

Mapy témat (Topic maps)

Petr Křemen

Katedra kybernetiky, FEL ČVUT

Co nás čeká

1 Thesaury

2 Mapy témat

Úvod

- dnes – propracovanější sémantické sítě :

Úvod

- dnes – propracovanější sémantické sítě :
 - **thesaurý**

Úvod

- dnes – propracovanější sémantické sítě :
 - **thesaurý**
 - **mapy témat**

Úvod

- dnes – propracovanější sémantické sítě :
 - **thesaurý**
 - **mapy témat**
- motivací pro obě technologie je snaha o urychlení vyhledávání nad dokumenty, neboli o inteligentní **indexaci**.

Úvod

- dnes – propracovanější sémantické sítě :
 - **thesaurý**
 - **mapy témat**
- motivací pro obě technologie je snaha o urychlení vyhledávání nad dokumenty, neboli o inteligentní **indexaci**.
- <http://www.willpower.demon.co.uk/thesprin.htm>

Thesaurus

taxonomie je hierarchie pojmů

thesaurus je taxonomie obohacená o další typy relací. Může být buď jednojazykový (ISO 2788:1986) nebo vícejazykový (ISO 5964:1985) :

BT/NT (broader/narrower term) = hierarchie pojmů.

Příklad

skopové maso → BT → maso

SN (scope note) vysvětluje význam daného termu.

Příklad

škola → SN → instituce sloužící k vzdělávání

USE/UF slouží k popisu deprecated synonym

Příklad

vteřina → USE → sekunda.

Thesaurus – pro a proti

- 😊 čitelnější než striktně formální jazyky pro sémantický web – snadněji tvořitelné.

Thesaurus – pro a proti

- 😊 čitelnější než striktně formální jazyky pro sémantický web – snadněji tvořitelné.
- 😞 problematická sémantika – špatně definované inferenční procedury.

Thesaurus – pro a proti

- 😊 čitelnější než striktně formální jazyky pro sémantický web – snadněji tvořitelné.
- ☹️ problematická sémantika – špatně definované inferenční procedury.

Příklad

Relace BT může být použita ve významech:

Thesaurus – pro a proti

- 😊 čitelnější než striktně formální jazyky pro sémantický web – snadněji tvořitelné.
- 😞 problematická sémantika – špatně definované inferenční procedury.

Příklad

Relace BT může být použita ve významech:

subsumpce , např. jablko BT ovoce,

Thesaurus – pro a proti

- 😊 čitelnější než striktně formální jazyky pro sémantický web – snadněji tvořitelné.
- ☹️ problematická sémantika – špatně definované inferenční procedury.

Příklad

Relace BT může být použita ve významech:

subsumpce , např. jablko BT ovoce,

instance , např. David BT člověk,

Thesaurus – pro a proti

- 😊 čitelnější než striktně formální jazyky pro sémantický web – snadněji tvořitelné.
- ☹️ problematická sémantika – špatně definované inferenční procedury.

Příklad

Relace BT může být použita ve významech:

- subsumpce** , např. jablko BT ovoce,
- instance** , např. David BT člověk,
- být částí** , např. kapota BT auto.

Thesaurus – pro a proti

- 😊 čitelnější než striktně formální jazyky pro sémantický web – snadněji tvořitelné.
- ☹️ problematická sémantika – špatně definované inferenční procedury.

Příklad

Relace BT může být použita ve významech:

subsumpce , např. jablko BT ovoce,

instance , např. David BT člověk,

být částí , např. kapota BT auto.

...

Mapy témat – témata

- ISO standard – ISO/IEC 13250:2003

Mapy témat – témata

- ISO standard – ISO/IEC 13250:2003
- jsou vlastně trochu strukturovanějšími *sémantickými sítěmi*

Mapy témat – témata

- ISO standard – ISO/IEC 13250:2003
- jsou vlastně trochu strukturovanějšími *sémantickými sítěmi*
- tři typy objektů : **témata** (topics), **výskyty** (occurrences) a **asociace** (associations).

Mapy témat – témata

- ISO standard – ISO/IEC 13250:2003
- jsou vlastně trochu strukturovanějšími *sémantickými sítěmi*
- tři typy objektů : **témata** (topics), **výskyty** (occurrences) a **asociace** (associations).
- témata

Mapy témat – témata

- ISO standard – ISO/IEC 13250:2003
- jsou vlastně trochu strukturovanějšími *sémantickými sítěmi*
- tři typy objektů : **témata** (topics), **výskyty** (occurences) a **asociace** (associations).
- témata
 - představují konceptuální pojmy - třídy, instance, vlastnosti, atd.

Mapy témat – témata

- ISO standard – ISO/IEC 13250:2003
- jsou vlastně trochu strukturovanějšími *sémantickými sítěmi*
- tři typy objektů : **témata** (topics), **výskyty** (occurences) a **asociace** (associations).
- témata
 - představují konceptuální pojmy - třídy, instance, vlastnosti, atd.
 - téma může mít několik tzv. **typů témat** (topic types). Vztah “mítí typ” vytváří hierarchii témat (analogie *isa* relace v *sémantických sítích*, nebo vlastnosti *rdf:type* v RDF(S), viz. příští přednáška).

Mapy témat – témata

- ISO standard – ISO/IEC 13250:2003
- jsou vlastně trochu strukturovanějšími *sémantickými sítěmi*
- tři typy objektů : **témata** (topics), **výskyty** (occurrences) a **asociace** (associations).
- témata
 - představují konceptuální pojmy - třídy, instance, vlastnosti, atd.
 - téma může mít několik tzv. **typů témat** (topic types). Vztah “mít typ” vytváří hierarchii témat (analogie *isa* relace v sémantických sítích, nebo vlastnosti *rdf:type* v RDF(S), viz. příští přednáška).
 - tématu může být přiřazeno několik **jmen** (např. přezdívka, formální jméno, přihlašovací jméno, atd.), každé z nich v několika **variantách** (např. zobrazení vs. třídění).

Mapy témat – výskyty

- výskyty

Mapy témat – výskyty

- výskyty
 - představují “odkazy” témat na reálné dokumenty/informační zdroje.

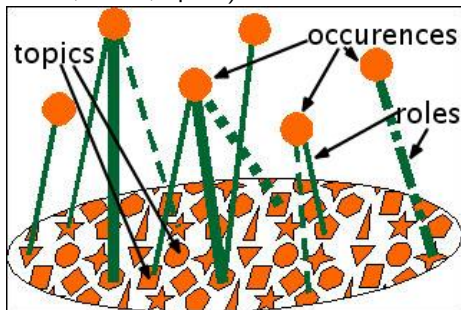
Mapy témat – výskyty

- výskyty
 - představují “odkazy” témat na reálné dokumenty/informační zdroje.
 - téma je s výskytem svázáno pomocí tzv. **role**, která určuje typ výskytu (web. stránka, článek, kniha, apod.).

Mapy témat – výskyty

- výskyty

- představují “odkazy” témat na reálné dokumenty/informační zdroje.
- téma je s výskytem svázáno pomocí tzv. **role**, která určuje typ výskytu (web. stránka, článek, kniha, apod.).



(<http://www.ontopia.net/topicmaps/materials/tao.html>)

Mapy témat – asociace

- asociace

Mapy témat – asociace

- asociace
 - představují vztahy mezi tématy – analogie n-árních relací

Mapy témat – asociace

- asociace
 - představují vztahy mezi tématy – analogie n-árních relací
 - asociaci je přiřazen **asociační typ**, který je tématem („a dále typ tématu je speciálním asociačním typem“).

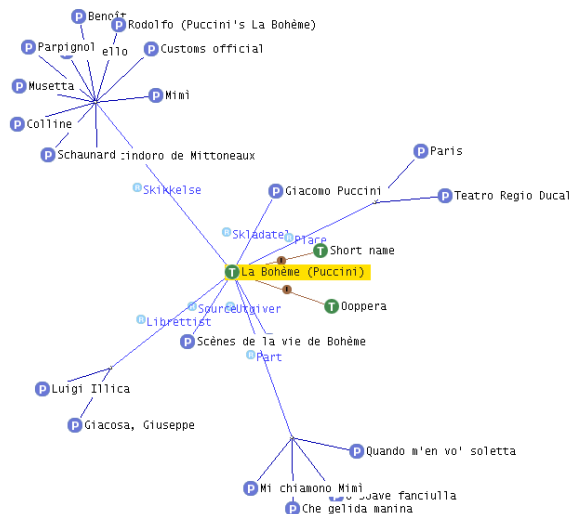
Mapy témat – asociace

- asociace
 - představují vztahy mezi tématy – analogie n-árních relací
 - asociaci je přiřazen **asociační typ**, který je tématem („a dále typ tématu je speciálním asociačním typem“).
 - témata vystupují v asociacích v tzv. **asociačních rolích**

Mapy témat – asociace

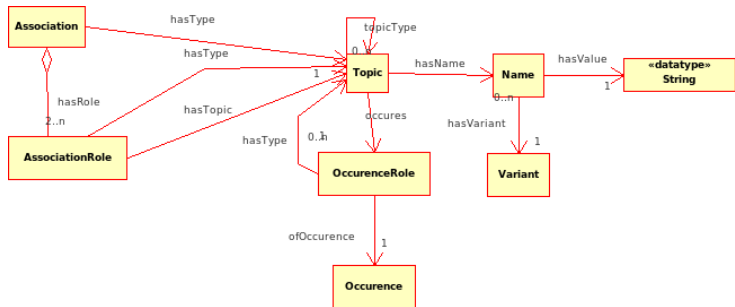
- asociace
 - představují vztahy mezi tématy – analogie n-árních relací
 - asociaci je přiřazen **asociační typ**, který je tématem (,a dále typ tématu je speciálním asociačním typem).
 - témata vystupují v asociacích v tzv. **asociačních rolích**
 - též asociační roli je přiřazen **typ asociační role**, který je tématem

Mapy témat – příklad



- T ... témata
- P ... částečně rozvinutá témata (kromě typů témat)
- R ... asociace

Mapy témat – model



Nástroje a odkazy

- nástroje:

Nástroje a odkazy

- nástroje:
 - Ontopia (Ontopoly, Omnigator, Vizigator) – hlavní tahoun vývoje v TM

Nástroje a odkazy

- nástroje:
 - Ontopia (Ontopoly, Omnigator, Vizigator) – hlavní tahoun vývoje v TM
 - TM4L

Nástroje a odkazy

- nástroje:
 - Ontopia (Ontopoly, Omnigator, Vizigator) – hlavní tahoun vývoje v TM
 - TM4L
 - TM4J

Nástroje a odkazy

- nástroje:
 - Ontopia (Ontopoly, Omnigator, Vizigator) – hlavní tahoun vývoje v TM
 - TM4L
 - TM4J
 - ... a mnoho dalších

Nástroje a odkazy

- nástroje:
 - Ontopia (Ontopoly, Omnigator, Vizigator) – hlavní tahoun vývoje v TM
 - TM4L
 - TM4J
 - ... a mnoho dalších
- odkazy:

Nástroje a odkazy

- nástroje:
 - Ontopia (Ontopoly, Omnigator, Vizigator) – hlavní tahoun vývoje v TM
 - TM4L
 - TM4J
 - ... a mnoho dalších
- odkazy:
 - <http://www.ontopia.net/topicmaps/materials/tm-vs-thesauri.html>

Nástroje a odkazy

- nástroje:
 - Ontopia (Ontopoly, Omnigator, Vizigator) – hlavní tahoun vývoje v TM
 - TM4L
 - TM4J
 - ... a mnoho dalších
- odkazy:
 - <http://www.ontopia.net/topicmaps/materials/tm-vs-thesauri.html>
 - <http://www.kosek.cz/xml/tmtut/>

TM4L Viewer

The screenshot displays the TM4L Viewer interface with three main panels:

- File Scope Help:** Shows the current file path: `M:file:/home/kremen/downloads/zsyopera.xtm`.
- Topic Maps Indexes:** A list of subject topics including 'Voto', 'Work by Paul de Choudens', 'A Masked Ball', 'A peasant', 'Abbess', 'Abbé', 'Abbé de Chazeuil', 'Abdual', 'Abhijnanasakuntala', 'Abigail', 'Abreschwiler', 'Acciano', 'Adam Mickiewicz', 'Adami', 'Adhemar de Monthell', 'Admèto', 'Adonella', 'Adriana Lecouvreur', 'Adrienne Lecouvreur', 'Agean island', 'Aira', 'Agnese', 'Ahi Per l'ultima volta', 'Ah, il suo nome', 'Ah-je', 'Ahi, Vergine Maria', 'Aida', 'Aida (character)', 'Alcindo de Mittonaux', 'Aida di Fatenow', 'Aldobrandino del Rangoni', 'Aldwort', 'Alexis', 'Alefano', 'Afu', 'Alfredo Germont', 'Alice', 'Alice Ford', 'Aligi', 'Aims Collector', 'Absace', 'Alternativnav', 'Atchiana', 'Avaro', 'Alvise Badero', 'Atzira', 'Atzira (character)', 'Atzira, ou Les Américains', and 'Amalia'.
- Tree View:** A hierarchical tree structure for 'La Bohème (Puccini)'. The root node is 'La Bohème (Puccini)', which branches into 'Dopperra', 'Short name', 'Web page', 'Skikkelse', 'Alcindo de Mittonaux', 'Parisi', 'Colline', 'Customs official', 'Marcello', 'Mimi', 'Musetta', 'Pargnion', 'Rodolfo (Puccini's La Bohèr)', and 'Schaunard'. Further sub-nodes include 'Faster', 'Source', 'Librettist', 'Place', 'Video recording', 'Sound clip', 'Article', 'Libretto', 'Illustration', 'Synopsis', 'Utgiver', 'Première date', and 'Part'.
- Graph/Text View:** A network graph showing the relationships between the nodes in the tree. Nodes are represented by colored circles with text labels. The central node is 'La Bohème (Puccini)'. It is connected to 'Paris', 'Giacosa, Giuseppe', 'Luigi Illica', 'Boppera', 'Giacoso Puccini', 'Part', 'Scènes de la vie de Bohèr', 'Ricordi', 'Librettist', 'Short name', 'Skikkelse', 'Pargnion', 'Customs official', 'Marcello', 'Rodolfo (Puccini's La Bohème)', 'Mimi', 'Musetta', 'Schaunard', and 'Benoit'. The graph also shows connections between 'Paris' and 'Teatro Regio Duca', 'Giacosa, Giuseppe' and 'Luigi Illica', 'Boppera' and 'Giacoso Puccini', 'Giacoso Puccini' and 'Part', 'Part' and 'Scènes de la vie de Bohèr', 'Scènes de la vie de Bohèr' and 'Ricordi', 'Ricordi' and 'Librettist', 'Librettist' and 'Short name', 'Short name' and 'Skikkelse', 'Skikkelse' and 'Pargnion', 'Pargnion' and 'Customs official', 'Customs official' and 'Marcello', 'Marcello' and 'Rodolfo (Puccini's La Bohème)', 'Rodolfo (Puccini's La Bohème)' and 'Mimi', 'Mimi' and 'Musetta', 'Musetta' and 'Schaunard', and 'Schaunard' and 'Benoit'.
- Log:** A message at the bottom: 'Cannot activate TopicMapReference, because its in state StoreState (opening)'.

Browser window showing the Omnigator interface for the topic "La Bohème (Puccini)".

URL: http://localhost:8080/omnigator/model/topic_complete.jsp?tm=zzyopera&tmId=la-boheme

Navigation: Home | Manage | WikiData | Support | About

Search: Open... Refresh Hot Indexed

zyOpera | Customise | Filter | Export | Merge | Statistics | Query | Edit | No schema | Vajpat

La Bohème (Puccini)

Type(s): Opera

Untyped Names (1)

- La Bohème (Puccini)
 - Bohème (Puccini) - Scope: Sort

Subject Identifiers (1)

- http://en.wikipedia.org/wiki/La_Bohème

Short name (1)

- La Bohème - Scope: Puccini, Giacomo
 - Bohème - Scope: Puccini, Giacomo; Sort

Internal Occurrences (2)

- Première date
 - 1896-02-01
- Video recording
 - 100 046

Associations (21)

- Based on
 - Scènes de la vie de Bohème
- Composed by
 - Puccini, Giacomo
- Contains
 - Che gelida manina
 - Mi chiamano Mimì
 - Osava fanciulla
 - Quando m'en voi' solletta
- Dramatic personas
 - Alicandro de Metzouauk
 - Benoit
 - Colline
 - Customs official
 - Harcelle
 - Mimi
 - Musetta
 - Pargipol
 - Rodolfo (Puccini's La Bohème)
 - Schaunard
- First performed at
 - Teatro Regio Ducal
- Libretto by
 - Giacosa, Giuseppe
 - Illica, Luigi
- Published by
 - Ricordi
- Takes place in
 - Paris

External Occurrences (11)

- Article
 - http://en.wikipedia.org/wiki/La_Bohème - Scope: Web; Wikipedia
- Illustration
 - <http://localhost:8080/operamap/occurs/opera/puccini/la-boheme-poster-1.jpg> - Scope: Local
- Libretto
 - <http://opera.stanford.edu/opera/Puccini/LaBoheme/libretto.html> - Scope: Opera Glass; Web
- Poster
 - <http://www.ds.com/opera/puccinian/pictures/boheme-score.htm> - Scope: OperaResource; Web
 - http://www.ds.com/opera/puccinian/pictures/bohema_poster.htm - Scope: OperaResource; Web
- Sound clip
 - <http://localhost:8080/operamap/occur/la-boheme.wav> - Scope: Local
 - <http://www.ontopia.net/topicmaps/examples/operamap/occurs/la-boheme.wav> - Scope: Web
- Synopsis
 - http://localhost:8080/operamap/occurs/la_syn.php3.htm - Scope: Arizona Opera; Local
 - <http://www.azopera.com/team/synopsis/boheme.shtml> - Scope: Arizona Opera; Web
 - <http://www.metopera.org/synopses/boheme.html> - Scope: Opera News; Web
- Web page
 - <http://opera.stanford.edu/opera/Puccini/LaBoheme/main.html> - Scope: Opera Glass; Web

Scoped Names (1)

- Rodolfo (Puccini's La Bohème) (Rodolfo (Puccini's La Bohème))

Scoped Association Types (2)

Další aspekty map témat

- kromě přehledu uvedeného výše je možné sdružovat mapy témat (spec. typ tématu) do **kontextů** (scopes, themes). Kontexty představují filtrační mechanismus

Příklad

Modelujeme-li doménu *turismus*, jiná data by měla být dostupná zájemci o cestování a jiná pracovníkovi cestovní kanceláře.

- XTM je XML formát pro ukládání map témat.
- dotazování pomocí TMQL, nebo tologu (podobná syntax jako SQL).