere e il Regista, questi due elementi sono a mio parere gli aspetti più caratteristici per un film.

Il Genere contribuisce al 70% sulla compatibilità, mentre il Regista il 30%; queste due percentuali andranno poi sommate e la percentuale risultante verrà utilizzata per la raccomandazione.

ESEMPIO: Gli ultimi 5 film visti da un Utente X sono stati 5 film di fantascienza, tutti e 5 diretti da 'John Carpenter'. Il 100% degli ultimi 5 film hanno come genere Fantascienza e Regista Carpenter, quindi per un nuovo Film mai visualizzato che ha come genere Fantascienza e come regista Carpenter avremo:

$$\text{COMPATIBILITÀ} = 70\% + 30\% = 100\%$$

Input: ID Utente

Output: I 5 film mai visti dall'Utente con la "compatibilità" più alta.

Frequenza stimata: 20.000 volte al giorno.

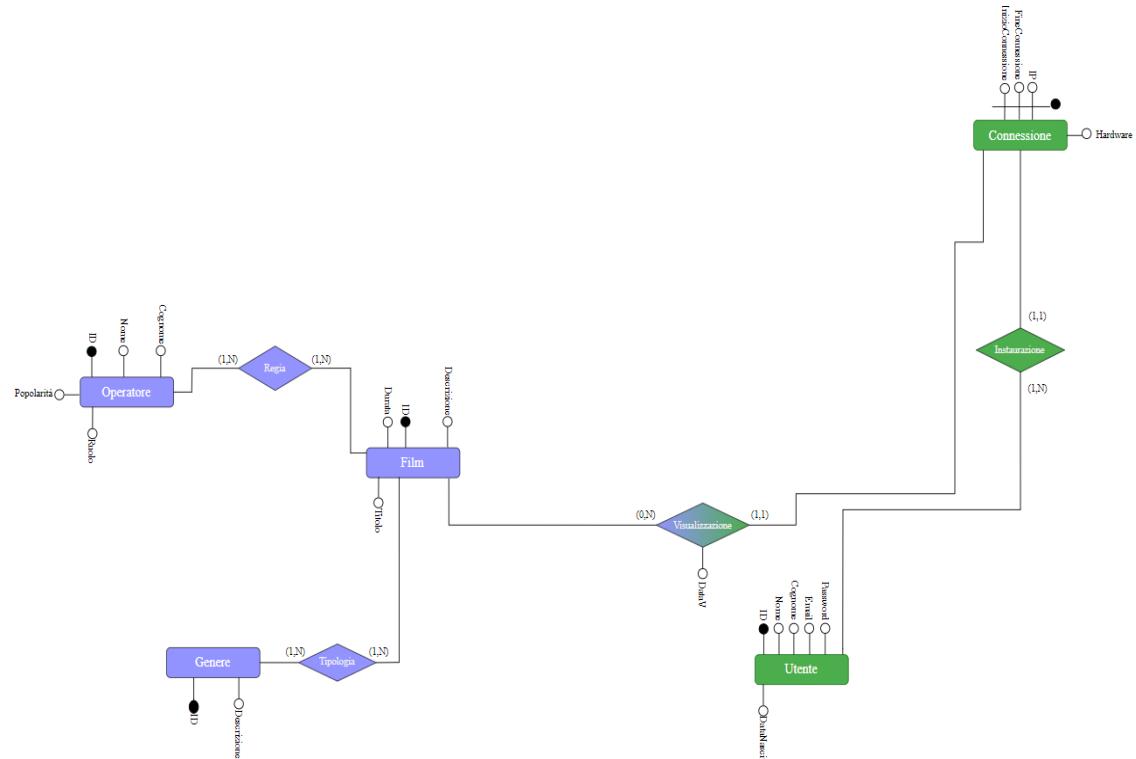


Figura 10: Porzione del diagramma E-R interessata

4.8.1 Porzione della tavola dei volumi interessata

Concetto	Tipo	Volume
Utente	E	100.000
Connessione	E	5.000.000
Film	E	1000
Operatore	E	1500
Genere	E	16
Instaurazione	R	5.000.000
Visualizzazione	R	5.000.000
Regia	R	1.200
Tipologia	R	3.000

4.8.2 Tavola degli Accessi

ID	Nome	TipoConcetto	TipoAccesso	Accessi	Descrizione
1	Instaurazione	R	L	5 milioni	Leggo per intero la Relazione per trovare le Connessioni dell'Utente.
2	Visualizzazione ⁸	R	L	250.000	Leggo la Relazione circa 5.000 volte (ogni Film è visualizzato in media 5.000 volte), per ogni Connessione trovata (in media 50) per vedere i Film visti dall'Utente, salvo tutte le visualizzazioni, in particolare le ultime 5. $5.000 * 50$

⁸Alla fine di quest'operazione ho una lista dei Film visti dall'utente e grazie ad essa posso trovare i Film mai visti dall'utente e gli ultimi 5 Film avendo a disposizione la data della visualizzazione.

ID	Nome	TipoConcetto	TipoAccesso	Accessi	Descrizione
3	Regia	R	L	6.000	Leggo per intero la Relazione per trovare i Registi degli ultimi 5 Film. 1.200 * 5
4	Tipologia	R	L	15.000	Leggo per intero la Relazione per trovare i generi degli ultimi 5 Film. 3.000 * 5

4.8.3 Accessi Giornalieri

$$20.000 * (5.000.000 + 250.000 + 6.000 + 15.000) = 105.420.000.000$$

4.9 Analisi ridondanze

Per valutare l'eventuale convenienza di una ridondanza analizziamo gli accessi di una o più operazioni in presenza e in assenza di essa.

4.9.1 PremiatoP

L'attributo aggiunto è un booleano che vale **TRUE** se il Film in questione è stato diretto da un Regista premiato, **FALSE** altrimenti.

L'aggiornamento della ridondanza è ottenuto tramite *immediate refresh*, effettuato ogniqualvolta viene assegnato un premio a un regista. Tenendo conto del fatto che vengono assegnati 8 premi all'anno ai registi (secondo la Tavola dei Volumi) il refresh verrà effettuato 8 volte all'anno e quindi circa 0,02 volte al Giorno.

Analizziamo l'operazione 4.5.

Assenza di ridondanza

ID	Nome	TipoConcetto	TipoAccesso	Accessi	Descrizione
1	Operatore	E	L	1500	Accedo all'entità per leggere il Ruolo.
2	PremiazioneR	R	L	200.000	Leggo l'intera relazione per ogni regista. 400*500
3	Regia	R	L	48.000	Leggo la relazione interamente per ogni regista premiato (40 registi, valore ipotetico). 40*1200

ID	Nome	TipoConcetto	TipoAccesso	Accessi	Descrizione
4	Film	E	L	80	Leggo l'entità 2 volte per ogni regista premiato in quanto un regista dirige in media 2 film. $2*40$

Accessi Giornalieri: $500 * (1500 + 200.000 + 48.000 + 80) = 124.790.000$

Presenza di ridondanza

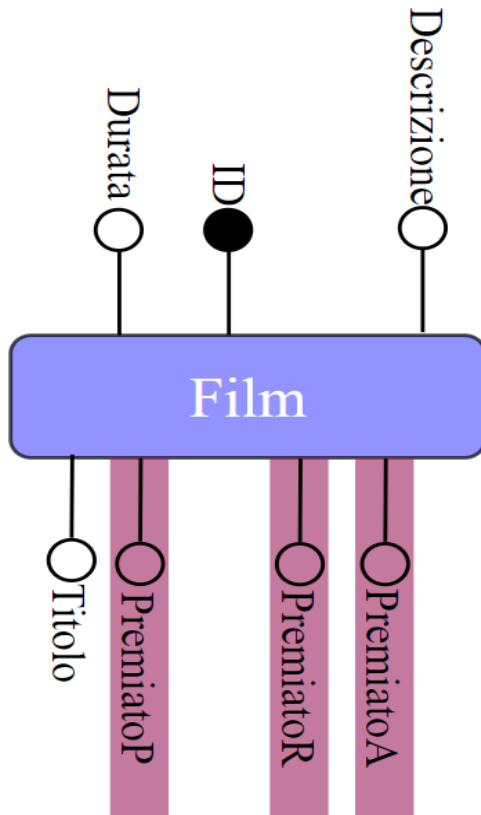


Figura 11: Porzione del diagramma E-R interessata

ID	Nome	TipoConcetto	TipoAccesso	Accessi	Descrizione
1	Film	E	L	1000	Leggo l'intera entità per vedere i Film che hanno l'attributo booleano <i>PremiatoR</i> = TRUE

Accessi Giornalieri: $500 * (1000) = 500.000$

Aggiornamento: Per quanto riguarda l'aggiornamento, effettuato circa 0,02 volte al giorno, la Tavola degli Accessi è la seguente⁹:

⁹Si tenga conto che quando viene assegnato un premio noi conosciamo l'ID del Premiato

ID	Nome	TipoConcetto	TipoAccesso	Accessi	Descrizione
1	Regia	R	L	1200	Leggo l'intera relazione per leggere l'ID dei film che il Regista ha diretto
2	Film	E	L	2	Leggo l'entità due volte (un regista dirige in media due film) per controllare se l'attributo <i>PremiatoR</i> = FALSE
3	Film	E	S	2	Se l'attributo <i>PremiatoR</i> = FALSE lo cambio in TRUE

Accessi Giornalieri Aggiornamento ridondanza: $0,02 * (1200 + 2 + 2 * 2) = 24,12$, arrotondiamo a 25.

Accessi Giornalieri senza ridondanza: 124.790.000

Accessi Giornalieri con ridondanza + Aggiornamento ridondanza: $500.000 + 25 = 500.025$.

La ridondanza è quindi molto conveniente. Si noti che sono stati inseriti anche gli attributi PremiatoP e PremiatoA che potrebbero servire per alleggerire il peso in operazioni simili a questa ma riguardanti gli Attori o il Film in sè e per sè.

4.9.2 Valutazione

L'attributo aggiunto è un numero reale compreso tra 0 e 10 calcolata come visto nell'operazione 4.7. L'aggiornamento della ridondanza è ottenuto tramite *deferred refresh* per poter effettuare l'aggiornamento di notevoli dimensioni in ore nelle quali il traffico di dati è meno intenso. Analizziamo l'operazione 4.7.

Assenza di ridondanza

ID	Nome	TipoConcetto	TipoAccesso	Accessi	Descrizione
1	RecensioneU	R	L	100.000	Leggo per intero la Relazione per tenere traccia del Voto di ogni Recensione.
2	RecensioneC	R	L	4.000	Leggo per intero la Relazione per tenere traccia del Voto di ogni Recensione.
3	Regia	R	L	1.200	Leggo per intero la Relazione per trovare i Registi del Film
4	Interpretazione	R	L	10.000	Leggo per intero la Relazione per trovare gli Attori del Film
5	PremiazioneR	R	L	480	Leggo per intero la Relazione per ogni Regista del Film (1.2 registi in media per Film). $400 * 1.2$

ID	Nome	TipoConcetto	TipoAccesso	Accessi	Descrizione
6	PremiazioneA	R	L	6.000	Leggo per intero la Relazione per ogni Attore del Film (10 attori in media per Film). 600*10
7	PremiazioneF	R	L	1000	Leggo per intero la Relazione
8	Premio	E	L	12	Leggo l'attributo Peso, per ogni Premio del Film o Artista. Valore Stimato.

Accessi Giornalieri: $5.000 * (100.000 + 4.000 + 1.200 + 10.000 + 480 + 6.000 + 1.000 + 12) = 613.460.000$

Presenza di ridondanza

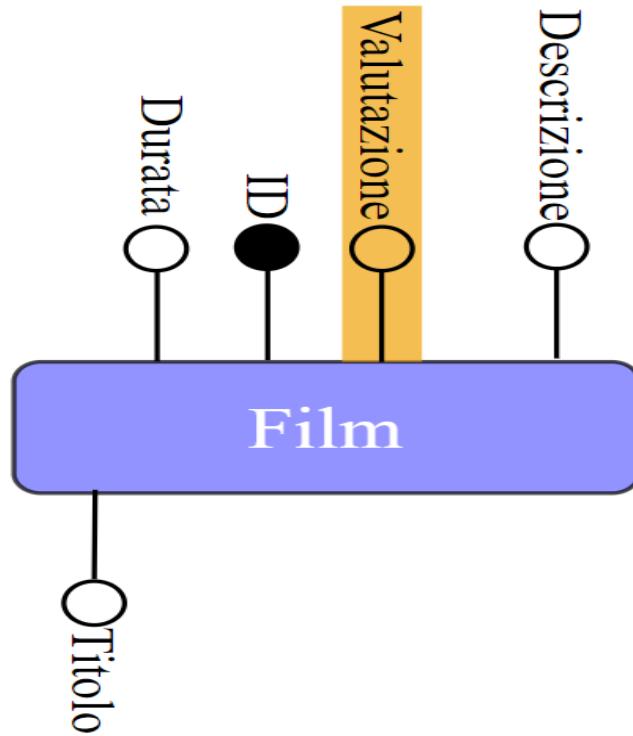


Figura 12: Porzione del diagramma E-R interessata

ID	Nome	TipoConcetto	TipoAccesso	Accessi	Descrizione
1	Film	E	L	1	Dato l'ID del Film leggo la Valutazione

Accessi Giornalieri: $5.000 * 1 = 5.000$

Aggiornamento: Per quanto riguarda l'aggiornamento, esso viene effettuato durante le ore notturne ogni giorno, per ogni Film presente sul sito. La Tavola degli Accessi è identica a quella dell'operazione 4.7 a meno di una scrittura sull'attributo Valutazione.

Accessi Giornalieri Aggiornamento ridondanza: $1.000 * (100.000 + 4.000 + 1.200 + 10.000 + 480 + 6.000 + 1.000 + 12 + 2) = 122.694.000$

Accessi Giornalieri senza ridondanza: 613.460.000

Accessi Giornalieri con ridondanza + Aggiornamento ridondanza: $5.000 + 122.694.000 = 122.699.000$

La ridondanza è quindi conveniente sia a livelli di accessi, sia perché l'aggiornamento, ovvero l'operazione con peso più rilevante, è effettuata in orari notturni.

4.10 Aggiornamento dizionario Entità

Nel dizionario riportato qui sotto verranno inserite solo le entità che sono state modificate durante la ristrutturazione dello schema E-R e la traduzione verso lo schema logico.

Nome Entità	Attributi	Sinonimi	Descrizione
Film	ID , Durata, Descrizione, Titolo, Valutazione, PremiatoP, PremiatoR, PremiatoA	Pellicola	Contenuto disponibile nella piattaforma Streaming
Operatore	ID , Nome, Cognome, Popolarità, Ruolo	Organico	Addetti ai lavori facenti parte della troupe di un Film, comprendente due figure artistiche: Attore e Regista
Lingua	NomeLingua , Tipologia	Idioma	Lingue disponibili per l'audio e per i sottotitoli di un film
Fattura	Codice , DataEmissione, Quota, Scadenza, DataPagamento		Documento che il venditore trasmette al compratore nel momento in cui avviene l'acquisto dell'abbonamento, si tiene traccia delle fatture già pagate e quelle non ancora pagate
Formato	Codice , DataAggiornamento , Bitrate, DimensioneFile, TipoFormato, QualitàA, QualitàV, Lunghezza, RapportoAspetto, Risoluzione		Sistema di codifica dei dati di un Film riguardante formato audio e formato video

4.11 Aggiornamento dizionario Relazioni

Relationship	Cardinalità	Attributi	Descrizione
Regia	Film (1,N) Operatore(1,N)		Direzione della realizzazione di un film da parte di un Regista
Interpretazione	Film (0,N) Operatore (1,N)		Presenza di un Attore in un Film
PremiazioneR	Operatore (0,N) Premio (0,N)	Data	Vincita di un premio da parte di un Regista
PremiazioneA	Operatore (0,N) Premio (0,N)	Data	Vincita di un premio da parte di un Attore
Pagamento	Carta (0,N) Fattura(0,1)		Attestazione di un Pagamento di una Fattura tramite Carta di Credito

5 Area Analytics

5.1 Classifiche

Descrizione: Tra i piani su cui competono i servizi di streaming spicca la scelta dei contenuti, è quindi utile per un'azienda sapere su cosa investire. Questa funzionalità serve a capire su quali Film e su quali Formati investire. I gusti e le preferenze sono disposte in modo eterogeneo in tutti i Paesi del mondo, è quindi utile, se non necessario, differenziare le classifiche per ogni Stato. La funzionalità si basa sul numero delle visualizzazioni degli Utenti, differenziati in termini di Nazionalità e piano di Abbonamento, per distinguere utenti "più interessati" e quindi disposti a spendere più soldi per avere un servizio più completo, e utenti che si "accontentano" di servizi base; i primi sono sicuramente quelli di cui le aziende terranno più conto. La procedura restituisce la classifica dei film più visualizzati dato il Paese e il piano di Abbonamento con associato il Formato Audio e Video più frutto. È importante specificare che questa classifica riguarda solo i film presenti su FilmSphere, non può quindi aiutare l'azienda a sapere su quali nuovi Film investire. L'azienda può utilizzare questa funzionalità per capire quali Film e Formati presenti sulla piattaforma rimuovere e/o su quali Film e Formati investire meno denaro.

Input: Stato, piano di Abbonamento.

Output: La classifica dei Film più visualizzati dagli Utenti con associato il Formato Audio e Video¹⁰.

Frequenza stimata: 2 volte all'anno per ogni combinazione Paese-Abbonamento.
 $2 * (5 * 195) = 1950$ all'anno. Quindi circa 6 volte al giorno.

¹⁰Questa associazione non dà in realtà una vera e propria classifica dei formati ma restituisce i Formati associati per l'appunto ai Film più visualizzati, questo risultato può far ragionare l'azienda, per esempio se nei primi 3 film è presente lo stesso tipo di formato video, allora forse è il caso di puntare su di esso.

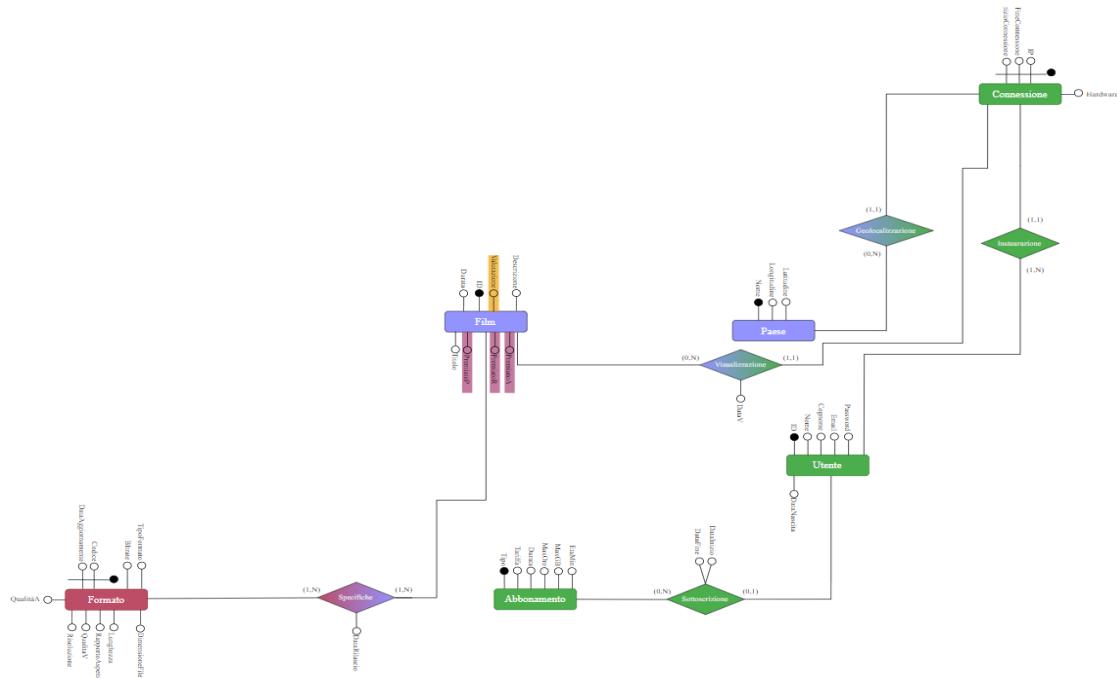


Figura 13: Porzione del diagramma E-R interessata

5.1.1 Porzione della tavola dei volumi interessata

Concetto	Tipo	Volume
Abbonamento	E	5
Utente	E	100.000
Connessione	E	5.000.000
Paese	E	195
Film	E	1000
Formato	E	60
Sottoscrizione	R	60.000
Instaurazione	R	5.000.000
Geolocalizzazione	R	5.000.000
Visualizzazione	R	5.000.000
Specifiche	R	3.000

5.1.2 Tavola degli Accessi

ID	Nome	TipoConcetto	TipoAccesso	Accessi	Descrizione
1	Geolocalizzazione	R	L	5 milioni	Leggo per intero la Relazione per trovare il numero di Connessioni associate al Paese dato.
2	Sottoscrizione	R	L	60.000	Leggo per intero la Relazione per risalire agli ID di tutti gli Utenti con il tipo di Abbonamento dato.

ID	Nome	TipoConcetto	TipoAccesso	Accessi	Descrizione
3	Instaurazione	R	L	3.100	Leggo la relazione tot volte quante sono le connessioni instaurate dagli utenti di uno stato abbonati ad un determinato tipo di abbonamento. ¹¹
4	Visualizzazione	R	L	15 milioni e 500 mila	Leggo la relazione 5.000 volte (Numero di visualizzazioni per film in media) per ogni connessione trovata (circa 3.100). 5.000*3.100

In questo momento io ho la classifica dei Film più visualizzati, mi resta solo da determinarne il Formato Audio e Video più frutto.

¹¹Sono state svolte delle stime basate su proporzioni: $100.000 \text{ utenti} / 195 = 513 \text{ utenti per stato}$, di cui il 20% è abbonato ad un determinato piano di abbonamento = 62, ogni utente ha in media 50 connessioni = 3.100

ID	Nome	TipoConcetto	TipoAccesso	Accessi	Descrizione
5	Specifiche	R	L	3.000.000	Leggo interamente la relazione per ogni Film disponibile su FilmSphere. 1.000 * 3.000.
6	Formato	E	L	3.000	Avendo la chiave di Formato, effettuo un singolo accesso per ogni Formato per leggere l'attributo TipoFormato

5.1.3 Accessi Giornalieri

$$6 * (5.000.000 + 60.000 + 3.100 + 15.500.000 + 3.000.000 + 3.000) = 141.396.600$$

6 Progettazione Logica

6.1 Traduzione in Schema Logico

Viene ora presentato il Database in Modello Logico Relazionale, ottenuto tramite l'algoritmo di traduzione da schema concettuale. I nomi delle Relazioni sono indicati in **Grassetto**, mentre gli identificatori tramite sottolineatura.

Film(ID, Descrizione, Durata, Titolo, Valutazione, PremiatoP, PremiatoR, PremiatoA, PaeseProduzione, AnnoProduzione)

Operatore(ID, Nome, Cognome, Popolarità, Ruolo)

Regia(Film, Regista)

Interpretazione(Film, Attore)

Premio(Nome, Categoria, Peso)

PremiazioneR(NomePremio, CategoriaPremio, Regista, Data)

PremiazioneA(NomePremio, CategoriaPremio, Attore, Data)

PremiazioneF(NomePremio, CategoriaPremio, Film, Data)

Critico(ID, Nome, Cognome)

RecensioneC(Film, Critico, Voto, Testo)

Lingua(NomeLingua, Tipologia)

Linguaggio(Film, NomeLingua, TipologiaLingua)

Genere(ID, Descrizione)

Tipologia(Film, Genere)

Paese(Nome, Longitudine, Latitudine)

Connessione(IP, InizioConnessione, FineConnessione, Hardware, Paese, Utente, Film)

Utente(ID, Nome, Cognome, Email, Password, DataNascita, SottoscrizioneAbbonamento, DataInizioAbb, DataFineAbb)

RecensioneU(Film, Utente, Voto, Testo)

Fattura(Codice, DataEmissione, Quota, Scadenza, DataPagamento, FatturazioneUtente, CartaPagamento)

Carta(PAN, CVV, DataScadenza, NomeIntestatario, CognomeIntestatario, UtenteProprietario)

Abbonamento(Tipo, Tariffa, Durata, MaxOre, MaxGB, EtaMin)

Offerta(Abbonamento, Funzionalità)

Funzionalità(Nome, Descrizione)

Restrizione(Abbonamento, StatoRestrittivo)

StatoRestrittivo(Stato, RestrizioneF, RestrizioneA)

Esclusione(StatoRestrittivo, CodiceFormato, AggiornamentoFormato)

Formato(Codice, DataAggiornamento, Bitrate, TipoFormato, QualitàA, QualitàV, DimensioneFile, Lunghezza, RapportoAspetto, Risoluzione)

Aggiornamento(CodiceFormato, Codec, Data)

Codec(FileCodec, Specifiche)

Specifiche(Film, CodiceFormato, DataAggiornamento, DataRilascio)

6.2 Vincoli di Integrità

6.2.1 Vincoli di Integrità Referenziale

I vincoli di integrità referenziale impongono ai valori di una relazione R_1 di comparire come valori della chiave primaria della relazione R_2 . Vediamo qui sotto i vincoli del Database.

1. Film(PaeseProduzione) in Paese(Nome).
2. Regia(Film) in Film(ID).
3. Regia(Regista) in Operatore(ID) ove l'attributo Ruolo associato all'Operatore vale 'R'.
4. Interpretazione(Film) in Film(ID).
5. Interpretazione(Attore) in Operatore(ID) ove l'attributo Ruolo associato all'Operatore vale 'A'.
6. PremiazioneR(NomePremio) in Premio (Nome), ove l'attributo Categoria associato al Premio vale 'R'.
7. PremiazioneR(CategoriaPremio) in Premio (Categoria)
8. PremiazioneR(Regista) in Operatore (ID) ove l'attributo Ruolo associato all'operatore vale 'R'.
9. PremiazioneA(NomePremio) in Premio (Nome), ove l'attributo Categoria associato al Premio vale 'A'.
10. PremiazioneA(CategoriaPremio) in Premio (Categoria)
11. PremiazioneA(Attore) in Operatore (ID), ove l'attributo Ruolo associato all'operatore vale 'A'.
12. PremiazioneF(NomePremio) in Premio (Nome), ove l'attributo Categoria associato al Premio vale 'F'.

13. PremiazioneF(CategoriaPremio) in Premio (Categoria)
14. PremiazioneF(Film) in Film (ID).
15. RecensioneC(Film) in Film(ID).
16. RecensioneC(Critico) in Critico(ID).
17. Linguaggio(Film) in Film(ID).
18. Linguaggio(NomeLingua) in Lingua(NomeLingua).
19. Linguaggio(TipologiaLingua) in Lingua(Tipologia).
20. Tipologia(Film) in Film(ID).
21. Tipologia(Genero) in Genero(ID).
22. Connessione(Paese) in Paese(Nome).
23. Connessione(Utente) in Utente(ID).
24. Connessione(Film) in Film(ID).
25. RecensioneU(Film) in Film(ID).
26. RecensioneU(Utente) in Utente(ID).
27. Utente(SottoscrizioneAbbonamento) in Abbonamento(Tipo).
28. Fattura(FatturazioneUtente) in Utente(ID).
29. Fattura(CartaPagamento) in Carta(PAN).

30. Carta(UtenteProprietario) in Utente(ID).
31. Offerta(Abonnamento) in Abbonamento(Tipo).
32. Offerta(Funzionalità) in Funzionalità(Nome).
33. Restrizione(Abonnamento) in Abbonamento(Tipo).
34. Restrizione(StatoRestrittivo) in StatoRestrittivo(Stato), ove l'attributo RestrizioneA associato allo StatoRestrittivo vale 'TRUE'.
35. Esclusione(StatoRestrittivo) in StatoRestrittivo(Stato), ove l'attributo RestrizioneF associato allo StatoRestrittivo vale 'TRUE'.
36. Esclusione(CodiceFormato) in Formato(Codice).
37. Esclusione(AggIORNAMENTOFormato) in Formato(DataAggiornamento).
38. Aggiornamento(CodiceFormato) in Formato(Codice).
39. Aggiornamento(Codec) in Codec(FileCodec).
40. Specifiche(Film) in Film(ID).
41. Specifiche(CodiceFormato) in Formato(Codice).
42. Specifiche(AggIORNAMENTOFormato) in Formato(DataAggiornamento).

6.2.2 Vincoli Intra-Relazionali

Per i vincoli Intra-Relazionali il soddisfacimento è definito rispetto a una singola relazione della base di dati. Essi sono stati gestiti in SQL tramite CHECK.

Vincoli di Dominio

Nell'elenco non sono stati aggiunti i vincoli di data $<= \text{CURRENT DATE}$ in quanto li ho reputati banali.

1. Film(Valutazione) ≥ 0.0 AND Film(Valutazione) ≤ 10.0
2. Operatore(Popolarità) ≥ 0 AND Operatore(Popolarità) ≤ 5
3. Operatore(Ruolo) = 'R' OR Operatore(Ruolo) = 'A'
4. Premio(Categoria) = 'R' OR Premio(Categoria) = 'A' OR Premio(Categoria) = 'F'
5. Premio(Peso) ≥ 1 AND Premio(Peso) ≤ 5
6. RecensioneC(Voto) ≥ 0 AND RecensioneC(Voto) ≤ 10
7. Lingua(Tipologia) = 'S' OR Lingua(Tipologia) = 'D'
8. Paese(Longitudine) ≥ -180.0 AND Paese(Longitudine) ≤ 180.0
9. Paese(Latitudine) ≥ -90.0 AND Paese(Latitudine) ≤ 90.0
10. Connessione(IP) $\geq 000.000.000.001$ AND Connessione(IP) $\leq 255.255.255.255$
11. RecensioneU(Voto) ≥ 0 AND RecensioneU(Voto) ≤ 10
12. Carta(PAN) $\geq 0000\ 0000\ 0000\ 0000$ AND Carta(PAN) $\leq 9999\ 9999\ 9999\ 9999$
13. Carta(CVV) ≥ 000 AND Carta(CVV) ≤ 999 OR Carta(CVV) ≥ 0000 AND Carta(CVV) ≤ 9999
14. Formato(Bitrate) ≥ 400 AND Formato(Bitrate) ≤ 14000
15. Formato(Risoluzione) ≥ 480 AND Formato(Risoluzione) ≤ 9999

Vincoli di Tupla

1. Connessione(InizioConnessione) < Connessione(FineConnessione)
2. Utente(DataInizioAbb) < Utente(DataFineAbb)
3. Fattura(DataEmissione) < Fattura(Scadenza)
4. Fattura(DataEmissione) < Fattura(DataPagamento)

6.2.3 Vincoli Inter-Relazionali

Il soddisfacimento è definito rispetto a più relazioni della base di dati.

1. Film(AnnoProduzione) < PremiazioneF(YEAR(Data))
2. Film(AnnoProduzione) < Specifiche(YEAR(DataRilascio))
3. Film(AnnoProduzione) < Connessione(YEAR(DataVisualizzazioneFilm))

6.3 Dipendenze Funzionali

Nelle specifiche del progetto è richiesto che il database progettato sia in forma normale di Boyce-Codd, quindi per ogni dipendenza funzionale non banale $X \rightarrow Y$ X deve essere una superchiave della tabella.

Qui sotto sono elencate le dipendenze funzionali per ogni tabella e specificato se le relazioni sono in BCNF oppure no.

1. **Film**(ID, Descrizione, Durata, Titolo, Valutazione, PremiatoP, PremiatoR, PremiatoA, PaeseProduzione, AnnoProduzione)
 $ID \rightarrow$ Descrizione, Durata, Titolo, Valutazione, PremiatoP, PremiatoR, PremiatoA, PaeseProduzione, AnnoProduzione. Quindi è in BCNF.
2. **Regia**(Film, Regista)
3. **Interpretazione**(Film, Attore)
4. **Operatore**(ID, Nome, Cognome, Popolarità, Ruolo)
 $ID \rightarrow$ Nome, Cognome, Popolarità, Ruolo. Quindi è in BCNF
5. **PremiazioneR**(NomePremio, CategoriaPremio, Regista, Data)
6. **PremiazioneA**(NomePremio, CategoriaPremio, Attore, Data)
7. **PremiazioneF**(NomePremio, CategoriaPremio, Film, Data)
8. **Premio**(Nome, Categoria, Peso)
 $Nome, Categoria \rightarrow Peso$. Quindi è in BCNF.
9. **RecensioneC**(Film, Critico, Voto, Testo)
 $Film, Critico \rightarrow Voto, Testo$. Quindi è in BCNF.
10. **Critico**(ID, Nome, Cognome)
 $ID \rightarrow Nome, Cognome$. Quindi è in BCNF.
11. **Linguaggio**(Film, NomeLingua, TipologiaLingua)
12. **Lingua**(NomeLingua, Tipologia)
13. **Tipologia**(Film, Genere)

14. **Genere**(ID, Descrizione)
 $ID \rightarrow Descrizione$. Quindi è in BCNF.
15. **Paese**(Nome, Longitudine, Latitudine)
 $Nome \rightarrow Longitudine, Latitudine$. Quindi è in BCNF.
16. **Connessione**(IP, InizioConnessione, FineConnessione, Hardware, Paese, Utente, Film)
 $IP, InizioConnessione, FineConnessione \rightarrow Hardware, Paese, Utente, Film$.
 $IP \rightarrow Paese$.
Quindi non è in BCNF.
17. **RecensioneU**(Film, Utente, Voto, Testo)
 $Film, Utente \rightarrow Voto, Testo$. Quindi è in BCNF.
18. **Utente**(ID, Nome, Cognome, Email, Password, DataNascita, SottoscrizioneAbbonamento, DataInizioAbb, DataFineAbb)
 $ID \rightarrow Nome, Cognome, Email, Password, DataNascita, SottoscrizioneAbbonamento, DataInizioAbb, DataFineAbb$.
 $Email \rightarrow ID, Nome, Cognome, Password, DataNascita, SottoscrizioneAbbonamento, DataInizioAbb, DataFineAbb$.
La relazione è in BCNF in quanto l'Email è identificatore della relazione ma, come richiesto nelle specifiche, è stato comunque aggiunto un codice identificativo come identificatore primario.¹²
19. **Fattura**(Codice, DataEmissione, Quota, Scadenza, DataPagamento, FatturazioneUtente, CartaPagamento)
 $Codice \rightarrow DataEmissione, Quota, Scadenza, DataPagamento, FatturazioneUtente, CartaPagamento$. Quindi è in BCNF.
20. **Carta**(PAN, CVV, DataScadenza, NomeIntestatario, CognomeIntestatario, UtenteProprietario)
 $PAN \rightarrow CVV, DataScadenza, NomeIntestatario, CognomeIntestatario, UtenteProprietario$. Quindi è in BCNF.
21. **Abbonamento**(Tipo, Tariffa, Durata, MaxOre, MaxGB, EtaMin)
 $Tipo \rightarrow Tariffa, Durata, MaxOre, MaxGB, EtaMin$. Quindi è in BCNF.
22. **Offerta**(Abbonamento, Funzionalità)
23. **Funzionalità**(Nome, Descrizione)
 $Nome \rightarrow Descrizione$. Quindi è in BCNF.

¹²Si noti che si potrebbe pensare che SottoscrizioneAbbonamento e DataInizioAbb possano determinare DataFineAbb, ho però pensato che i Termini di Servizio riguardo alla durata degli abbonamenti possano cambiare da un giorno all'altro e quindi non rendere più corretta la dipendenza funzionale.

24. **Restrizione**(Abbonamento, StatoRestrittivo)

25. **StatoRestrittivo**(Stato, RestrizioneF, RestrizioneA)
 $\text{Stato} \rightarrow \text{RestrizioneF}$, RestrizioneA . Quindi è in BCNF.

26. **Esclusione**(StatoRestrittivo, CodiceFormato)

27. **Formato**(Codice, DataAggiornamento, Bitrate, TipoFormato, QualitàA, QualitàV, DimensioneFile, Lunghezza, RapportoAspetto, Risoluzione)
 $\text{Codice}, \text{DataAggiornamento} \rightarrow \text{Bitrate}, \text{TipoFormato}, \text{QualitàA}, \text{QualitàV}, \text{DimensioneFile}, \text{Lunghezza}, \text{RapportoAspetto}, \text{Risoluzione}$. Quindi è in BCNF.

28. **Aggiornamento**(CodiceFormato, Codec, Data)
 $\text{CodiceFormato}, \text{Codec} \rightarrow \text{Data}$. Quindi è in BCNF.

29. **Codec**(FileCodec, Specifiche)
 $\text{FileCodec} \rightarrow \text{Specifiche}$. Quindi è in BCNF.

30. **Specifiche**(Film, CodiceFormato, DataAggiornamento, DataRilascio)
 $\text{Film}, \text{CodiceFormato}, \text{DataAggiornamento} \rightarrow \text{DataRilascio}$. Quindi è in BCNF.

6.4 Normalizzazione

L'unica relazione non in BCNF è Connessione, applichiamo dunque l'algoritmo per decomposizione in BCNF.

Abbiamo ottenuto 2 relazioni:

- IPPaese(IP, Paese) con Dipendenza Funzionale:
 $\text{IP} \rightarrow \text{Paese}$.

- DettagliConnessione(IP, InizioConnessione, FineConnessione, Hardware, Utente, Film) con Dipendenza Funzionale:
 $\text{IP}, \text{InizioConnessione}, \text{FineConnessione} \rightarrow \text{Hardware}, \text{Utente}, \text{Film}$.

6.5 Trattazione vincoli

I vincoli di integrità referenziale sono stati gestiti attraverso l'uso delle *FOREIGN KEY*.

I vincoli intra-relazionali, come già specificato sono stati gestiti tramite *CHECK*.

I vincoli Inter-Relazionali ed altri vincoli non presenti nella documentazione, ma rilevanti ai fini della gestione del Database, sono stati gestiti tramite l'utilizzo di trigger, qui di seguito sono riportati i nomi di essi:

1. validate_regista_before_insert
2. validate_attore_before_insert
3. check_anno_produzione_before_insert_premio
4. check_anno_produzione_before_insert_specifiche
5. check_anno_produzione_before_insert_visualizzazione
6. check_categoria_premio_before_insert_premiazioneR
7. check_categoria_premio_before_insert_premiazioneA
8. check_categoria_premio_before_insert_premiazioneF
9. SetDataFineAbb_before_insert
10. CheckAge_
11. trg_before_insert_restrizione
12. before_insert_esclusione

6.6 Events

Tramite gli Event sono stati gestiti due casi:

- Il primo caso riguarda il reset degli attributi riguardanti l'abbonamento di un Utente una volta scaduto il proprio abbonamento. Esso è gestito tramite un Event *ResetAbbonamentiScaduti_* eseguito una volta al giorno.
- Il secondo caso riguarda la cancellazione dal Database delle Connessioni che "compiono" un Anno, gestito tramite l'Event *evento_cancellazioneconnessioni* eseguito una volta al giorno.

6.7 Trattazione Ridondanze

Come già specificato precedentemente la ridondanza PremiatoR e di conseguenza anche PremiatoA e PremiatoP sono gestite attraverso un Trigger per compiere un *Immediate Refresh*. Qui di seguito sono riportati i nomi dei trigger creati.

- premiatoR_immediateRefresh_
- premiatoA_immediateRefresh_
- premiatoP_immediateRefresh_

La Valutazione del Film è invece gestita attraverso un *Deferred Refresh* e quindi attraverso una Procedure *calcolaValutazioneFilm* che viene chiamata da un Event *aggiornaValutazione* ogni notte alle ore 01:00:00.

6.8 Popolamento

Il popolamento non rispetta i Volumi espressi nella Tavola dei Volumi, e non riesce dunque a rispecchiare le infinite possibilità e dunque le possibili fragilità del Database che possono sorgere in un contesto reale.

Ogni riferimento a persone esistenti o a fatti realmente accaduti è puramente casuale, dato che parte del Popolamento è avvenuto tramite generazione casuale con l'ausilio di ChatGpt 4.