

MODELLO RELAZIONALE

NICOLA MORETTO (MATR. 578258)

4 ottobre 2012

Il documento presenta il modello relazionale per l'integrazione dei nuovi criteri di classificazione nella piattaforma.

VERSIONE	DATA	MODIFICHE
0.1	2-10-2012	Prima stesura del documento.
0.2	3-10-2012	Aggiunto lo SCHEMA RELAZIONALE.
0.3	3-10-2012	Redazione della sezione OPERAZIONI.
0.4	4-10-2012	Aggiornamento e revisione della sezione OPERAZIONI.
0.5	5-10-2012	Redazione della sezione SCHEMA RELAZIONALE.

Tabella 1: Registro delle modifiche

INDICE

1	SCHEMA RELAZIONALE	5
1.1	Sintassi	6
1.1.1	etichette	6
1.2	Semantica	6
1.2.1	entità	6
1.2.2	tipi_entità	6
1.2.3	sinonimi	6
1.2.4	gerarchia	6
1.3	Contenuti	6
1.3.1	classificazione	6
2	OPERAZIONI	7
2.1	Etichette	7
2.1.1	Ricerca di un'etichetta	7
2.1.2	Ricerca delle accezioni di un'etichetta	7
2.1.3	Ricerca dell'etichetta primaria	8
2.1.4	Ricerca del numero di assegnazioni di un'etichetta	8
2.1.5	Inserimento di un'etichetta	8
2.1.6	Eliminazione di un'etichetta	9
2.1.7	Calcolo di affinità tra due etichette primarie	9
2.2	Accezioni	9
2.2.1	Eliminazione di un'accezione	9
2.3	Entità	10
2.3.1	Ricerca dei figli	10
2.3.2	Ricerca dei padri	10
2.3.3	Ricerca delle etichette	11
2.3.4	Aggiungere una relazione padre-figlio	11
2.4	Contenuti	11
2.4.1	Assegnazione di un'etichetta	11
2.4.2	Ricerca delle etichette assegnate	12
2.4.3	Ricerca di contenuti generici	12
2.4.4	Ricerca di contenuti specifici	12

SCHEMA RELAZIONALE

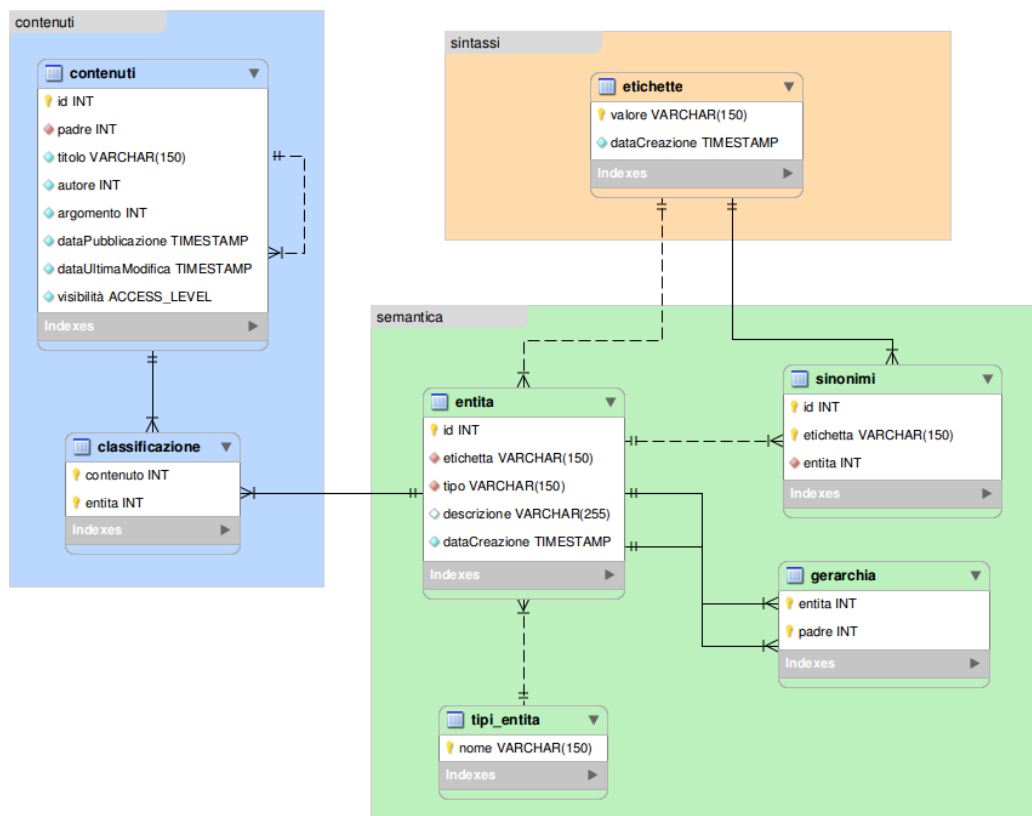


Figura 1: Schema relazionale del sistema di classificazione

La progettazione dello schema relazionale per il sistema di classificazione ha tenuto conto di diverse esigenze:

Obiettivi

- separare le informazioni sintattiche (ETICHETTE) da quelle semantiche (ACCEZIONI, ENTITÀ);
- evitare di dover ricorrere ad interrogazioni complesse ed onerose sul piano computazionale per reperire o inserire informazioni;
- garantire la massima flessibilità ed espandibilità per adattarsi alle evoluzioni tecnologiche della piattaforma (assegnazione automatica delle etichette, ricerca *full-text*, ...).

1.1 SINTASSI

1.1.1 *etichette*

Le etichette, intese come singole parole o brevi espressioni aventi lunghezza massima fissata (150 caratteri) e codificate secondo un formato standard, rappresentano lo strumento per identificare una determinata entità e sono a loro volta univocamente distinguibili per la struttura sintattica.

Un'etichetta possiede diverse accezioni, ciascuna delle quali individua un'entità ben precisa del dominio e può essere chiave o sinonimica, a seconda che identifichi univocamente l'entità in questione o meno. Nel primo caso l'etichetta viene definita primaria (rispetto all'entità), altrimenti secondaria.

Accezioni

1.2 SEMANTICA

1.2.1 *entità*

Ciascuna entità è identificata da un'etichetta primaria, o più precisamente da un'accezione chiave di tale etichetta. La tabella *entità* rappresenta allo stesso tempo le singole entità, cui associa la corrispondente etichetta primaria, e un'accezione chiave per la medesima etichetta.

Ciascuna etichetta può avere diverse accezioni chiave, ossia identificare univocamente diverse entità in altrettante accezioni, ma ciascuna entità ha una sola etichetta primaria.

1.2.2 *tipi_entità*

Le entità del dominio possono essere convenientemente distinte organizzate in sotto domini a seconda di ciò che rappresentano: un luogo, una persona, un evento, un prodotto, un concetto,

1.2.3 *sinonimi*

Ciascuna entità è richiamabile e riferibile dall'utente mediante etichette secondarie: le accezioni secondarie di ciascuna etichetta sono presenti in tale tabella e ciascuna di esse è associata chiaramente ad un'entità del dominio.

1.2.4 *gerarchia*

Le relazioni gerarchiche tra le etichette vengono conservate in questa tabella e vengono aggiornate automaticamente in caso di aggiornamento o cancellazione di un'accezione chiave/entità.

1.3 CONTENUTI

1.3.1 *classificazione*

A ciascun contenuto possono essere assegnate diverse etichette primarie, ciascuna declinata in un'accezione specifica che individua precisamente l'entità riferita nel contenuto.

OPERAZIONI

In questo capitolo vengono illustrate le operazioni essenziali di gestione e utilizzo del sistema di classificazione progettato ed implementato secondo le specifiche descritte nella sezione 1.

Per ciascuna operazione, si descrive lo scopo e si fornisce il codice *SQL* con cui realizzarle, motivando ove opportuno le scelte di progettazione del database e mostrando i vantaggi ottenuti.

Di seguito si assume che le etichette rispettino un formato standard, a garanzia dell'assenza di duplicati; le variabili nel codice *SQL*, il cui valore è noto al momento dell'esecuzione dell'istruzione, sono riconoscibili dal prefisso '@'.

2.1 ETICHETTE

Le etichette rappresentano la componente sintattica mediante la quale gli utenti interagiscono con il sistema di classificazione, catalogando e recuperando i contenuti.

2.1.1 Ricerca di un'etichetta

Una stringa digitata dall'utente può rappresentare il prefisso di un'etichetta o l'etichetta stessa: in entrambi i casi per cercare riscontri nel dizionario è sufficiente considerare la tabella *etichette*.

CORRISPONDENZA ESATTA

Restituisce - ove presente - l'etichetta corrispondente alla stringa inserita.

```
SELECT *
  FROM etichette
 WHERE valore='@etichetta'
```

CORRISPONDENZA PARZIALE

Restituisce - ove presenti - le etichette corrispondenti alla stringa inserita o di cui quest'ultima sia prefisso.

```
SELECT *
  FROM etichette
 WHERE valore LIKE '@etichetta%'
```

2.1.2 Ricerca delle accezioni di un'etichetta

Individuata un'etichetta, è spesso necessario recuperare le relative accezioni per identificare un'entità precisa del dominio. Tale operazione richiede di consultare separatamente le tabelle *entita* e *sinonimi* e di unire i risultati, eventualmente ordinandoli alfabeticamente.

```
{
  SELECT id
    FROM entita
```

```

        WHERE etichetta=@etichetta'
    }
    UNION
    {
        SELECT entita
            FROM sinonimi
            WHERE etichetta=@etichetta'
    }
    ORDER BY id

```

2.1.3 Ricerca dell'etichetta primaria

Se l'etichetta inserita dall'utente risulta - rispetto ad un'accezione - secondaria, è necessario risalire alla corrispondente primaria. Ciò richiede di consultare le tabelle entita e sinonimi: è possibile effettuare tale operazione mediante un'unica istruzione SQL, che richiede il JOIN tra le due tabelle, o due distinte.

ISTRUZIONE SINGOLA

```

SELECT e.etichetta
    FROM entita AS e JOIN sinonimi AS s
        ON (e.id=s.entita)
    WHERE s.etichetta=@etichetta' AND s.id=@id'

```

DUE ISTRUZIONI

```

SELECT entita
    FROM sinonimi
    WHERE s.etichetta=@etichetta' AND s.id=@id'

```

```

SELECT etichetta
    FROM entita
    WHERE c.etichetta=@entita'

```

2.1.4 Ricerca del numero di assegnazioni di un'etichetta

Per determinare il numero di contenuti cui sia stata assegnata una data etichetta è sufficiente accedere alla tabella classificazione.

```

SELECT COUNT(*) AS num
    FROM classificazione
    WHERE etichetta=@etichetta'

```

2.1.5 Inserimento di un'etichetta

L'inserimento di una nuova etichetta si rende necessario qualora un utente cerchi di riferire un'entità con una stringa (parola o breve espressione) non presente nel dizionario (vedi sezione 2.1.3).

A questo punto l'etichetta viene inserita nella tabella etichette:

```

INSERT INTO etichette
    VALUES (@stringa,current_timestamp)

```


Infine si deve associarle almeno un'accezione, che chiarisca l'entità cui si riferisce: tale operazione può essere svolta dall'utente o in maniera (semi)automatica, mediante l'assegnazione di valori predefiniti.¹

Di seguito si assume che l'utente abbia stabilito se si tratti di etichetta primaria (identifica una nuova entità) o secondaria (identificatore alternativo per un'entità definita), sfruttando eventualmente l'organizzazione per tipi.

ACCEZIONE CHIAVE

Se l'etichetta è primaria, si fornisce una descrizione dell'entità in questione.

```
INSERT INTO entita(etichetta, tipo)
VALUES (@etichetta, @tipo)
```

ACCEZIONE SINONIMICA

Se l'etichetta è secondaria, è necessario indicare la corrispondente accezione chiave.

```
INSERT INTO sinonimi(etichetta, entita)
VALUES (@etichetta, @entita)
```

2.1.6 Eliminazione di un'etichetta

Il processo di eliminazione di un'etichetta richiede di prendere in esame ciascuna accezione ed eliminarla, secondo le modalità specificate nella sezione 2.2.1.

Solo una volta cancellate le relative accezioni, è possibile procedere all'eliminazione dell'etichetta dal dizionario.

```
DELETE FROM etichette
WHERE valore=@etichetta
```

2.1.7 Calcolo di affinità tra due etichette primarie

L'affinità tra due etichette primarie è espressa dal numero di contenuti in cui compaiano entrambe le entità riferite.

```
SELECT contenuto, COUNT(*)
FROM classificazione
WHERE entita='@entita1' OR entita='@entita2'
GROUP BY contenuto
HAVING COUNT(*)>=2
```

2.2 ACCEZIONI

2.2.1 Eliminazione di un'accezione

L'eliminazione di un'accezione associata ad un'etichetta richiede di considerare due casi distinti, a seconda che l'accezione stessa sia chiave o sinonimica.

ACCEZIONE SINONIMICA

In questo caso si può procedere all'eliminazione diretta, poiché l'entità in questione continua ad essere identificata da un'altra etichetta primaria (la cui accezione è chiave).

¹ L'assegnazione automatica dell'accezione è possibile solo per la prima.

```
DELETE FROM sinonimi
      WHERE etichetta='@etichetta' AND id='@id'
```

ACCEZIONE CHIAVE

In quest'altro caso, prima di procedere all'eliminazione, occorre individuare un'etichetta secondaria da rendere primaria, promuovendo l'accezione corrispondente da sinonimica (tabella sinonimi) a chiave (tabella entita).

Per prima cosa si individua un'accezione sinonimica con cui sostituire l'accezione chiave (ove non ve ne sia nessuna, l'intera operazione va annullata).

```
SELECT *
      FROM sinonimi
      WHERE entita='@entita'
      LIMIT 1
```

A questo punto si sostituisce l'accezione chiave con la sinonimica, aggiornando l'etichetta assegnata alla prima con il valore della seconda. Così facendo le restanti accezioni sinonimiche riferenti la medesima entità si aggiorneranno automaticamente.

```
UPDATE entita
      SET etichetta='@etichetta_sinonimica'
      WHERE id='@entita'
```

Infine si elimina l'accezione sinonimica.

```
DELETE FROM sinonimi
      WHERE etichetta='@etichetta_sinonimica' AND id='@id'
```

2.3 ENTITÀ

Ciascuna entità è identificata univocamente nel sistema di classificazione da un'accezione chiave e - sul piano sintattico - dalla relativa etichetta primaria. Per tale motivo la maggior parte delle operazioni coinvolge le tabelle *accezioni_chiave* e *gerarchia*.

2.3.1 Ricerca dei figli

La ricerca dei figli di un'entità consiste nell'individuare le entità di cui essa sia il padre.

```
SELECT e.etichetta AS figlio
      FROM gerarchia AS g JOIN entita AS e
            ON (g.figlio=e.id)
      WHERE padre='@entita'
```

2.3.2 Ricerca dei padri

La ricerca dei padri di un'entità risulta analoga e simmetrica rispetto alla precedente.

```
SELECT e.etichetta AS padre
      FROM gerarchia AS g JOIN entita AS e
            ON (g.padre=e.id)
      WHERE figlio='@entita'
```

2.3.3 Ricerca delle etichette

Per ottenere la lista completa delle etichette con cui è riferibile una determinata entità del dominio, è sufficiente recuperare l'etichetta primaria e quelle sinonimiche.

```
{
  SELECT etichetta
    FROM entita
   WHERE entita='@entita'
}
UNION
[
  SELECT etichetta
    FROM sinonimi AS s JOIN entita AS e
      ON (s.entita=e.id)
   WHERE s.entita='@entita'
]
```

2.3.4 Aggiungere una relazione padre-figlio

L'aggiunta di una relazione padre-figlio tra entità esistenti richiede di specificare padre e figlio nella tabella gerarchia.

```
INSERT INTO gerarchia(entita,padre)
VALUES (@figlio,@padre)
```

Un controllo elementare per evidenziare l'incompatibilità della relazione tra due entità consiste nel verificare se risulti già definita una relazione ove la coppia (padre,figlio) sia scambiata. L'operazione può essere effettuata automaticamente all'inserimento della nuova riga grazie ad un trigger.

```
DELIMITER $
CREATE TRIGGER controlloRelazionePadreFiglioEntita
BEFORE INSERT ON classificazione
FOR EACH ROW BEGIN
  DECLARE records INT;

  SELECT COUNT(*) INTO records
    FROM classificazione
   WHERE padre=NEW.figlio AND figlio=NEW.padre;

  IF records>0 THEN
    SIGNAL SQLSTATE '45000'
    SET MESSAGE_TEXT = 'Relazione padre-figlio inversa già presente';
  END IF;
END; $
DELIMITER ;
```

2.4 CONTENUTI

2.4.1 Assegnazione di un'etichetta

L'assegnazione di un'etichetta (più precisamente un'accezione chiave) ad un contenuto assume che sia stata individuata l'etichetta primaria per l'entità in

questione (v. sezioni 2.1.1 e 2.1.2). A questo punto si procede inserendo una nuova riga nella tabella `classificazione`.

```
INSERT INTO classificazione(contenuto,entita)
VALUES (@contenuto,@entita)
```

2.4.2 Ricerca delle etichette assegnate

Il reperimento della lista delle etichette primarie assegnate ad un contenuto richiede di consultare le tabelle `classificazione` ed `entita`.

```
SELECT e.etichetta
FROM classificazione AS c JOIN entita AS e
ON (c.entita=e.id)
WHERE contenuto='@contenuto'
```

2.4.3 Ricerca di contenuti generici

La ricerca dei contenuti cui sia stata assegnata una certa etichetta richiede di consultare la tabella `classificazione`.

```
SELECT contenuti
FROM classificazione
WHERE entita='@entita'
```

2.4.4 Ricerca di contenuti specifici

La ricerca di contenuti specifici consiste nell'individuare i contenuti cui siano state assegnate - tra le altre - determinate etichette (due o più) specificate dall'utente.

```
SELECT contenuto, COUNT(etichetta)
FROM classificazione
WHERE entita='@entita1' [OR entita='@entita2' ...]
GROUP BY contenuto
HAVING COUNT(*)>=@num_etichette
```