Tutorial: Social Semantics
Dr Matthew Rowe, Knowledge Media Institute

Tutorial: Social Semantics, Wed 23/05/12 14:00-15:00
Outline

- Introduction:
  - Social Semantics, so what?
- Representing Social Data Semantically
  - The Social Semantic Web Stack
  - Ontologies and Vocabularies
- Applications of Social Semantics
  - Collaborative Environments
  - Exporting Social Semantic Data
  - Current Research
- Conclusions

Next session: Hands-on!
  - Using social semantic data to detect Web citations
SOCIAL SEMANTICS, SO WHAT?
The Social Web

- A long time ago (~2001/2002) the Web 1.0 began to change… (to 2.0)
  - B2.0 = The Document Web
    - Geocities_web_page -> Tripod_personal_page
  - A2.0 = The Social Web
    - Alice knows Bob
    - Alice uploads a revealing photo of Bob, Bob gets in trouble with his boss

- Social networking came to the fore… people became connected through objects
  - Photos: Flickr
  - Videos: Youtube
  - Music: MySpace
  - Status: Twitter
  - People: Facebook

- Technologies emerged:
  - Blogs and Microblogs, Wikis, Bookmarks, Geo-social networking
The Social Web, Digital Identity, ‘Celebrity’

“All the world's a stage,  
And all the men and women merely players”

- We can directly and spontaneously shape *who* we are
  - Intentional identity
  - Self-presentation
- We can publish content to *our* audience
  - Attention Culture
  - Self-efficacy and self-affirmation
- Our ‘reach’ is global
  - Videos going viral (Star Wars kid)
  - We are all celebrities
- There are consequences however… (as you will see)
The Social Web

- And today….Masses of social data now being published
  - Facebook likes (2.7 billion likes per day)
  - Foursquare (Average user checks in 563 times per year)
  - Twitter (On a busy day ~175 million tweets are published)
  - Instagram (5 million+ photos uploaded per day)

- However…
  - Social data is contained in silos
    - Can I take it with me?
    - Heterogeneous formats
    - Isolated communities
    - Distributed identities
  - We experience social fatigue
    - Sign up, after sign up
    - Facebook Connect, reliable?
The Semantic Web... a ‘minor’ digression

Goal: enable machine-readability and interpretation of information

Solution: build technologies to:

- A) provide common metadata representations
- B) enable shared conceptualisation of some domain (football = soccer?)
- C) facilitate inferences (let machines find things out for you)

Agreed set of technologies:

- URI = To name things
- RDF = To represent metadata
- RDFS/OWL = To formalise ontologies
- SPARQL = To query data
- SWRL/RIF = To infer things
- Linked Data = To connect things together
Ta Da! The Social Semantic Web

The Social Semantic Web

- Use RDF to represent social data
- Use common, shared ontologies for the semantics of social data
  - People
  - Relationships
  - Social objects (content)

- Overcome the heterogeneity of social data
  - Common understanding that a person in Facebook is a person on Twitter

- Data is no longer hidden away in silos
  - Climb the walls of the walled gardens
  - Link people together across those boundaries

- Data portability is enabled
  - Export your data in a common format
  - Take it with you… liberation!
REPRESENTING SOCIAL DATA SEMANTICALLY
The Social Semantic Web Stack
People / Social Networks: FOAF

- Friend of a Friend Ontology (FOAF)

- It describes people:
  - _:Winston_Smith a foaf:Person

- And their properties (both datatype and object):
  - _:Winston_Smith foaf:name “Winston Smith”
  - _:Winston_Smith foaf:mbox <wsmith@ministy_of_truth.org.uk>

- And who they know:
  - _:WinstonSmith foaf:knows _:TomParsons
  - _:WinstonSmith foaf:knows _:MrsParsons
  - _:TomParsons foaf:knows _:MrsParsons

- Allows people to be linked to one another across social web platforms
  - Distributed social networks!
People / Content / Social Networks: Schema.org

- Collection of tags to define item types (classes) and their attributes
- Covers a range of types:
  - Creative works, Event, Organisation, Person, Place, Produce, Review
- And relations between people:
  - Knows, colleague, children, parent, sibling, relatedTo

### Person

<table>
<thead>
<tr>
<th>Property</th>
<th>Expected Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>description</td>
<td>Text</td>
<td>A short description of the item.</td>
</tr>
<tr>
<td>image</td>
<td>URL</td>
<td>URL of an image of the item.</td>
</tr>
<tr>
<td>name</td>
<td>Text</td>
<td>The name of the item.</td>
</tr>
<tr>
<td>url</td>
<td>URL</td>
<td>URL of the item.</td>
</tr>
<tr>
<td>additionalName</td>
<td>Text</td>
<td>An additional name for a Person, can be used for a middle name.</td>
</tr>
<tr>
<td>address</td>
<td>PostalAddress</td>
<td>Physical address of the item.</td>
</tr>
<tr>
<td>affiliation</td>
<td>Organization</td>
<td>An organization that this person is affiliated with. For example, a school/university, a club, or a team.</td>
</tr>
<tr>
<td>alumniOf</td>
<td>EducationalOrganization</td>
<td>An educational organizations that the person is an alumni of.</td>
</tr>
<tr>
<td>award</td>
<td>Text</td>
<td>An award won by this person or for this creative work. (legacy spelling; see singular form, award)</td>
</tr>
<tr>
<td>awards</td>
<td>Text</td>
<td>Awards won by this person or for this creative work. (legacy spelling; see singular form, award)</td>
</tr>
<tr>
<td>birthDate</td>
<td>Date</td>
<td>Date of birth.</td>
</tr>
<tr>
<td>children</td>
<td>Person</td>
<td>A child of the person.</td>
</tr>
</tbody>
</table>
People / Social Networks: Microformats

- Lightweight formats embedded within XHTML elements
- ‘Lowercase semantics’
- hCard: defines people
  - Semantics contained in the ‘class’ attribute

```html
<span class="vcard">
  <span class="fn">Winston Smith</span>
  <span class="url">http://wsmith.ministry_of_truth.org.uk</span>
  <span class="email">wsmith@ministry_of_truth.org.uk</span>
</span>
```

- XFN (XHTML Friends Network): defines links/relations between people
  - Semantics of the relation is in the ‘rel’ attribute

```html
<a href="http://wsmith.ministry_of_truth.org.uk" rel="me">My homepage</a>
<a href="http://tparsons.ministry_of_truth.org.uk" rel="friend">Tom Parsons</a>
<a href="http://www.charringtons.co.uk" rel="met acquaintance">Mr Charrington</a>
```
Social Networks: SemSNA

- Semantic Social Network Analysis:
  - Understanding the structure of networks and the positions of nodes
  - Enables: assessing information flow, trust assessments, role analysis
  - [http://ns.inria.fr/semsna/2008/11/02/voc.rdf](http://ns.inria.fr/semsna/2008/11/02/voc.rdf)
Content / Online Communities: SIOC

- Semantically Interlinked Online Communities (SIOC) Ontology
  - [http://sioc-project.org/ontology](http://sioc-project.org/ontology)

- Captures the semantics of:
  - Forum, Community, User, Content Item

- Allows multiple accounts from disparate platforms to be reconciled

- Conceptualises the context of users and actions
Combining FOAF + SIOC

FOAF + SIOC + SKOS

Combining FOAF + SIOC

UserAccount
- Post
- ChatChannel
- Post
- UserAccount
- Post
- MailingList
- Post
- UserAccount
- Post
- UserAccount
- Post
- Weblog
- Post
- UserAccount
- Post

foaf: Person

UserAccount
- Post
- MessageBoard
- Post

foaf: account

UserAccount
- Post
- Weblog
- Post

foaf: account

UserAccount
- Post

foaf: account

UserAccount
- Post

foaf: account

UserAccount
- Post

skos:Concept

skos:narrower
Online Presence Ontology (OPO)
- [http://online-presence.net/ontology.php](http://online-presence.net/ontology.php)
- Models the online presence of users
- Enables cross platform representation, i.e. status updates
Folksonomies: MOAT

- Folksonomy = tuple (user, tag, resource)
- Problem: tag ambiguity
  - “Watched the Broncos win in last night’s Football #mondaynightfootball”
- Solution: Meaning of a Tag (MOAT)
  - Align tags to concepts
  - [http://moat-project.org/](http://moat-project.org/)

Notable Mention:
- Abundance of Entity Recognisers available:
  - [http://tomazkovacic.com/blog/](http://tomazkovacic.com/blog/)
  - [http://www.semantic-web-journal.net/content/aligning-tweets-events-automation-semantics](http://www.semantic-web-journal.net/content/aligning-tweets-events-automation-semantics)
Behaviour: OUBO

- People, social networks, communities… behaviour
  - User behaviour is contextual: different places, different times
  - Behaviour is adaptive: we change to suit our environment
- How to capture this…. Open University Behaviour Ontology (OUBO)
  - http://purl.org/net/oubo/0.3
  - Represents the dimensions of behaviour and captures contextual aspects
  - Enables Role inference via SPARQL rules (as you will see)
APPLICATIONS OF SOCIAL SEMANTICS
Collaborative Environments: Semantic Wikis

- Wikis contain knowledge that is human decipherable
  - Nineteen Eighty-Four (first published in 1949) by George Orwell is a dystopian novel about Oceania

- Semantic wikis enable machine-readability of content via annotations
  - Annotation syntax that allows properties to be embedded
    - Nineteen Eighty-Four (first published in 1949) by [[Has Author : George_Orwell]]

- Semantic MediaWiki (Kötzsch, Vrandečić and Völkel)
  - Semantic forms for publishing semantic data

- Ontowiki (Auer, Dietzold and Riechart)
  - Visualisation of knowledge base as an information map
  - Creation of instance data through WYSIWYG editors

- For an overview of semantic wikis:
Semantic CMS: Drupal 7 & 8

- Semantic Content Management Systems use boiler plating to generate semantic content
  - Each page is machine-readable
- Drupal 7 & 8 provide RDF ‘out of the box’
  - Common ontologies included: FOAF, SIOC, SKOS, DC
  - Each user has a URI on the site, so too does each content item
- SPARQL Views
  - Enables Drupal to consume linked data via SPARQL queries
  - Integration with existing Views module to render SPARQL results
    - http://drupal.org/project/sparql_views
- Schema.org Module
  - Annotate content with schema.org semantics
  - Enables Google’s Rich Snippets to be placed in your page!
    - http://drupal.org/project/schemaorg
Facebook Open Graph

- Designed to link profile pages into Facebook’s social graph
  - Profile pages: movies, products, people, events, etc
  - Bootstrapping web content for recommendations!
- Facebook deployed their own vocabulary of terms: Open Graph Protocol

Web developers encode markup into page `<head>` element:

```html
<html xmlns="http://www.w3.org/1999/xhtml"
     xmlns:og="http://ogp.me/ns#">
  <head>
    <title>The Rock (1996)</title>
    <title>Dead Snow (2009) - IMDb</title>
    <meta property="og:url" content="http://www.imdb.com/title/tt1278340/" />
    <meta property='og:image' content='http://ia.media-imdb.com/…SY140_.jpg'>
    <meta property='og:type' content='movie' />
    <meta property='og:title' content='Dead Snow (2009)' />
    <meta property='og:site_name' content='IMDb' />
    <meta property='og:director' content='Tommy Wirkola' />
  </head>
</html>
```
Facebook Open Graph

Dead Snow (2009)

Directed by Tommy Wirkola. With Jeppe Laursen, Charlotte Frogner, Jenny Skavlan, Vegar Hoel. A ski vacation turns horrific for a group of medical students, as they find themselves confronted by an unimaginable menace: Nazi zombies.

Matthew Rowe likes a link.

Dead Snow (2009)

www.imdb.com

6 seconds ago · Like · Comment · Share
Exporting Social Semantic Data

- Facebook
  - Exports your social graph as RDF from Facebook containing FOAF and Geo data

- Twitter
  - SemanticTweet - Dead

- Flickr
  - FlickrExporter - Dead

- Passant. :me owl:sameAs flickr:33669349@N00. Linked Data on the Web Workshop, WWW2009. 2009
Reconciling Distributed Identities

Now that social graphs are provided in a common format…
- Digital Identity Reconciliation
- Multiple accounts are combined together

Explicit Linking:
- ?x owl:sameAs ?y

Implicit Linking: inverse functional
CONSTRUCT { ?x owl:sameAs ?y }
WHERE {
    ?x foaf:homepage ?h .
    ?y foaf:homepage ?h .
    FILTER (?person1 != ?person2) }


Useful for: recommendations, knowledge discovery, and …
Disambiguating Identity Web References

- Elevated web presence = problems: identity theft, lateral surveillance
- Finding all information out there about you is laborious and repetitive
- Semantic social graphs can be used to disambiguate identity web references
- We have: (a) seed data, (b) candidate web references
- We need to disambiguate references, thereby detecting web citations

SPARQL Rules:

```
CONSTRUCT { <http://www.dcs.shef.ac.uk/~mrowe/foaf.rdf#me> foaf:page ?url }
WHERE {
  ?q foaf:name ?n .
  ?e foaf:name #sam
  <http://www.dcs.shef.ac.uk/~mrowe> foaf:homepage .
  ?q foaf:name ?n .
  ?e foaf:name #sam
  <http://www.dcs.shef.ac.uk/~mrowe> foaf:homepage .
}
```
Disambiguating Identity Web References (II)

- Graph traversals and clustering:
Semi-supervised Machine Learning:
- Features: RDF instances
- Feature-weighting: (a) Graph equivalence, (b) Inverse Functional Properties, (c) RDF entailment
- Class: positive (cites the person), negative (does not)

Want to know more?
**Twitcident: Semantics and Twitter**

- Allows for exploration, search and analysis of event incidents on Twitter:  
  - [http://wis.ewi.tudelft.nl/twitcident/](http://wis.ewi.tudelft.nl/twitcident/)
- Connects to emergency broadcasts
- Monitors Twitter for incident data
- **Semantic Enrichment of Tweets:**
  - Named Entity Recognition
  - Classification: {reports, damages, risks}
  - Linkage: concept assignment
- Facilitates semantic search of Tweets
- Real-time analysis of incidents

Temporal Categorisation of POIs

- Social awareness streams (i.e. Twitter) provide information about mood and topics
- Also published on the move using ubiquitous devices

- Categorise Points of Interest (POIs) using awareness stream data:
  - Retrieves messages from a social awareness stream associated with POI
  - Perform Semantic Enrichment
  - Perform Semantic Categorisation
  - Induce categorisation function using entropy-based measures

- Enables events to be detected within geographical regions
  - Using volatility, measured by entropy, to identify the events

- Cano et al. Volatile Classification of Point of Interests based on Social Activity Streams. Social Data on the Web workshop, ISWC 2011. 2011
Role Analysis in Online Communities

- Community hosts need to know how their community is functioning and how it will evolve
- Derive the role composition of a community and correlate with its health
- For most communities we can accurately predict health changes based on role compositions – i.e. we can forecast changes
CONCLUSIONS
Conclusions

- Social Web
  - Abundance of heterogeneous data
  - Data is contained within disparate silos
- Semantic Web
  - Machine-readable information using shared ontologies
- The Social Semantic Web
  - People, Social Networks, Content, Communities, Collaboration, Folksonomies and Behaviour
  - Enables machine readability of the masses of social data
  - Collaborative environments generate and consume semantic data
  - Integration of disparate social graphs from multiple social web platforms
  - Facilitates the disambiguation of web references using social data
  - Aids recommendations by combining social data together