

## Modello del Dominio

### Survey

Il Modello del Dominio, che costituisce il deliverable dell'attività B.4, è incentrato sulla rappresentazione degli oggetti della vendita, ossia i prodotti. In particolare, tale modello definisce le entità (i.e., Categoria, Prodotto e Bundle), le proprietà (e.g., anagrafica prodotto, etc.) e le relazioni (e.g., similarità, accessorietà, etc.) utili a rappresentare e classificare i prodotti.

In termini di innovazione, la modellazione dei dati in InViMall, è incentrata sull'uso pervasivo di tecnologie emergenti dell'area della rappresentazione della conoscenza, ossia le tecnologie semantiche. Infatti, i modelli previsti in InViMall saranno progettati in accordo ad una vista concettuale (Ontologie), che descrive il dominio in cui ci muoviamo, piuttosto che mediante un approccio orientato ai dati (ossia, Database Relazionale). In particolare, la definizione del modello del dominio è principalmente rivolta alla rappresentazione delle sole informazioni abilitanti particolari benefici funzionali e, pertanto, sono tralasciate informazioni peculiari per le quali non è utile una rappresentazione ontologica e che sono tipicamente già disponibili e gestite da una generica piattaforma di e-commerce disponibile allo stato dell'arte.

L'adattabilità del modello del dominio a diverse aree merceologiche implica che la definizione del modello esibisca caratteristiche di flessibilità e scalabilità rispetto alla inclusione nel progetto InViMall di altri domini merceologici di vendita anche non specificatamente analizzati. Ciò è ottenibile dall'adozione di un approccio che alterni fasi di analisi bottom-up e fasi di analisi top-down al fine di determinare a partire dallo studio di domini specifici le informazioni applicabili in senso generale e viceversa. Nello specifico, saranno preferite modalità di rappresentazione che consentano di separare la logica di rappresentazione dei prodotti dallo specifico dominio in cui essi sono collocabili mediante la separazione, ad esempio, delle dimensioni di classificazione (e.g., caratteristiche tecniche, etc.) dei prodotti da entità utili alla rappresentazione del prodotto stesso (e.g., prezzo, etc.).

Le entità concettuali e le modalità di rappresentazione dei prodotti mediante il modello del dominio sono rivolte a supportare:

- attività di information filtering guidate dalla semantica dei criteri di classificazione;
- attività di query augmentation predisponendo le basi per ampliare, attraverso opportuni suggerimenti, il raggio di interesse dell'utente nonché il soddisfacimento delle desiderata d'acquisto;
- aggregazione dei risultati anche sulla base della composizione di offerte per prodotti composti ottenuti per assemblaggio di prodotti elementari.

Nella descrizione del modello del dominio le informazioni mantenute sono caratterizzate rispetto alla loro natura:

- esplicite, ovvero dichiarate quando si istanzia l'entità modellata, o la si aggiorna. Al più, potrebbero essere dichiarate esplicite anche le informazioni dedotte in modo certo ed inequivocabile;

- implicite, ovvero dedotte attraverso meccanismi di inferenza basati sull'analisi delle informazioni esplicite mantenute sia localmente, nel modello cui l'entità è definita, che globalmente, negli altri modelli correlati. In alcuni casi, i meccanismi di inferenza potrebbero lavorare anche su alcune informazioni implicite, evitando, però, di incorrere in fenomeni di loop;

e rispetto alla loro validità temporale:

- a lungo termine, ovvero informazioni anche variabili nel tempo, ma che accompagnano le entità modellate per tempi lunghi;
- a breve termine, ovvero informazioni valide per periodo temporali ristretti, relativi, ad esempio, a particolari attività circoscritte nel tempo.

La Figura 1 (a) sintetizza la vista macroscopica delle varie tipologie di entità concettuali e le relazioni che tra queste sussistono, mentre la Figura 1 (b) mostra le modalità con cui sono rappresentati prodotti e venditori istanziando opportunamente le meta-informazioni disponibili all'interno del modello del dominio (Figura 1 (a)).

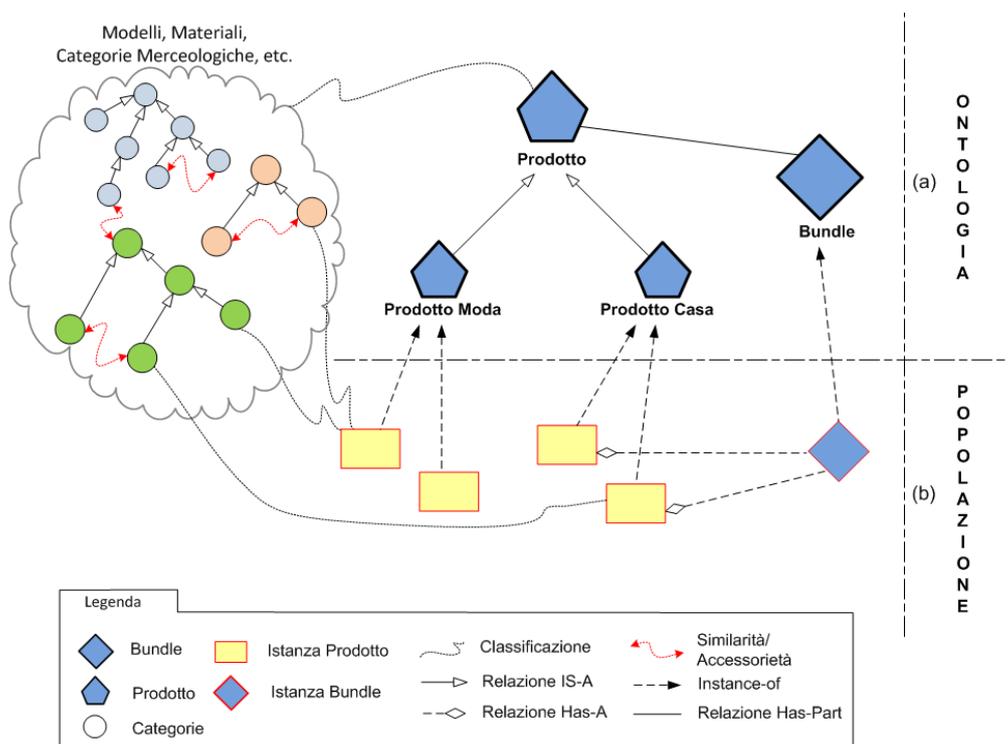


Figura 1. Rappresentazione dettagliata del modello del dominio distinguendo le entità concettuali dalla popolazione del modello

In particolare, le entità concettuali individuate nel Modello del Dominio sono:

- **Categoria:** identifica una etichetta utile a classificare i prodotti in base a specifici criteri di classificazione. Tali criteri spesso coincidono con le caratteristiche tecniche del prodotto stesso (e.g., tipologia, colore, materiale, dimensioni, etc.). Dunque, i valori assunti dall'entità Categoria rappresentano l'insieme delle possibili classi cui un prodotto può appartenere. Tali valori possono afferire a diverse dimensioni di classificazione ed in dipendenza di queste, possono essere tra loro gerarchizzate o meno. Infatti, se consideriamo l'insieme delle Categorie Merceologiche (e.g.,

Pantalone, Maglia, Maglia a maniche corte, etc.) i valori dell'entità Categoria sono relazionati tra loro attraverso relazioni di specializzazione/generalizzazione. Mentre, se consideriamo i "Colori" le categorie potrebbero essere semplicemente dei valori foglia disgiunti e rappresentabili come tipo di dati enumerato. In particolare, distinguiamo diverse tipologie di criteri di classificazione in base alle modalità con cui sono strutturati i relativi valori dell'entità Categoria, ossia:

- *Tassonomie e vocabolari controllati*, i cui valori sono organizzati attraverso relazioni semantiche di tipo "broader term", "narrower term", "similar to". Tra i criteri di classificazione inclusi in tale tipologia, il modello del dominio in InViMall intende supportare indipendentemente dalla filiera di riferimento la Categoria Merceologica, il Designer, il Target cui i prodotti sono rivolti, i Materiali dei prodotti.
- *Enumerati*, che ammette valori foglia non necessariamente relazionati tra loro. Sono tipi di dati che fanno parte di un predefinito insieme, statico, di valori con uno specifico ordine. Tra i criteri di classificazione inclusi in tale tipologia, il modello del dominio in InViMall intende supportare indipendentemente dalla filiera di riferimento i Colori e lo Stile.
- *Continui*, che ammette valori nel continuo ma che possono essere opportunamente discretizzati attraverso opportune procedure di analisi automatica dei dati. In tal senso, si ritiene di estrarre opportuni range (e.g., fasce di prezzo, etc.) in modo da ottenere una classificazione dei prodotti costruibile on-the-fly sulla base della discretizzazione dei valori continui assunti dai prodotti selezionati. Tra i criteri di classificazione inclusi in tale tipologia, il modello del dominio in InViMall intende supportare indipendentemente dalla filiera di riferimento, il criterio "Fasce di prezzo".
- *Composti*, che ammette valori derivanti dalla composizione di due o più campi eterogenei afferenti alle tipologie sopra indicate (e.g., tassonomie, etc.). Dunque la categoria in questione, diventa la composizione dei valori assunti dai vari campi. Ad esempio, se consideriamo la composizione "viscosa: 90%", i campi sono due, il valore del primo campo è il materiale, e il valore del secondo campo è un valore continuo o discreto.

La Figura 2 riassume in linea di massima le informazioni individuate per l'entità Categoria. In particolare, la figura differenzia la natura, implicita od esplicita, e la validità, a breve o lungo termine, delle proprietà individuate.

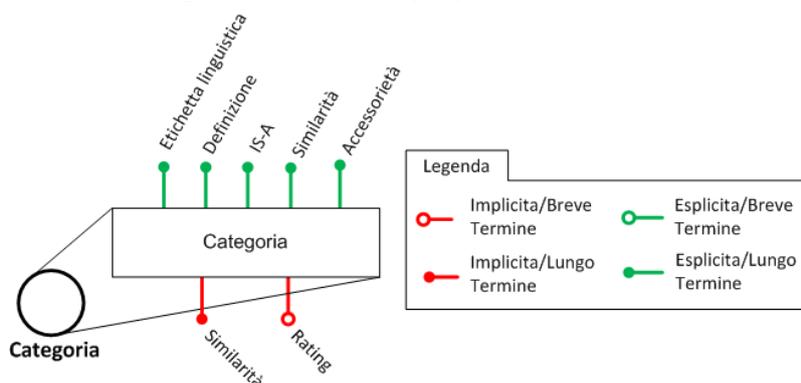


Figura 2. Informazioni di massima associate al concetto Categoria

- Prodotto:** ossia l'entità che definisce gli oggetti destinati alla vendita. Nell'ottica del riuso di talune caratteristiche definite nella rappresentazione di altri prodotti, il modello del dominio intende supportare la gerarchizzazione dell'entità Prodotto che, dunque, può essere opportunamente estesa. Infatti, un'adeguata gerarchizzazione del concetto Prodotto, consente sia di differenziare tipologie di prodotto aventi caratteristiche specifiche e non comuni a tutti gli altri sia di ereditare proprietà e relazioni già definite. In particolare, nel contesto applicativo di InViMall si intende definire una piattaforma di e-commerce in grado di supportare più filiere di vendita eventualmente aggiunte anche in fasi successive alla progettazione del sistema stesso. Dunque, il modello del dominio prevede la specializzazione dell'entità Prodotto per differenziare prodotti di filiere eterogenee (i.e., moda, casa, alimentari, etc.) ed eventualmente utili a rappresentare prodotti con caratteristiche peculiari. In particolare, la gerarchizzazione del nodo Prodotto non è utile a veicolare la rintracciabilità dell'oggetto ma a definire relazioni altrimenti non includibili nel modello (e.g., un prodotto di moda probabilmente avrà una taglia, un prodotto alimentare probabilmente avrà un peso, un prodotto casa avrà peso, lunghezza, spessore e larghezza, e così via).

La Figura 3 riassume in linea di massima le informazioni individuate per l'entità Prodotto. In particolare, la figura differenzia la natura, implicita od esplicita, e la validità, a breve o lungo termine, delle proprietà individuate che saranno dettagliate nelle sezioni seguenti.

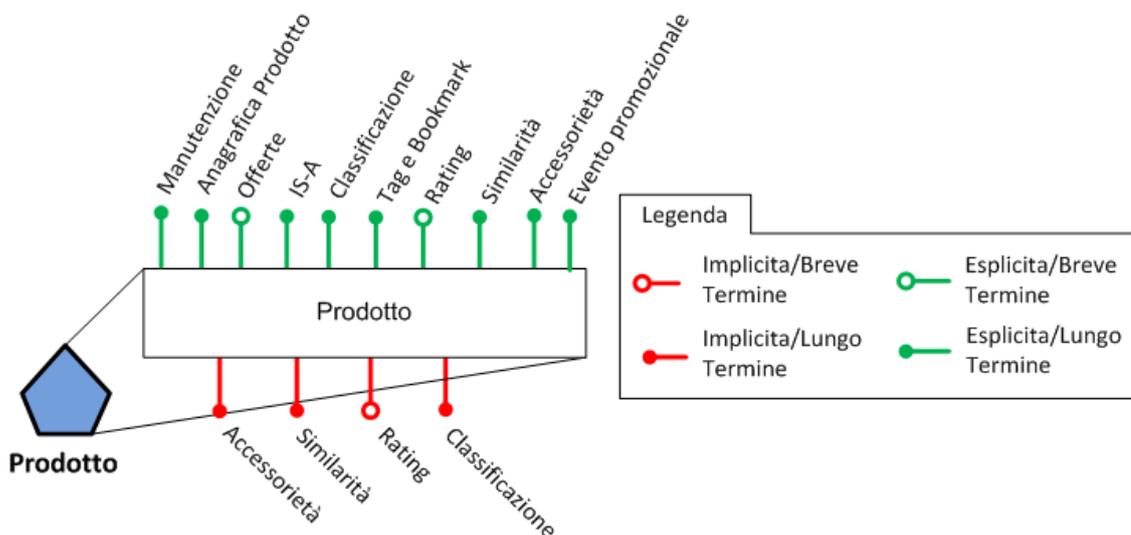


Figura 3. Informazioni di massima associate al concetto Prodotto

- Bundle:** ossia l'entità che rappresentano un prodotto composto (i.e., insieme di prodotti elementari) che può essere indicizzato, reperito e venduto come elemento singolo. Mediante relazioni di meronimia ed olonimia è possibile predisporre nel modello di dominio di InViMall la vendita di oggetti composti ottenuti come aggregazione di prodotti elementari. Nello specifico, il concetto di Bundle rappresenta un prodotto bundle che può essere indicizzato, reperito e venduto come elemento singolo. Il Bundle può essere definito sia in modo esplicito, composizione espressamente definita dal gestore della vendita o dai clienti (supervisionata dai

venditori), che in modo implicito sulla base di specifiche regole di composizione/compatibilità estratte dai Bundle definiti in modo esplicito.

La Figura 4, riassume in linea di massima le informazioni individuate per il Bundle. In particolare, la figura differenzia la natura, implicita od esplicita, e la validità, a breve o lungo termine, delle proprietà individuate che saranno dettagliate nelle sezioni seguenti.

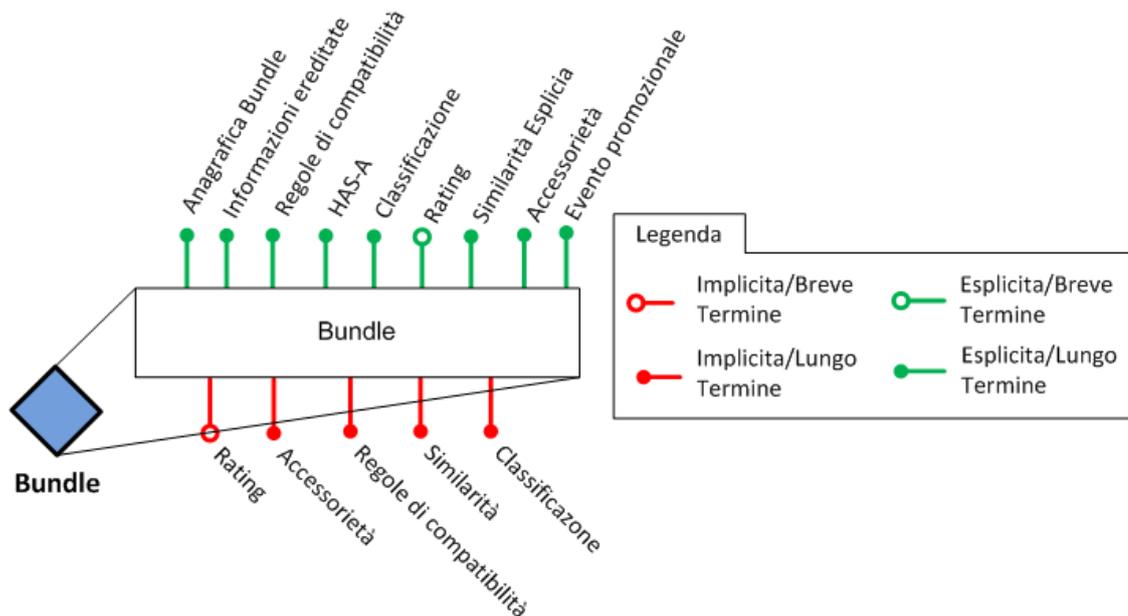


Figura 4. Informazioni di massima associate al concetto Bundle

Le informazioni esplicite individuate saranno definite da chi disegna (e/o gestisce) la base di conoscenza di InViMall (esperti del dominio), e si differenziano in proprietà e relazioni. Per quanto riguarda la deduzione delle informazioni implicite associate a ciascuna entità concettuale, sono state definite differenti tecniche:

- **Tecniche di aggregazione dei Rating puntuali.** Tali tecniche definiscono le modalità di aggregazione delle informazioni puntuali dei rating impliciti ed espliciti provenienti dal modello del Cliente e relativi a Prodotto, Bundle e Categoria. Pertanto, il punteggio (i.e., rating) complessivo di un'entità, non è un altro rating indipendente dagli altri, ma, è visto come una funzione di aggregazione  $f$  dei singoli rating forniti per quell'entità dall'utente. Formalmente:

$$f(r_1, \dots, r_k)$$

In altre parole, questo approccio presuppone che il rating complessivo (esplicito o implicito) ha una certa relazione con i singoli rating. Nello specifico, la Figura 5 mostra le relazioni che sussistono tra i valori puntuali raccolti dal Modello del Cliente e i valori aggregati di rating esplicito ed implicito dedotti nel Modello del Dominio.

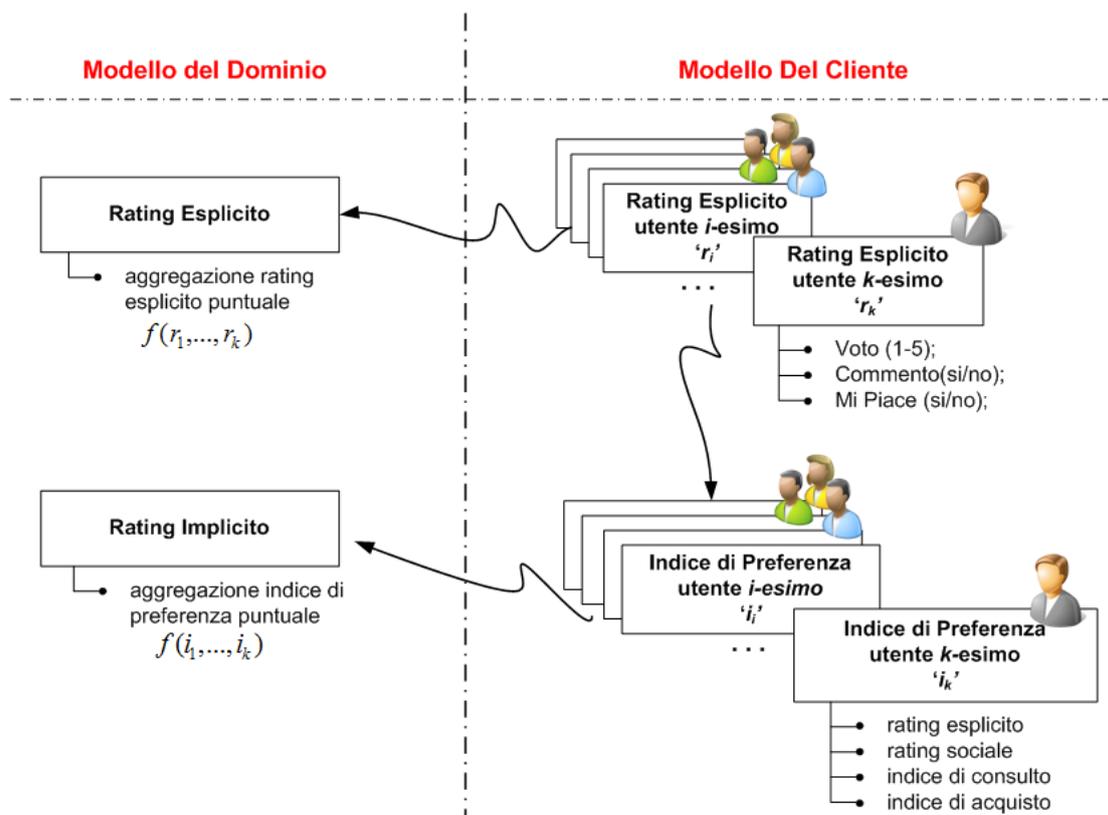


Figura 5. Aggregazione dei rating puntuali

- **Tecniche di supporto alla deduzione della Similarità implicita.** Utili a supportare l'implicazione di selezione e l'ampliamento della classificazione dei prodotti, tali tecniche definiscono le modalità con cui calcolare similarità semantica sia su base esplicita, definite dall'esperto del dominio, che su base implicita, tra prodotti e tra categorie. Attraverso la definizione di opportune matrici di adiacenza "prodotto/prodotto", "categoria/categoria", etc. sarà possibile ottenere di volta in volta una misura sintetica della similarità (implicita e/o esplicita) tra i vari elementi;
- **Tecniche di supporto alla deduzione dell'Accessorietà implicita.** Sulla base dell'accessorietà esplicita definita tra le entità Categoria sarà possibile dedurre la relazione di accessorietà tra entità Prodotto per supportare l'implicazione d'acquisto. In tal senso, saranno introdotte matrici di adiacenza (che a differenza delle matrici di similarità non sono simmetriche) che rappresentano il valore di accessorietà tra prodotti;
- **Tecniche di supporto alla Classificazione dei prodotti.** Tecniche per la classificazione implicita adottate per ampliare la classificazione esplicita di un prodotto (definita dall'esperto del dominio e/o venditore) e quindi garantire una maggiore reperibilità dello stesso. In tal senso, sarà valutata la possibilità di supportare funzionalità di:
  - **Classificazione implicita:** Tale funzionalità è finalizzata ad arricchire la classificazione dei prodotti fornita in fase di catalogazione degli stessi, attraverso opportune tecniche di analisi dei dati, sfruttando le informazioni



*InViMall - Intelligent Virtual Mall*  
Progetto MSE MI01 – 123



(anche storiche) relative ai prodotti esistenti e le relative categorie disponibili sul mall virtuale.

- *Classificazione tag based:* Attraverso l'analisi dei tag attribuiti ai prodotti dagli utenti, clienti del mall, si vuole estrarre in modo implicito le categorie dei prodotti, e di conseguenza una classificazione dei prodotti rispetto a tali valori.
- *Deduzione di categorie implicite:* Analizzare i dati dei prodotti al fine di dedurre, sulla base dell'analisi di talune proprietà, criteri di classificazione non esplicitamente disponibili nei dati (i.e., deduzione "on the fly" di categorie quali ad esempio "fasce di prezzo" e deduzione di categorie aggregate).

\*\*\* QUESTA E' L'ULTIMA PAGINA DEL DOCUMENTO \*\*\*