Evaluation RDF Assignment

• Goal: Get some experience in working with RDF (related technologies)

• Some Observations:
  • Deadline problems: 58% of the submissions in the last hour before the deadline
  • Some people reported tool problems
  • Most people performed quite well on the assignments

<table>
<thead>
<tr>
<th>10</th>
<th>9</th>
<th>8</th>
<th>7</th>
<th>6</th>
<th>5</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>6</td>
<td>6</td>
<td>10</td>
<td>12</td>
<td>7</td>
<td>4</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
</tbody>
</table>
RDF Assignment – Common Problems
Proper Subclassing

Versus

Vehicle

Car

SportsCar

SUV

Car

Engine

Crank

Valve

Exhaust

subclass subclass

subclass subclass

subclass subclass

subclass subclass

/ name of department
Recap: Why RDF?

- Consider a typical web page
- Markup consists of:
  - Rendering information (e.g., font size and color)
  - Hyper-links to related content
- Semantic content is accessible to humans but not (easily) to computers…
What should the schema capture?

Technical Information
- Dimension
  - Width: 665px
  - Height: 850px
- Type: JPEG
- FileSize: 148 KB

Content Information
- Author
  - First Name: Maurits Cornelis
  - Last Name: Escher
- CreationDate: 1961
- Title: Waterfall
- Method: Litography
- Subject
  - Type: Optical Illusion
  - Type: Abstract
  - Depicts: Waterfall
  - Depicts: Watermill

Personal Information
- User
  - Name: Kees
  - Rating: 9.1
  - Comment: “Great Optical Illusion!”

Personal Information
- User
  - Name: Bart
  - Rating: 7.3
  - Comment: “Not his Best Work”
What does RDF give us?

A mechanism for annotating data and resources

• Single (simple) data model

• Syntactic consistency between names (URIs).

• Data model for the Web
  • Openness & Flexibility (use arbitrary properties)
  • Resource-centered (triple-based)
  • Datamodel easy to understand and manipulate

• RDF graphs can be simply merged (RDF merge is a monotonic operation!)
Subjects

- Examples of Semantic Web applications
  - CHIP
  - iFanzy
  - RHCe
  - Semantic Web Challenge
- Linked Data
  - Legacy Applications
  - Evolution
- RDFa
- Current Research Topics
CHIP – Rijksmuseum Amsterdam
providing semantic browsing, searching and semantic recommendations

Inside museum

Online
http://www.rijksmuseum.nl

7000 artworks

50000 artworks
Personalized Art Experience

Personalized Web Site

• Interactive user modeling
• Personalization of navigation & presentation

Personalized Museum Tour

• Printed tour maps through the Rijksmuseum

Personalized Tour on a Mobile Device

• Interactive user modeling
• Route adaptation inside the Rijksmuseum
Approach

• Making museum metadata available in RDF/OWL

• Making relevant vocabularies available in RDF/OWL

• Aligning & enriching vocabularies/metadata

• Using resulting RDF/OWL graph for building a combined (virtual and physical) user model

• Using the above results for (semi)automatic generation of virtual and physical museum tours
Artist: Rembrandt van Rijn

Roles:
- artist (preferred)
- draftsman
- printmaker
- etcher
- painter
- teacher

Gender: male

Birth and Death Places:
- Born: Leyden (South Holland, Netherlands) (inhabited place)
- Died: Amsterdam (North Holland, Netherlands) (inhabited place)

Events:
- active: 1631-1669 Amsterdam (North Holland, Netherlands) (inhabited place)

Related People or Corporate Bodies:
- teacher of .... Bol, Ferdinand
  (Dutch history and portrait painter, 1616-1680) [500013764]
- teacher of .... Dou, Gerrit
  (Dutch painter and draftsman, 1613-1675) [500115513]
- teacher of .... Drost, Willem
  (Dutch painter, printmaker, and draftsman, ca. 1630-after 1680) [500009474]
- teacher of .... Dullaert, Heyman
  (Dutch painter and poet, 1636-1684) [500021474]
- teacher of .... Enckhout, Gerbrand van den
Style: Baroque

Note: Refers to the style and period of architecture, visual art, decorative art, music, and literature of western Europe and the Americas about 1590 to 1750. The style is characterized by balance and wholeness, often with an emphasis on spectacle and emotion, as well as tendencies toward contrasts of light against dark, mass against void, and the use of strong diagonals and curves.

Terms:
Baroque (preferred, C.U.D.American English-P)

Facet/Hierarchy Code: F.F.L

Hierarchical Position:
- Styles and Periods Facet
  - Styles and Periods
    - <styles and periods by region>
      - European
        - <European styles and periods>
          - <Renaissance-Baroque styles and periods>
            - Baroque

Related concepts:
- Rococo related to
  - <Renaissance-Baroque styles and periods>, <European styles and periods>, ... Styles and Periods) [300107026]
Click the icon to view the hierarchy.

ID: 7006952

Coordinates:
Lat: 52.21 00 N  degrees minutes   Lat: 52.3500  decimal degrees
Long: 004 54 00 E  degrees minutes   Long: 4.9000  decimal degrees

Note: Located on over 90 islands in the IJ arm of the IJsselmeer. Early inhabitants built dikes on both sides Amstel River to prevent flooding, and a dam was built between the dikes in 1270. Chartered in 1306. Becar affluent in the 15th century due to trade with Baltic seaports, and was the financial center of the world by 1 century. United Dutch East India Company was founded in 1602, followed by the West India Company in 1st Capital of the Batavian Republic under Napoleon, later of the kingdom of Holland, and became part of the Empire in 1810. Under German occupation from 1940-1945. Center of the world’s diamond trade.

Names:
- Amsterdam (preferred, C,V,N,English-P,Dutch-P)
- Amstel-dam (H,V,N)  
  ............ documented in 13th cen., meaning "dam on the Amstel [river]"
- Amsteldam (H,V,N)
- Amstelledamme (H,V,N)
- Amstelodamum (H,O,N)
- Amsteadamum (H,O,N)
- Amsterdamum (H,O,N)
- Amstredamum (H,O,N)
- Amstredamense oppidum (H,O,N)

Place Types:
- inhabited place (preferred, C)  
  ............ there possibly was a Roman settlement in the area; modern town probably originated as a fishing village in 13th century
- city (C)
- capital (C)  
  ............ nominal capital of The Netherlands, though government is located in 's-Gravenhage
Semantic Recommendation

The Night Watch

creationSite

North Holland

part of

Amsterdam

creationDate

1642

material

Oil paint

Creator

Rembrandt

style

Baroque

subject

Militia

specific term

subject

Portraiture

studentOf

Pieter Lastman

teacherOf

Ferdinand Bol
Help Needed!

- To collect users' feedback on the effectiveness of the recommendation strategy, we invite you to participate in our online user study

- [http://www.chip-project.org/demoUserStudy3](http://www.chip-project.org/demoUserStudy3)
iFanzy: Personalized TV-guide

- http://www.nu.nl/tvgids/
  - Focus on the EPG view
  - Long list of channels
  - No personalization / adaptation / search function

- http://www.tvgids.nl/
  - Focus on program search
  - Limited personalization / adaptation
  - Hard to scale interface
Goal

- Personalized Web-based browser for digital television content
  - Harvest program information from different sources
  - Use the application on several type of machines and platforms
  - Give the user control over a large source of data
Integration!

- Integration of program data:
  - Crawling the Web to bring different of program information together
  - Searching for background knowledge to uncover extra connections between items
- Integration of user data
  - Looking at similar users to form different groups
  - Looking for different existing user profiles and integrate them to get a richer user model
- Integration of platform interaction
  - Different platforms show different user behavior
Content Integration

- RDFized IMDB dataset (multi-million triples, 13GB of data)
- Retrieved URLs photos and trailers from the Web
- Connected to Time, GEO and TVA-genre ontologies for reasoning purposes
- Relate IMDB with EPG=data by movie title in combination with director and actor names
Converting TV Metadata in RDF/OWL

Input source 1:

```xml
<program channel="NED1">
  <source>http://foo.bar/</source>
  <title>Sportjournaal</title>
  <start>20080309184500</start>
  <end>20080309190000</end>
  <genre>sport nieuws</genre>
</program>
```

Input source 2:

```xml
<program title="Match of the Day">
  <channel>BBC One</channel>
  <start>2008-03-09T19:45:00Z</start>
  <duration>PT01H15M00S</duration>
  <genre>sport</genre>
</program>
```

Translation to TV-Anytime in RDF/OWL

```xml
<TVA:ProgramInformation ID="crid://foo.bar/0001">
  <hasTitle>Sportjournaal</hasTitle>
  <hasGenre rdf:resource="TVAGenres:3.1.1.9"/>
</TVA:ProgramInformation>

<TVA:Schedule ID="TVA:Schedule_0001">
  <serviceIDRef>NED1</serviceIDRef>
  <hasProgram crid="crid://foo.bar/0001"/>
  <startTime rdf:resource="TIME:TimeDesc_0001"/>
</TVA:Schedule>

<TIME:TimeDescription ID= "TIME:TimeDesc_0001">
  <year>2008</year>
  <month>3</month>
  <day>9</day>
  <hour>18</hour>
  <minute>45</minute>
  <second>0</second>
</TIME:TimeDescription>
```
Converting Vocabularies in RDF/OWL

Translation of TV-Anytime genres to RDF/OWL using SKOS

```
<TVAGenres:genre ID="TVAGenres:3.1.1.9">
  <rdfs:label>Sport News</rdfs:label>
  <skos:broader rdf:resource="TVAGenres:3.1.1"/>
  <skos:related rdf:resource="TVAGenres:3.2"/>
</TVAGenres:genre>

<TVAGenres:genre ID="TVAGenres:3.2">
  <rdfs:label>SPORTS</rdfs:label>
  <Term termID="3.2.1">
    <Name xml:lang="en">Athletics</Name>
  </Term>
</TVAGenres:genre>

<TVAGenres:genre ID="TVAGenres:3.1.1">
  <rdfs:label>News</rdfs:label>
  <skos:narrower rdf:resource="TVAGenres:3.1.1.9"/>
  <skos:broader rdf:resource="TVAGenres:3.1"/>
</TVAGenres:genre>
```
Aligning and Enriching Vocabularies

• **Alignment of Genre vocabularies**
  - XMLTV:documentaire $\rightarrow$ TVA:”Documentary”
  - IMDB:Thriller $\rightarrow$ TVA:”Thriller”
  - IMDB:Sci-Fi $\rightarrow$ TVA:”Science Fiction”

• **Semantic enrichment of Genre vocabulary**
  - News -skos:narrower-> Sports News $\Rightarrow$ Original Term hierarchy
  - Sport News -skos:related-> Sport $\Rightarrow$ Partial label matches
  - Skating -skos:related-> ‘Ice skating’ $\Rightarrow$ Partial label matches
  - ‘American Football’ -skos:related-> Rugby $\Rightarrow$ Domain expert

• **Semantic enrichment of TV metadata with IMDB movie descriptions**
  - “Buono, il brutto, il cattivo, Il (1966)” $\Rightarrow$
  - “The Good, the Bad and the Ugly”

• **Alignment of date/time descriptions to Time ontology concepts to allow temporal reasoning**
  - “2006-01-01T12:00:00” $\Rightarrow$
    - <time:year>2006</time:year>
    - <time:day>01</time:day>
    - <time:hour>12</time:hour>
Semantic Graph for Recommendations

- Generating recommendations based on usage data and the RDF/OWL graph, behavior analysis
- Query expansion on search terms and UM values
  - WordNet: synonyms for search terms
  - skos:narrower/related relationships
- When asking for a recommendation, empty search fields like <genres> and <terms> are filled in by user preferences
  - When requested only specific contexts are considered. Context includes:
    - Time contexts e.g. preferences in morning, evening,…
    - Audience e.g. preferences for groups
    - Location e.g. preferences can differ per location
**iFanzy online: www.iFanzy.nl**

<table>
<thead>
<tr>
<th>DAGEN</th>
<th>ZO</th>
<th>MA</th>
<th>DI 1 JUL</th>
<th>WE</th>
<th>DO</th>
<th>VR</th>
</tr>
</thead>
<tbody>
<tr>
<td>UREN</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>18:00</td>
<td>19:00</td>
<td>20:00</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Ned1**
- Journaal
- Een Vandaag
- Sportjour
- Sportjour Lingo
- Hart voor rieren
- Journaal
- Blik op de weg

**Ned2**
- Dr. Quinn, Medicine
- That's the Question
- Van jonge leeu en oude grond
- Kinderen geen bezwaar
- Vroege vogels
- Netwerk
- Stop de...

**Ned3**
- ZipZoom
- Worldwide
- Het Klokhuys
- Jeugdjou
- Falcon Beach
- Journaal Mad Labs
- Journaal Acht

**RTL4**
- Nieuws
- Editie NL
- Wie is de chef
- De bloemenstaal
- Nieuws
- NL vertrekt
- Bouwval gezocht

**RTL5**
- Date My Dr. Phil
- The Nanny
- Just Shoot Me
- Flood
- CSI: Miami

**SBS6**
- J&J E.R.
- Hart van Nederland
- Shownieuws: Vroege Editie
- Trauma Centrum
- House, M.D.

**RTL7**
- rtl7
- Dagtv
- Stecked
- The Osbournes
- Aller Allez Zimbabwe
- The Package

**Net5**
- Tennis: Wimbledon
- Will & Grace
- Judging Amy
- Grey's Anatomy
- Gossip Girl

**Veron**
- The Fresh Prince of Bel Air
- According to Jim
- Friends
- Two Guys, a Girl and a Pizza Place
- Two and a Half Men
- Mr. 3000

**RTL8**
- Joan of Arcadia
- Het roer om
- Dr. Phil
- Bounce
Personal TV-guide
iFanzy STB

- STB interface
- Build to fit on a television screen
- Different layout, same server
- Works with a VOD source
1. Hatema, Helmond
2. Strijp S
3. Nette huishoudster
4. Uw woning, uw kasteel!
5. Adolfsfilm Steersel
6. Sigarettenfabriek NV Crescent Cie
• Regional Historic Center Eindhoven
  • Governs all historic material related to the region of Eindhoven
  • Govern historic material and provide citizens access to the archives (including promoting)
• Heterogeneous datasets
  • Videos, pictures, drawings, postcards, ownership records, birth marriage death records, maps, aerial pictures, meeting minutes, financial records, etc
CHI Browser

• RHCe Portal that provides navigation and personalization over the archives to the user

• Navigation Structure:
  • Objects in general are connected by shared ‘dimensions’: description keywords, time and location
  • Using these facets allow both searching and browsing the collections and connecting similar objects (over these dimensions)
  • Built specialized browsing paradigms for these dimensions
Tagging

What?

• Assigning keywords (or short phrases) to resource
  • Tags can be used like text in textual documents during indexing for retrieval

Why?

• It might be a bit complex for end users to edit an RDF graph
• Tagging is a simpler mechanism that users are already used to
  • No complexity: users can make up their own words
Desire

• Benefit from the advantages of the simple tagging mechanism, while also benefitting from the richer structure of the Semantic Web

Purpose Matching Component

• Relating tags to ontological concepts
String Matching

- Pattern Matching
  - Exact or substring
  - “day” matches “Friday”
- Levenshtein
  - Minimum number of edits to transform one word into another
  - “Hockie” -> “Hocke” -> “Hockey” (distance 2):
- Jaro-Winkler
  - Compare number of similar characters on similar positions
  - Hockie vs Hockey (4 exact and one transposition):
- Soundex
  - Use phonetics to compare sound of words
  - Hockie (H300) vs Hockey (H300)
Semantic Broadening

- Using structure of ontology to expand concepts by following properties
  - E.g. using rdfs:subClassOf or skos:narrower

![Diagram showing relationships between Hockey, URI_2, URI_1, and Sport with labels 'narrower' and 'broader']
Semantic Broadening (2)

• Or more complicated, by using a query

![Diagram showing concepts related to sports with labels: Hockey and Tennis]
• Context often allows to determine right context of ambigues word (especially names)
• Take several input tags
• Use context of those input tags for disambiguation
  • E.g. “Bill” + “President” versus “Bill” + “Microsoft”
• Configure concept-distance