

Metadati: Topic Maps

Marco Pirruccio
Riccardo Gentilucci



Topic Maps: introduzione

In seguito vedremo:

- ◆ Il modello Topic Map
- ◆ La sintassi di Topic Map: XTM
- ◆ Differenze e similitarità tra TM & RDF



Caratteristiche generali

Topic Maps è definito dallo standard ISO 13250 e rappresenta un meccanismo per la rappresentazione di metainformazioni, con l'intento di facilitare l'organizzazione, il mantenimento e la navigazione di risorse informative.

Un'istanza di documento Topic Maps (d'ora in poi *mappa*) crea una sovrastruttura semantica alla base informativa cui si riferisce.

Il paradigma Topic Maps può essere utilizzato in più contesti diversi, non necessariamente per l'indicizzazione e la caratterizzazione di risorse web; il web attraverso gli URI offre un meccanismo di indicizzazione che ne facilita l'utilizzo.

Permette la creazione di ontologie attraverso meccanismi di tipizzazione gerarchica e relazioni di classe-istanza e superclasse-sottoclasse tra concetti descritti.

Ontologia: l'insieme delle informazioni che definiscono la natura delle "cose" e le relazioni tra esse esistenti, relativamente ad un determinato dominio applicativo



Caratteristiche generali: i concetti

Nascono con l'obiettivo di creare un meccanismo di indicizzazione che si adatti ad ogni supporto informativo.

In genere, gli indici sono conformi alla struttura dell'informazione cui si riferiscono, quindi i loro modelli di meta-informazioni sono impliciti e diversi caso per caso.

Creando un modello formale generico per la descrizione di indici si ottiene l'interoperabilità necessaria ad automatizzare la loro gestione e la fusione di documenti di meta-informazioni aventi intersezione di contenuti.

Topic Maps definisce una serie di costrutti che creano una classificazione a livello di linguaggio delle meta-informazioni: i concetti che rappresentano i mattoni con cui costruire delle strutture semantiche.

Si possono distinguere tre costrutti fondamentali:

- ◆ Argomenti (d'ora in poi Topics)
- ◆ Occorrenze
- ◆ Associazioni



XTM: XML Topic Maps (1.0)

Il modello concettuale definito dallo standard ISO 13250 è indipendente da qualsiasi meccanismo di serializzazione.

Per favorire l'interscambio di mappe e l'automazione della loro gestione, un consorzio indipendente formato dalle principali aziende sostenitrici della tecnologia Topic Map ha creato XTM: grammatica formale per la rappresentazione di mappe, basata su XML e su XLink come meccanismo di referenziazione.

Esistono diversi formati di rappresentazione delle mappe. XTM, oltre ad essere il più diffuso, è recentemente diventato parte dello standard ISO.

Useremo questa grammatica negli esempi che seguono.



Topics

Sono il concetto fondamentale dello standard. Qualsiasi cosa che debba essere descritta viene fatta corrispondere ad un topic, cui eventualmente si associano ulteriori caratteristiche

Topic è un concetto simile a *risorsa* in RDF.

Tuttavia un topic può anche essere completamente scorrelato da qualsiasi riferimento a risorse esterne, e descrivere quindi un concetto astratto.

Le caratteristiche che possono essergli associate sono:

- ◆ un soggetto (subject identity)
- ◆ uno o più nomi (base-name), di cui uno indicabile come preferito nella visualizzazione
- ◆ uno o più tipi (relativamente ad una ontologia prestabilita)
- ◆ una o più occorrenze



Topic Names

E' possibile associare ai topic uno o più nomi per la visualizzazione.

Il doppio livello di tag è dovuto alla possibilità di contestualizzare il nome (vedi dopo), ossia associare degli attributi di validità o di significatività alla stringa che segue.

```
<topic id="mario_rossi">
  <baseName>
    <baseNameString>Rossi, Mario</baseNameString>
  </baseName>
  <baseName>
    <baseNameString>Mario Rossi</baseNameString>
  </baseName>
</topic>
```



Tipizzazione

La tipizzazione dei topic non è imposta, è il meccanismo essenziale mediante il quale utilizzare una ontologia preesistente.

Il tipo di un topic è un altro topic o un insieme di altri topic: è prevista la tipizzazione multipla.

Non esiste una separazione a livello di modello tra la parte dichiarativa e quella descrittiva di una mappa: non esistono vincoli che impongono ad un topic il ruolo di tipo o di istanza.

In assenza di un linguaggio di constraint che ponga dei vincoli, sta al creatore della mappa mantenere la coerenza interna.

Mediante il tag `<instanceOf>` di un `<topic>` è possibile specificare una lista di topic-types, che lo tipizzano.



Tipizzazione: Esempi

```
<topic id="tt-person">
  <baseName>
    <baseNameString>Person</baseNameString>
  </baseName>
</topic>
<topic id="tt-artist">
  <instanceOf>
    <topicRef xlink:href="#tt-person"/>
  </instanceOf>
  <baseName>
    <baseNameString>Artist</baseNameString>
  </baseName>
</topic>
<topic id="tt-musician">
  <instanceOf>
    <topicRef xlink:href="#tt-artist"/>
  </instanceOf>
  <baseName>
    <baseNameString>Musician</baseNameString>
  </baseName>
</topic>
```



Soggetto e Occorrenze

Sono i costrutti mediante i quali si creano correlazioni tra un topic di una mappa e le reali risorse informative.

- ◆ Il soggetto indica ciò che un topic rappresenta. Può essere qualsiasi risorsa indirizzabile, in particolare qualsiasi URI oppure un'altro topic, nel qual caso si attua una reificazione.
- ◆ Le occorrenze indicano una o più risorse indirizzabili (o definite online nella mappa stessa) nelle quali il concetto che il topic rappresenta “ricorre” in una qualsiasi accezione del termine “ricorrere”.

Distinguono i ruoli delle risorse referenziate all'interno di una mappa: una cosa è associare delle meta-informazioni ad una risorsa, un'altra è descrivere un concetto e associarlo ad una risorsa in cui tale concetto viene analizzato, espresso o semplicemente citato.

Un'occorrenza è una sottostruttura di un topic, intrinsecamente legata ad esso, caratterizzata da un ruolo (occurrence role type)



Soggetti e Occorrenze: Esempio

Esempio: la mia home page viene modellata con un topic che ha per soggetto il suo URL, e viene ulteriormente caratterizzata da un'occorrenza sulla pagina in cui è linkata, opportunamente tipizzata in modo da indicare il ruolo dell'occorrenza.

```
<topic id="pagina_di_link"/>
<topic id="la_mia_pagina_web">
  <subjectIdentity>
    <resourceRef xlink:href="http://cs.unibo.it/~gentiluc"/>
  </subjectIdentity>
  <occurrence>
    <instanceOf>
      <topicRef xlink:href="#pagina_di_link"/>
    </instanceOf>
    <resourceRef link:href="http://cs.unibo.it/students.html"/>
  </occurrence>
</topic>
```



Associazioni (1)

Sono il secondo costrutto cardine delle mappe. Un documento XTM è in effetti una collezione di elementi *topic* e *association*.

Servono (non sorprendentemente) a realizzare una relazione tra due o più topic, ma anche tra topic e risorse.

Si noti che già le occorrenze ed il soggetto realizzano un particolare tipo di associazione tra un topic e una risorsa: associazioni esplicite sono necessarie per tutti gli altri casi.

Esempio:

```
<association>
  <member>
    <topicRef xlink:href="#mario_rossi"/>
  </member>
  <member>
    <resourceRef xlink:href="http://www.mariorossi.com/index.html"/>
  </member>
</association>
```



Associazioni (2)

Anche le associazioni sono soggette a tipizzazione per evidenziare il tipo di relazione che sussiste fra i membri.

Le associazioni in TM sono inerentemente bidirezionali: $A \sim B$ $B \sim A$

Ciascun membro `<member>` può essere caratterizzato mediante un attributo che ne determini il "ruolo" nella relazione, mediante il tag `<roleSpec>`.

Aumentiamo il livello di dettaglio dell'esempio precedente:

```
<association>
  <instanceOf> <topicRef xlink:href="#creatore"/> </instanceOf>
  <member>
    <roleSpec> <topicRef xlink:href="#autore"/> </roleSpec>
    <topicRef xlink:href="#mario_rossi"/>
  </member>
  <member>
    <roleSpec> <topicRef xlink:href="#pagina_web"/> </roleSpec>
    <resourceRef xlink:href="http://www.mariorossi.com/index.html"/>
  </member>
</association>
```



Associazioni (3)

Un unico costrutto member può contenere più "players", ossia più riferimenti a risorse o ad altri topic che stanno nella stessa relazione e nello stesso ruolo

Esempio: Se Mario Rossi è il creatore di altre pagine web:

```
<association>
  <instanceOf> <topicRef xlink:href="#creatore"/> </instanceOf>
  <member>
    <roleSpec> <topicRef xlink:href="#autore"/> </roleSpec>
    <topicRef xlink:href="#mario_rossi"/>
  </member>
  <member>
    <roleSpec> <topicRef xlink:href="#pagina_web"/> </roleSpec>
    <resourceRef xlink:href="http://www.mariorossi.com/index.html"/>
    <resourceRef xlink:href="http://www.myhsot.org/mrossi.html"/>
  </member>
</association>
```



Published Subject Indicators

Nella definizione di ontologie, per evitare problemi di ambiguità e di sinonimia dei concetti descritti è utile ricorrere, quando possibile, a soggetti definiti pubblicamente.

I Published Subject Indicators (PSI) sono topic che rappresentano concetti basilari, utili nella caratterizzazione di argomenti.

Sono definiti su *namespace* noti e permettono di:

- ◆ facilitare l'interscambio di mappe e la loro fusione;
- ◆ fornire un meccanismo di base per la differenziazione tra parte dichiarativa e parte descrittiva delle mappe.

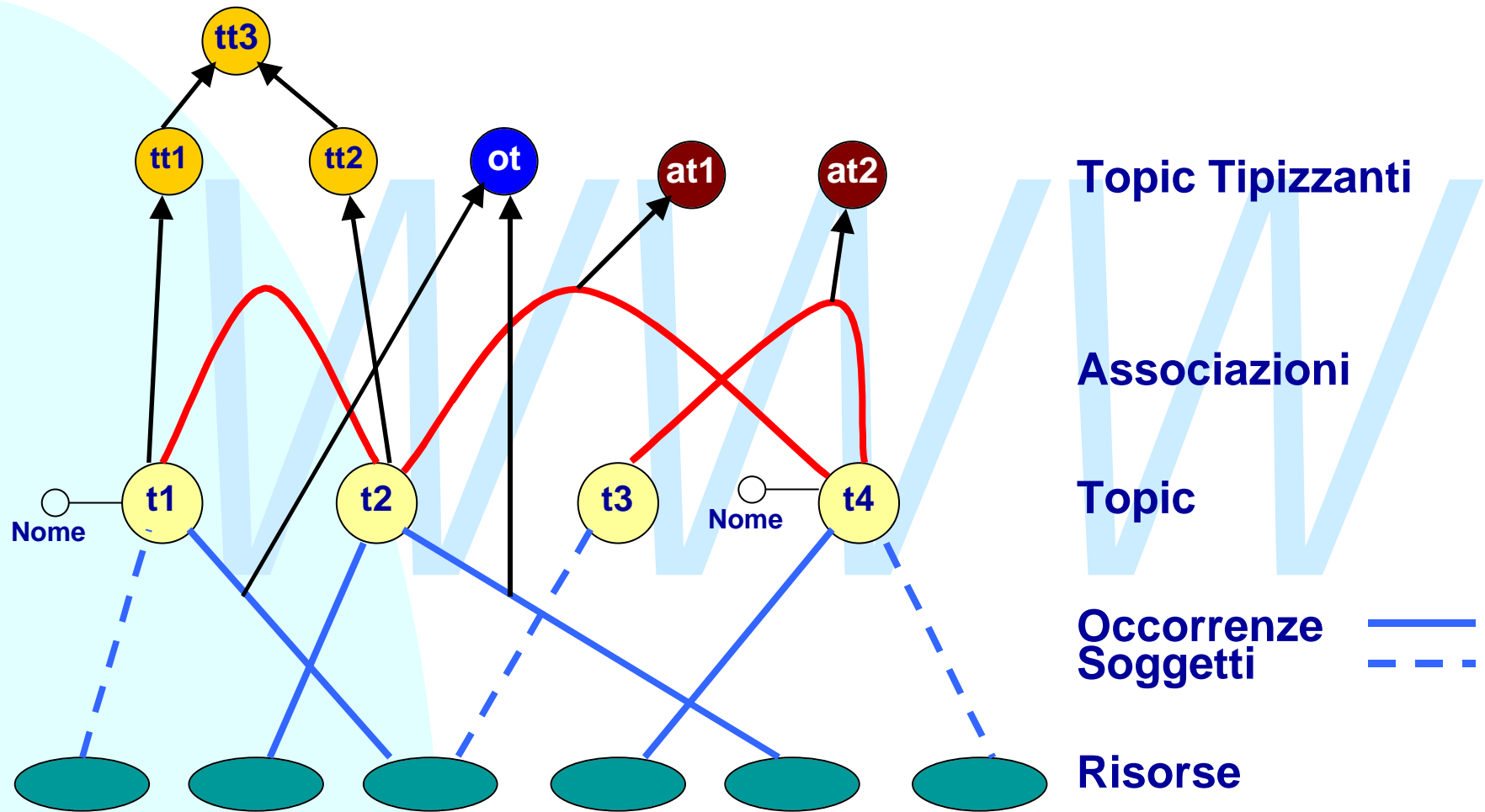
Ne sono un esempio quelli definiti all'url:

<http://www.topicmaps.org/xtm/1.0/core.xtm>

Sono essenzialmente dei topic con ruolo di tipizzazione.



Gerarchia dei concetti di TM



Scope (1)

Il modello Topic Maps prevede la possibilità di contestualizzare le associazioni, le occorrenze ed i nomi assegnabili ai topic col meccanismo degli scope.

Gli scope descrivono un contesto nel quale le caratteristiche descritte hanno significato: un'applicazione di navigazione potrà considerare tale informazione per filtrare le meta-informazioni rilevanti in ogni momento.

L'assenza di uno scope indica validità generale.

Ad esempio, nel seguente topic il basename "Rossi, Mario" deve essere usato per l'applicazione definita con topic "sorting", ovvero per l'ordinamento dei nomi.

```
<topic id="mario-rossi">  
  <baseName>  
    <scope>  
      <topicRef xlink:href="#sorting"/>  
    </scope>  
    <baseNameString>Rossi, Mario</baseNameString>  
  </baseName>  
</topic>
```



Scope (2)

Un topic usato per definire uno scope (nel caso precedente "sorting") viene detto tema (theme). In generale uno scope può essere costituito da un numero arbitrario di temi.

Esempio:

```
<association>
  <instanceOf> <topicRef xlink:href="#creatore"/> </instanceOf>
  <scope>
    <topicRef xlink:href="#web_authoring"/>
    <topicRef xlink:href="#completed_works"/>
  </scope>
  <member> <topicRef xlink:href="#mario_rossi"/> </member>
  <member>
    <resourceRef xink:href="http://www.mariorossi.com/index.html"/>
  </member>
</association>
```



Differenze con RDF

Topic Maps è semanticamente più ricco, tanto che un vero e proprio confronto non può essere fatto senza considerare i livelli superiori di RDF (RDF Schema, modelli di II ordine come DAML+OIL, etc.).

Similarità:

- ◆ Permettono di creare strutture di metadati complesse
- ◆ Permettono la definizione di ontologie
- ◆ Possono essere utilizzati come "building block" per il Semantic Web
- ◆ Al momento, entrambi poco maturi (nonostante ciò che dicono i rispettivi sostenitori)
- ◆ Si prestano entrambi ad una rappresentazione a grafo.

Differenze:

- ◆ Topic Maps consente delle relazioni n-arie multidirezionali invece di richiedere la creazione di $n(n-1)$ predicati monodirezionali di RDF.
- ◆ TM fornisce regole formali per il merging di mappe, ovvero la creazione di mappe unione di due o più mappe più semplici.
- ◆ Non esiste ancora uno standard per definire schemi di mappe (TMCL)



Riferimenti

- Steve Pepper, Graham Moore, *XML Topic Maps (XTM) 1.0*, <http://www.topicmaps.org/xtm/1.0/xtm1-20010806.html>, International Organization for Standardization, ISO/IEC13250:2000 Document description and processing languages. In *Topic Maps*, Ginevra, CH, 2000.
- S. Pepper. The TAO of Topic Maps, finding the way in the age of infoglut. *XML Europe 2000*, 2000.
- Kal Ahmed, Danny Ayers, Mark Birbeckand, Jay Cousinsand, David Doddsand, Joshua Lubelland, Miloslav Nicand, Daniel Rivers-Mooreand, Andrew Wattand, Rob Worden, and Ann Wrightson. *Professional XML Meta Data*. Wrox, 2001.
- Anitta Altenburger. Topic Maps Tutorial, 2001 <http://topicmaps.bond.edu.au/tutorial1>.

