



Politecnico
di Torino



Lab su progettazione mediante LLM: errori comuni

Luca Cagliero, Laura Farinetti, Giuseppe Gallipoli, Davide Napolitano
Dipartimento di Automatica e Informatica
Politecnico di Torino

Text2Schema

- Esempio di task da risolvere mediante LLM
 - Laboratorio
- Dato delle specifiche di progetto formulate in linguaggio naturale definire uno schema logico relazionale di una base di dati che rispetti le specifiche suddette
- Esempio di prompt con specifiche testuali
 - *«Si vuole progettare una base di dati relazionale per raccogliere le informazioni relative ai fornitori di un'azienda. I fornitori sono caratterizzati da un codice fornitore univoco; si vuole inoltre tenere traccia del nome ed eventualmente della sede del fornitore. Generare lo schema logico relazionale della base di dati sopra descritta»*
- Esempio di output atteso
 - Tabella Fornitori(CodiceFornitore, Nome, Sede*)

Demo Text2Schema su ChatGPT (<https://chatgpt.com/>)

INPUT

Si vuole progettare una base di dati relazionale per raccogliere le informazioni relative ai fornitori di un'azienda. I fornitori sono caratterizzati da un codice fornitore univoco; si vuole inoltre tenere traccia del nome ed eventualmente della sede del fornitore. Generare lo schema logico relazionale della base di dati sopra descritta.

OUTPUT 1/2

Per progettare uno schema logico relazionale per raccogliere le informazioni sui fornitori di un'azienda, seguiremo i requisiti descritti:

1. **Codice Fornitore:** Identificatore univoco per ciascun fornitore.
2. **Nome del Fornitore:** Nome del fornitore.
3. **Sede del Fornitore:** Informazione opzionale relativa alla sede.

Schema Logico Relazionale

Possiamo definire una tabella chiamata **Fornitori** che contenga le seguenti colonne:

Campo	Tipo	Descrizione
CodiceFornitore	VARCHAR o INT (PK)	Identificatore univoco del fornitore
Nome	VARCHAR	Nome del fornitore
Sede	VARCHAR (NULL)	Sede del fornitore (opzionale)

Chiave Primaria

- **CodiceFornitore** è la chiave primaria (PK) della tabella, che garantisce l'univocità di ogni fornitore.

Demo Text2Schema su ChatGPT (<https://chatgpt.com/>)

INPUT

Si vuole progettare una base di dati relazionale per raccogliere le informazioni relative ai fornitori di un'azienda. I fornitori sono caratterizzati da un codice fornitore univoco; si vuole inoltre tenere traccia del nome ed eventualmente della sede del fornitore. Generare lo schema logico relazionale della base di dati sopra descritta.

OUTPUT 2/2

Schema SQL

Ecco il codice SQL per creare questa tabella:

sql

Copia codice

```
CREATE TABLE Fornitori (  
  CodiceFornitore VARCHAR(50) PRIMARY KEY,  
  Nome VARCHAR(100) NOT NULL,  
  Sede VARCHAR(100)  
);
```

Spiegazione

- `CodiceFornitore` è definito come chiave primaria (`PRIMARY KEY`), quindi deve essere unico per ciascun fornitore.
- `Nome` è un campo obbligatorio (`NOT NULL`) che raccoglie il nome del fornitore.
- `Sede` è un campo facoltativo, quindi può essere nullo.

Questo schema soddisfa i requisiti per la gestione dei dati sui fornitori.

Relazioni molti a molti

SORGENTE	TIPOLOGIA DI ERRORE
LLM	chiave usata per identificare la N-N
Studente	attributo dell'entità anzichè della relazione
LLM	nonostante la specifica ambigua il modello riconosce il pattern
Studente	storicizzazione aggiunta in modo improprio
Studente	espliciti riferimenti a conoscenza di dominio in ambito db
Studente+LLM	mancono le foreign key
Studente	errore linguistico
LLM	non nullabilità errata
Studente	1-N anziché N-N

Generalizzazione totale, esclusiva

SORGENTE	TIPOLOGIA DI ERRORE
Studiante+LLM	tipo modellato in modo errato
Studiante	gerarchia non necessaria

Generalizzazione parziale, sovrapposta

SORGENTE	TIPOLOGIA DI ERRORE
Studente	Il testo indica una generalizzazione diversa da (p,s)
Studente+LLM	In caso di accorpamento nel padre gli attributi delle figlie non sono opzionali
LLM	Chiave primaria nelle tabelle figlie che non fa riferimento alla chiave della tabella padre
Studente	Le tabelle figlie non hanno nessuna chiave primaria

Attributo multivalore

SORGENTE	TIPOLOGIA DI ERRORE
Studente	multivalore confuso con relazione N-N
LLM	attributo composti impropriamente
LLM	tabella ridondante
Studente+LLM	attributo semplice confuso con multivalore
Studente	chiave errata nella tabella di traduzione del multivalore
Studente+LLM	sintassi errata

Entità debole

SORGENTE	TIPOLOGIA DI ERRORE
Studente+LLM	Nell'entità debole: non mettere la chiave esterna anche nella chiave primaria
Studente	Nell'entità debole non dichiarare la chiave esterna

Relazione binaria col tempo

SORGENTE	TIPOLOGIA DI ERRORE
Studente+LLM	Non c'è il tempo in chiave
Studente+LLM	Manca la relazione con il tempo
Studente+LLM	Storicizzata invece di binaria col tempo

Relazione ternaria col tempo

SORGENTE	TIPOLOGIA DI ERRORE
LLM	Chiave primaria (id) nella tabella della ternaria
Studente	Non c'è il tempo in chiave
Studente+LLM	Manca la tabella della relazione ternaria

Entità storicizzata

SORGENTE	TIPOLOGIA DI ERRORE
Studente	Manca la relazione storicizzata (sostituita con relazioni 1-N)
Studente+LLM	Confusa con relazione ternaria
Studente+LLM	Modellazione errata

Copyright

Attribution + Noncommercial + NoDerivative