

Adventures in Linked Data: Building a Connected Research Environment

Lisa Goddard, Memorial University Libraries
& INKE Research Group
Access 2012, Montreal QC
October 19th, 2012

Canadian Writing Research Collaboratory (CWRC)

CFI funded initiative to establish an online infrastructure for literary research in and about Canada.



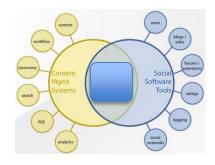
www.cwrc.ca

Canadian Writing Research Collaboratory (CWRC)

"To enable unprecedented avenues for studying the words that most move people in and about Canada."







Dissemination



Discovery

Collaboration

"We need to stop talking around the issue of the single-author monograph as the benchmark for excellence." (DH Manifesto, 2004)



Senior Scholars



Librarians



Programmers



Emergent Scholars



Project Managers

Why Linked Data for the Humanties?

Improved Search

Find all references to food in Joyce's Ulysses.

New Discovery Tools

Instead of looking for a needle in a haystack, an effective text mining tool will show you the shape of the haystack and tell you what needles are there that you'd want to stick in your foot. (Rockwell, 2011)

Interoperability



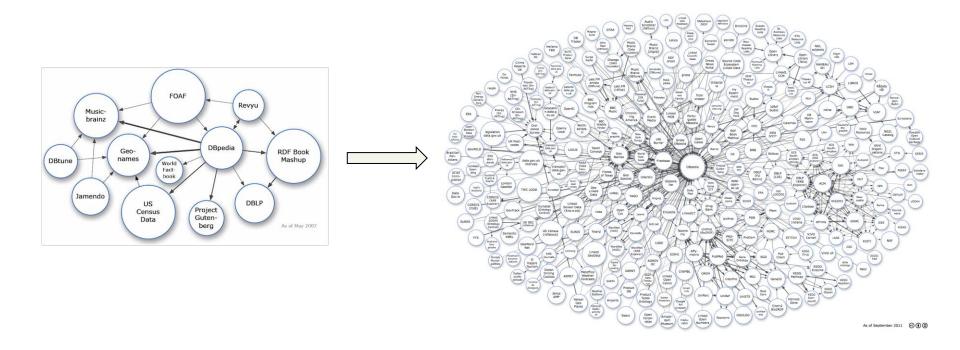


Disparate data sources and incompatible data structures are among the biggest obstacles for 21st century humanities researchers.

(RIN, 2011)

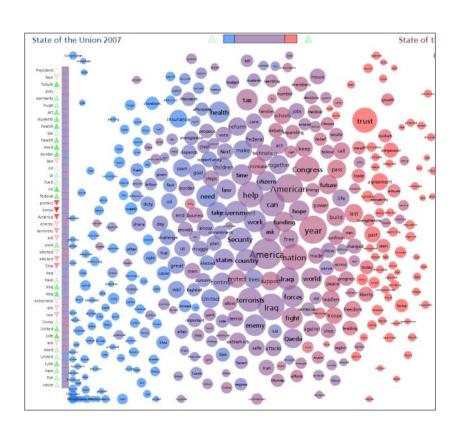


Collaboration



Linked data doesn't just accommodate collaboration, it enforces collaboration.

Big (Text) Data



Scale is a new horizon of intellectual inquiry. What kinds of humanistic phenomena appear only at scale? (Liu, 2012)

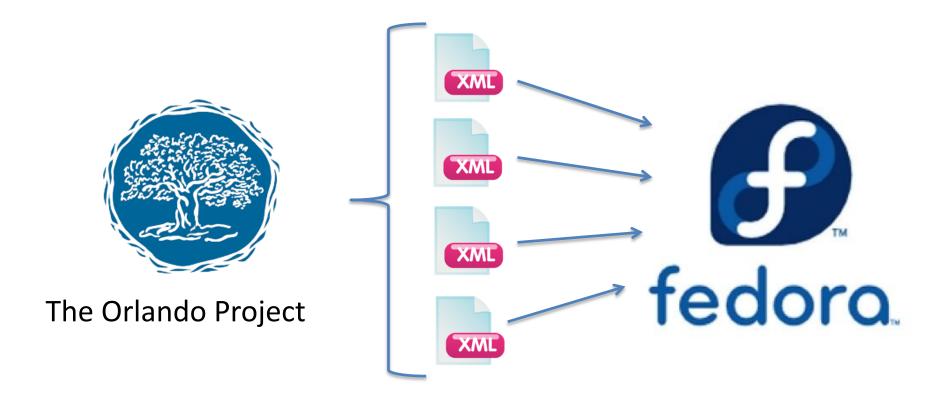
Heterogeneity

Text data is messy.

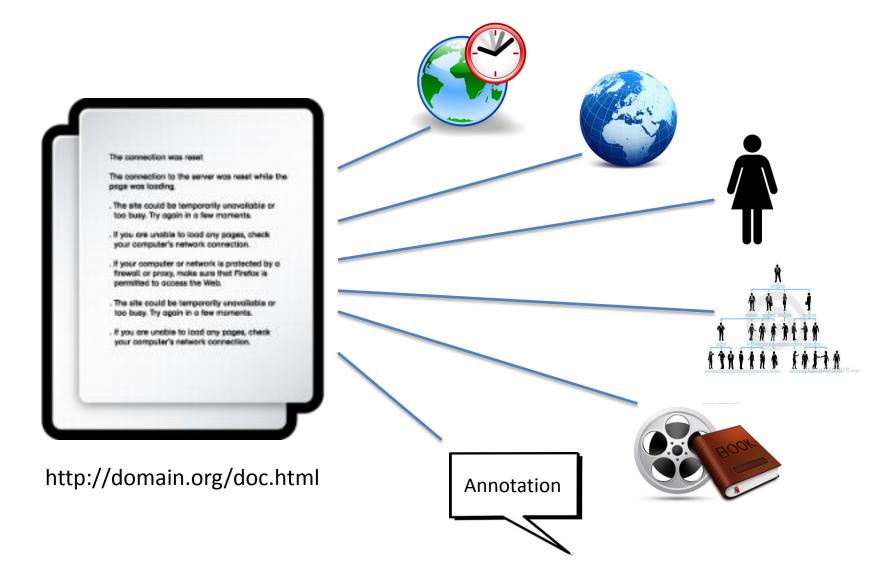


Getting Started

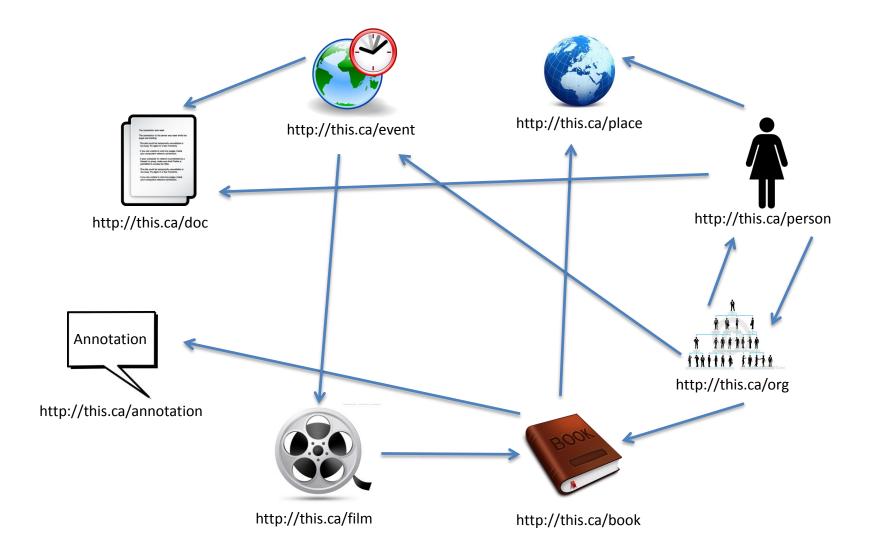
Dev Platform



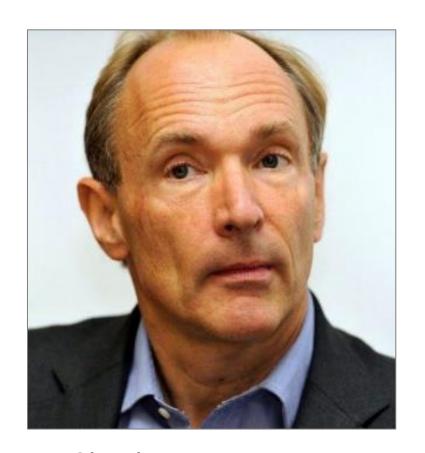
Identify Top Level Entities



Mint URIs for Entities



Cool URIs Don't Change



Sir Tim Berners-Lee is judging you.

Do you really feel that the old URIs cannot be kept running? If so, you chose them very badly.

- Sir Tim

Minting URIs

Abstract away from implementation details.

http://tiger.cwrc.ca/person.php?id=
virginia-woolf&format=rdf



http://cwrc.ca/person/virginia-woolf



Canonical URIs

http://dbpedia.org/resource/Virginia woolf

(303 redirects and content negotiation)



http://dbpedia.org/page/ Virginia_Woolf



http://dbpedia.org/data/ Virginia_Woolf.rdf

URI Patterns

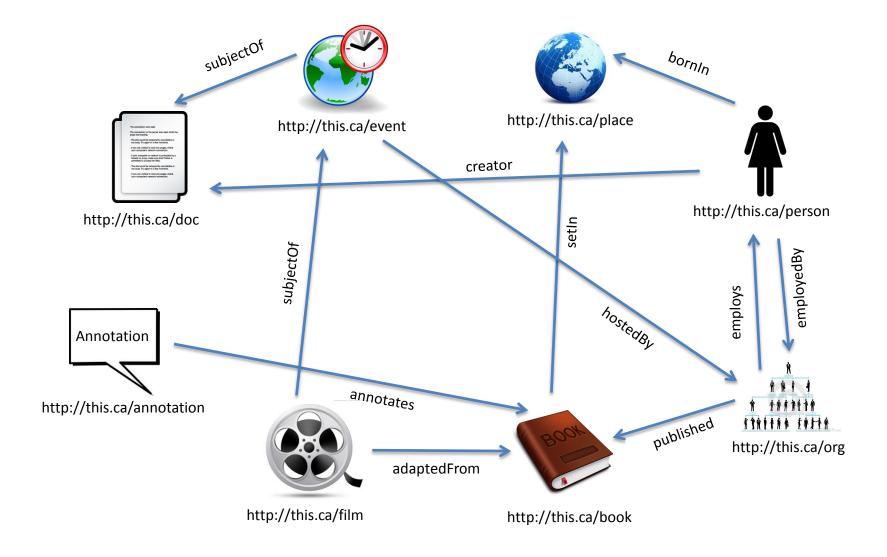
http://dbpedia.org/resource/Mary Shelley

http://id.loc.gov/authorities/names/n85300519

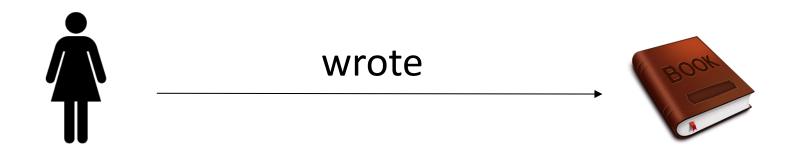
http://viaf.org/viaf/95216565/

http://cwrc.ca/person/6qMXIR5UNQSiupA

Define Relationships

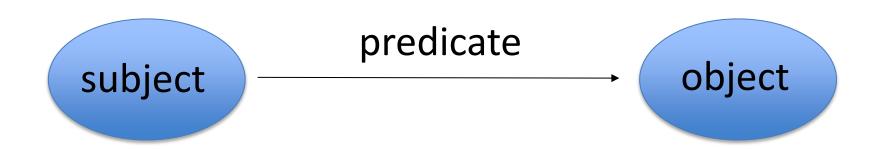


RDF Statement

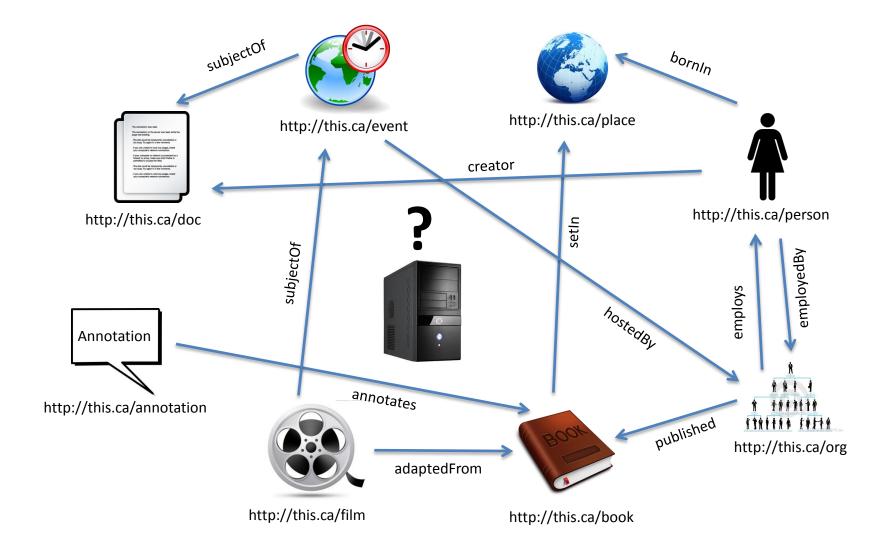


http://viaf.org/viaf/39385478/ http://dbpedia.org/resource/Mrs_Dalloway

http://purl.org/dc/elements/1.1/creator



Machine Readable Definitions



Ontologies

The semantic web is basically an accessibility initiative for machines.

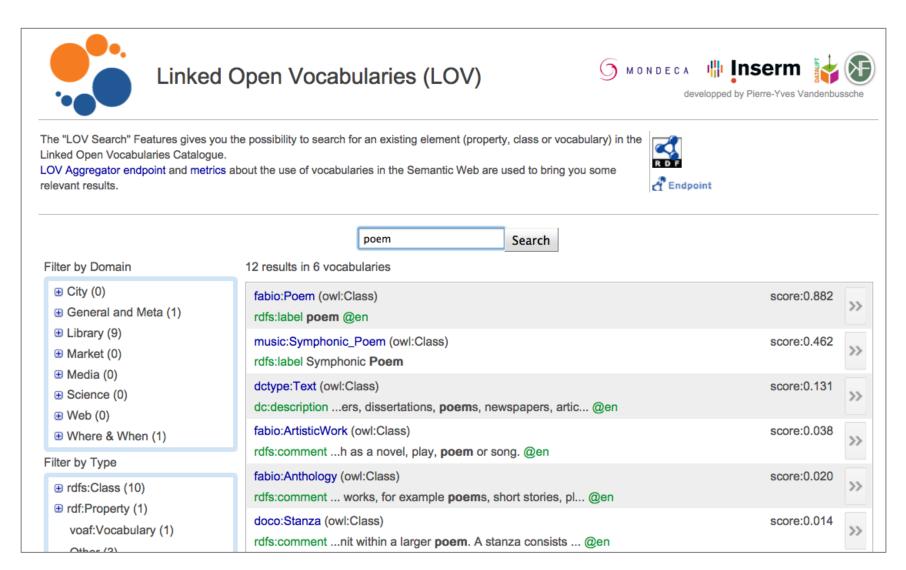
Role of Ontologies

- Define entities (classes)
- Define relationships (properties)
- Impose rules to support machine reasoning.
- Machine-readable (RDF)

Reuse or Build?

An ontology is for life.

Ontology Dowsing



foaf:Person (Class)

Class: foaf:Person

Person - A person.

Status: stable

Properties myersBriggs familyName publications lastName family name firstName currentProject surname knows workInfoHomepage pastProject geekcode schoolHomepage

workplaceHomepage img plan

Used

Of

with: knows

Subclass

Agent Spatial Thing Person

Disjoint

With:

Organization Project

The <u>Person</u> class represents people. Something is a <u>Person</u> if it is a person. We don't nitpic about whether they're alive, dead, real, or imaginary. The <u>Person</u> class is a sub-class of the <u>Agent</u> class, since all people are considered 'agents' in FOAF.

frbr:Work (Class)

work

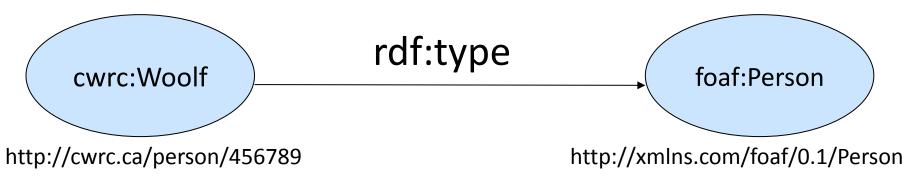
URI: http://purl.org/vocab/frbr/core#Work

A class whose members are an abstract notion of an artistic or intellectual creation. This class corresponds to the FRBR group one entity 'Work'.

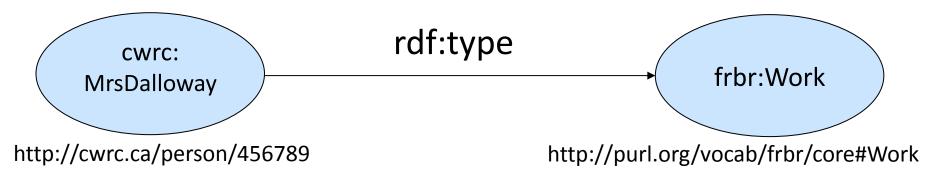
Semantics

Being a member of this class implies also being a member of endeavour and http://xmlns.com/wordnet/1.6/Work~2. No member of this class can also be a member of expression, manifestation or item. Having a realization, a creator or a subject implies being a member of this class. Things are a member of this class if they are the value of a realization of or a creator of.

Assigning Entities to Classes



http://www.w3.org/1999/02/22-rdf-syntax-ns#type



http://www.w3.org/1999/02/22-rdf-syntax-ns#type

Ontology: Predicates

Person to Person

e.g. Relationship Ontology (rel: EnemyOf)

Document to Document

e.g. Bibliographic Ontology (bibo: reviewOf)

Annotation to Document

e.g. Open Annotation Ontology (oa: hasTarget)

Ontology: Predicates



Life Partner of

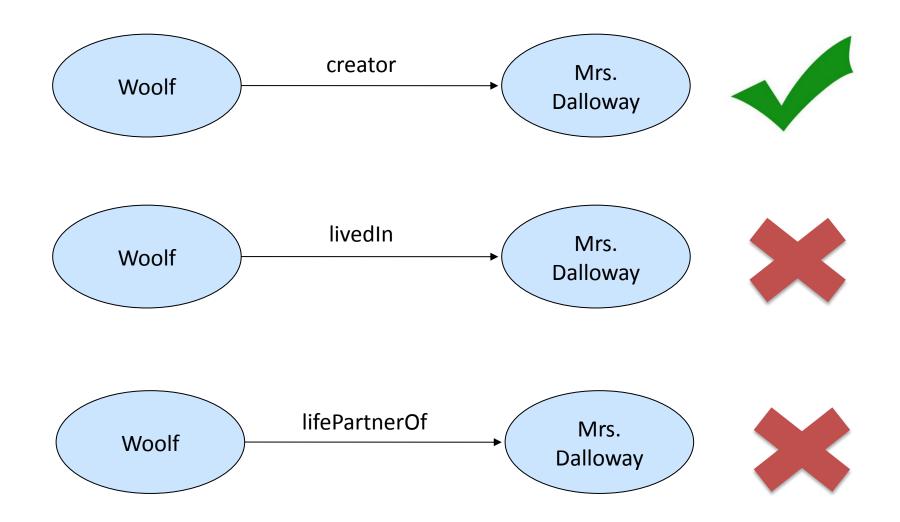
URI: http://purl.org/vocab/relationship/lifePartnerOf

A property representing a person who has made a long-term commitment to this person

Semantics

This property is symmetrical. Having this property implies being a <u>foaf:Person</u>. Every value of this property is a <u>foaf:Person</u>. It is a <u>sub-property</u> of <u>foaf:knows</u> and <u>owl:differentFrom</u>.

Classes and predicates work together to support reasoning.



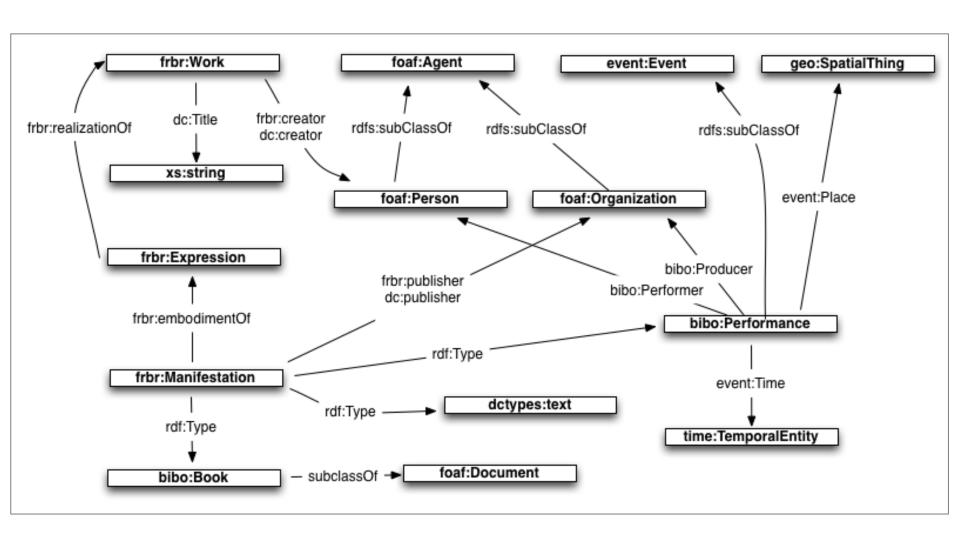
CWRC RDF Ontologies

Entity	Relevant Ontologies
Person	FOAF, MADS, EAC-CPF
Work	FRBR, MODS, DC
Place	WGS84 Geo Positioning, Geonames
Organization	FOAF, EAC-CPF
Event	OWL Time, Event
Annotation	Open Annotation, OAI-ORE

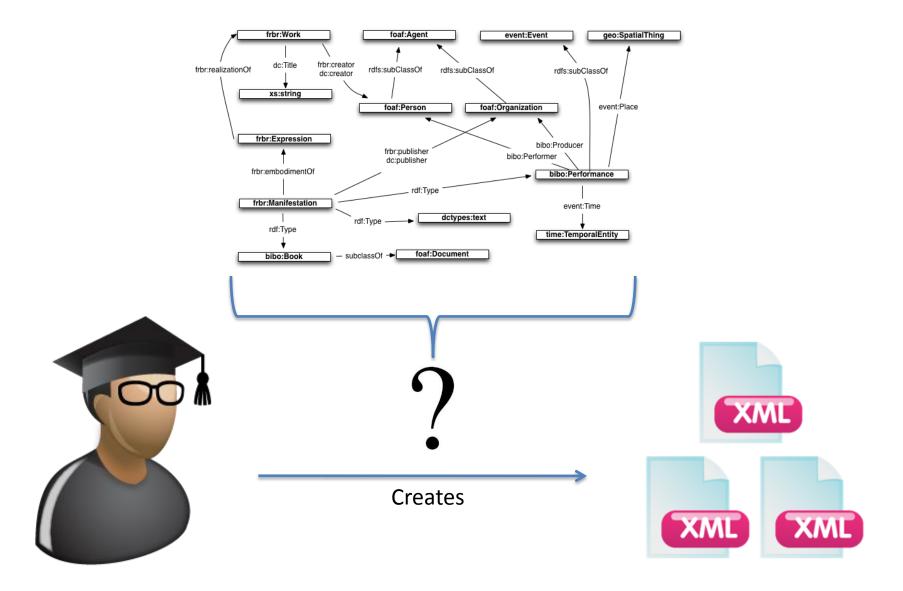
So Far

- defined our major entities and relationships
- minted URIs to represent those things within the CWRC data store
- selected ontologies that will help computers to reason about our entities and relationships

Data Model



Integrating the Data Model



CWRC-Writer

an open-source editor for online digital scholarship

This is an early prototype for demonstration purposes: certain features are not yet functional and there are many bugs. The editor works best with the current version of the Firefox browser (available here).

For more information or to help with testing CWRC-Writer, visit here.

Load an	Existing	Document
---------	----------	----------

Create a Letter

See a Sample Letter

Create a Poem

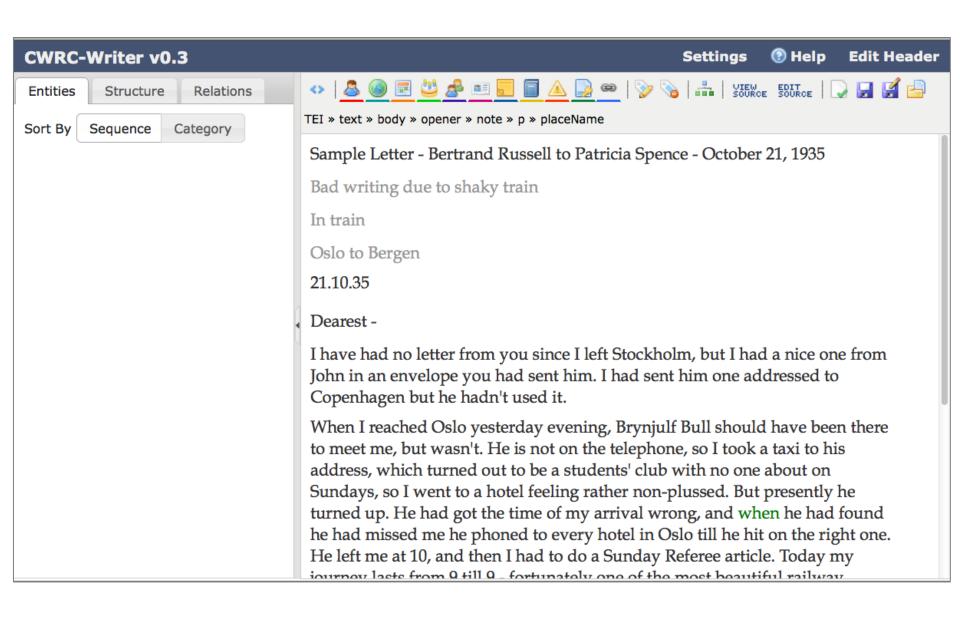
See a Sample Poem

Create Prose

Create an Event







CWRC-Writer Tagging Toolbar





























TEI » text » body » head » title

Sample Letter - Bertrand Russell to Patricia Spence - October 21, 1935

Bad writing due to shaky train

In train

Oslo to Bergen

21.10.35

Dearest -

I have had no letter from you since I left Stockholm, but I had a nice one from John in an envelope you had sent him. I had sent him one addressed to Copenhagen but he hadn't used it.

When I reached Oslo yesterday evening, Brynjulf Bull should have been there to meet me, but wasn't. He is not on the telephone, so I took a taxi to his



TEI » text » body » head » title

Sample Letter - Bertrand Russell to Pa

Bad writing due to shaky train

Results from Authority File.

Dearest -

I have had no letter from you since I l envelope you had sent him. I had sen used it.

When I reached Oslo yesterday eveni but wasn't. He is not on the telephone a students' club with no one about on non-plussed. But presently he turned when he had found he had missed might one. He left me at 10, and then I lasts from 9 till 9 - fortunately one of to Tomorrow I lecture at Bergen to the A lecture there Fri. and Sat. and then sta

Bull is a nice young man but incompe

Tag Person

Search Bertrand Russell

Ŧ

▼ Results from cwrc Project

Arthus Bertrand

Oscar Bertrand

Bertrand, Thérèse

Russell, Bertrand

Sir Bertrand: A Fragment

Jouvenel, Bertrand de

Arthus-Bertrand, Yann

Bertrand, Jean-Baptiste

Bronson, Bertrand H.

Smith, Hastings Bertrand Lees

Moleville, Antoine-François Bertrand de

Penn, Bertrand Edward,,, Viscount Dawson of

- ▶ Results from Web
- Alternate Identifier

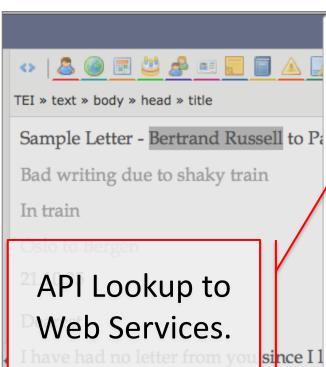
This identification is:

Definite

Reasonably Certain

Speculative

Too Danson Add Now Danson



envelope you had sent him. I had sen used it.

When I reached Oslo yesterday eveni but wasn't. He is not on the telephone a students' club with no one about on non-plussed. But presently he turned when he had found he had missed might one. He left me at 10, and then I lasts from 9 till 9 - fortunately one of to Tomorrow I lecture at Bergen to the A lecture there Fri. and Sat. and then sta

Bull is a nice young man but incompe

Tag Person

Search Bertrand Russell

- Results from cwrc Project
- Results from Web

Bertrand Russell, 1872-1970

Bertrand Russell peace foundation

Bertrand Russell society

Bertrand Russell Research Centre

Bertrand Russell Archives

Bertrand Russell Peace Foundation. Australian Branch

Bertrand Russell Centenary Symposium on Spheres of Influence in the Age of Imperialism 1972 Linz, Austria

Bertrand Russell peace foundation (New York)

Bertrand Russell Tribunal 2 1974 - 1976 Roma; Bruxelles

Bertrand Russell memorial logic conference 1971 Uldum, Danemark

Alternate Identifier

This identification is:

Definite Reasonably Certain

Speculative

Can Dawson And Mary Dawson

VIAF OpenSearch

http://viaf.org/viaf/search?query=local.names+all+%22mar garet%20atwood%22+&maximumRecords=100&startRecor d=1&sortKeys=holdingscount&httpAccept=application/rss %2Bxml

VIAF Search: margaret atwood

VIAF: Results of search: local.names all "margaret atwood"

Atwood, Margaret, 1939-....

21 September, 2012 9:54 AM

Atwood, Margaret

16 November, 2010 1:37 AM

Margaret Atwood Society

17 December, 2010 1:51 AM

Judson, Margaret Atwood, 1899-1991

22 May, 2012 4:37 AM

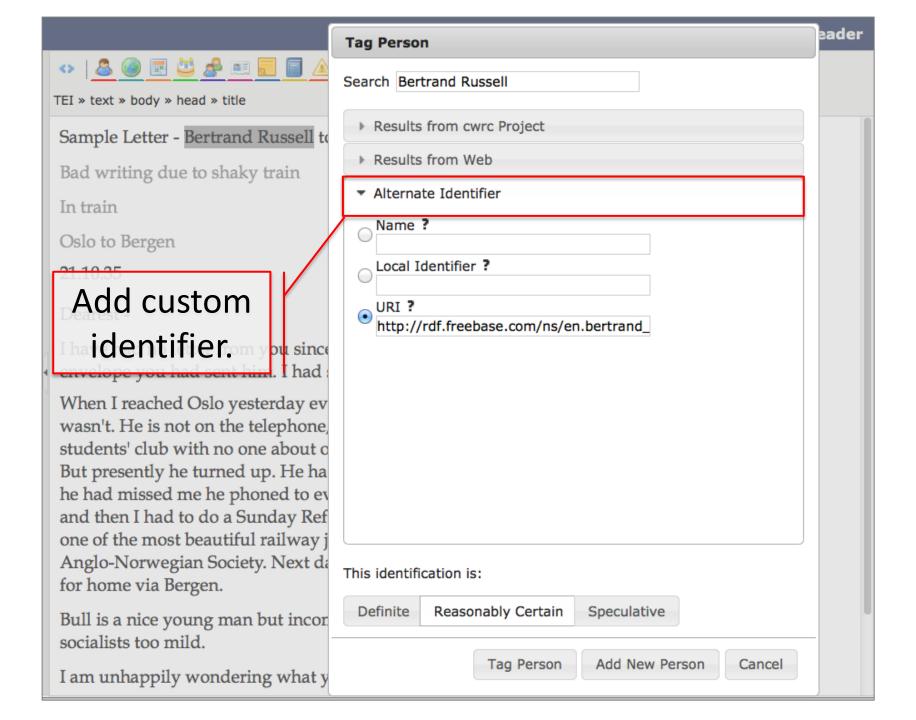
Robinson, Margaret A.

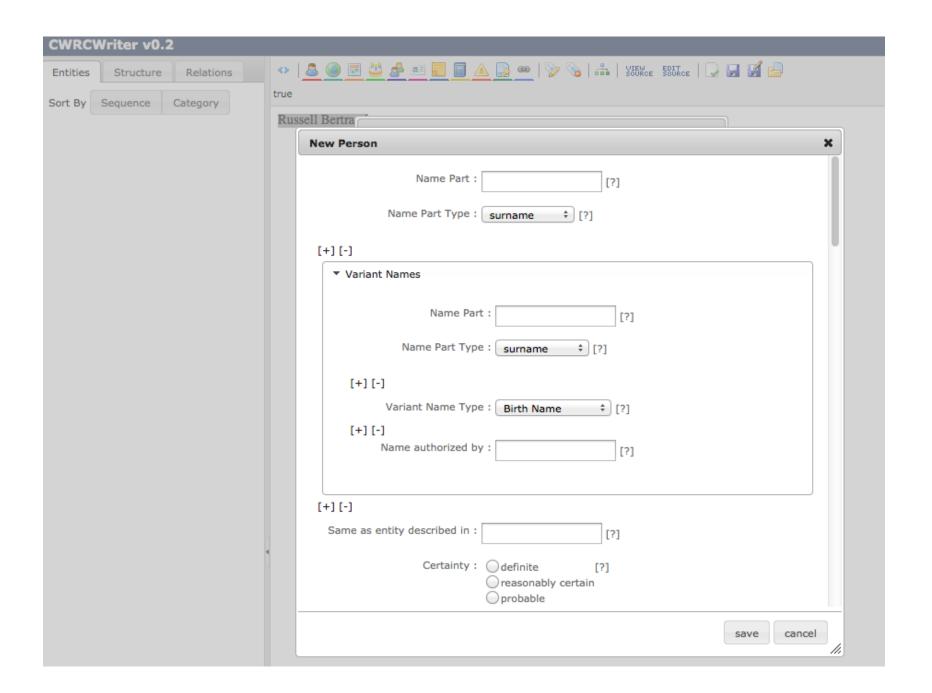
22 February, 2010 3:14 AM

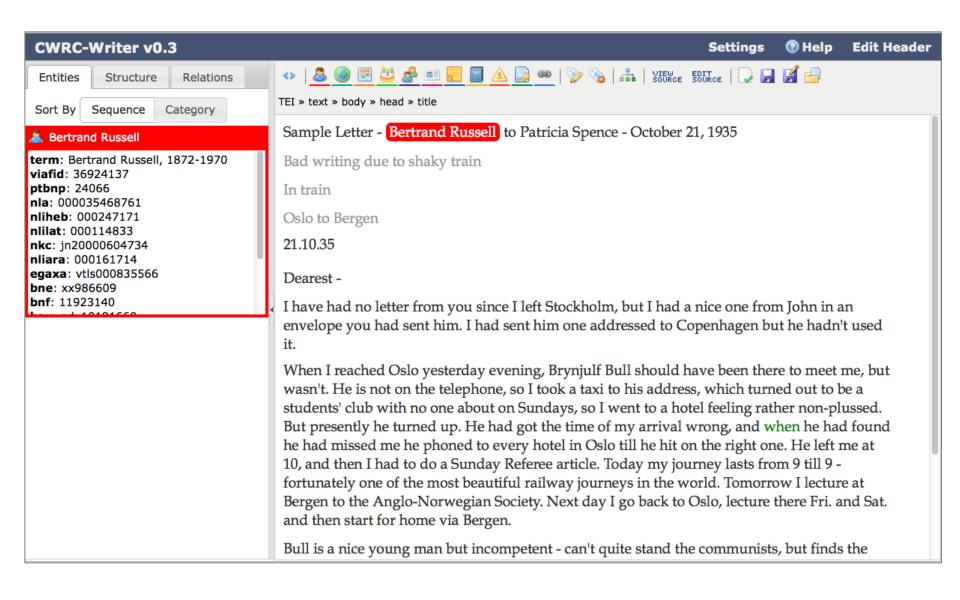
DBPedia RESTful API

curl -H "Content-Type: text/xml" -d @post.xml http://dbpedia.org/fct/service > result.xml

```
<?xml version="1.0"?>
<query
xmlns="http://openlinksw.com/services/facets/1.
0" inference="" same-as="">
    <text>Mary Shelley</text>
    <view type="text" limit="5" offset=""/>
</query>
```





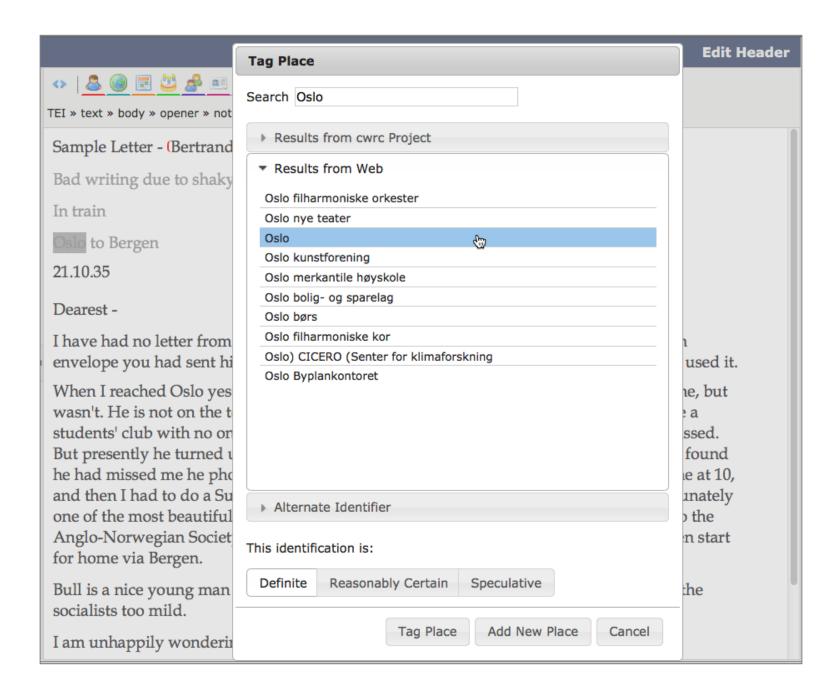


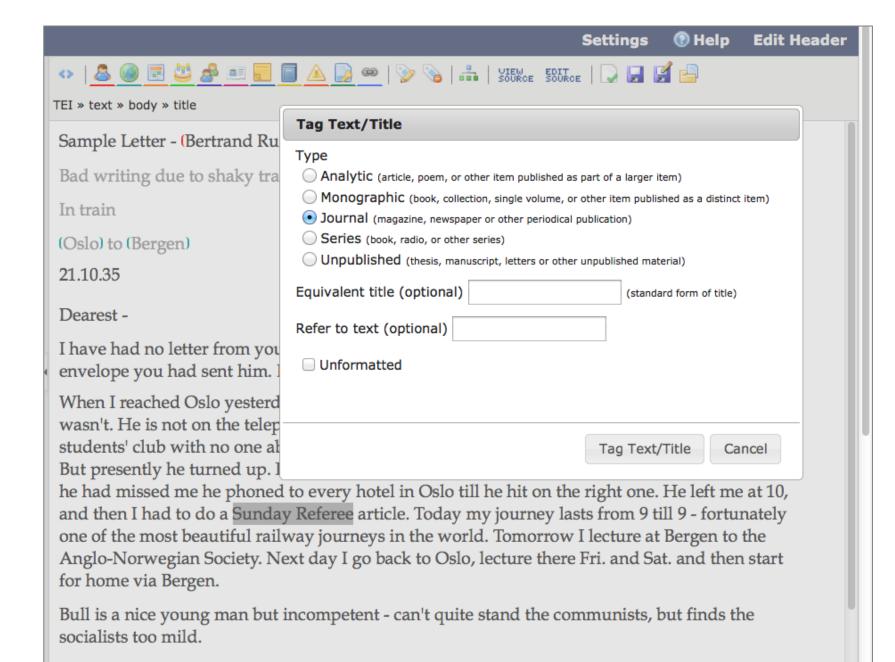
Entity in XML

```
<span name="ent_1798" class="entity person
start" type="person" entity="true" />
   Russell Bertrand
<span name="ent_1798" class="entity person
end" type="person" entity="true" />
```

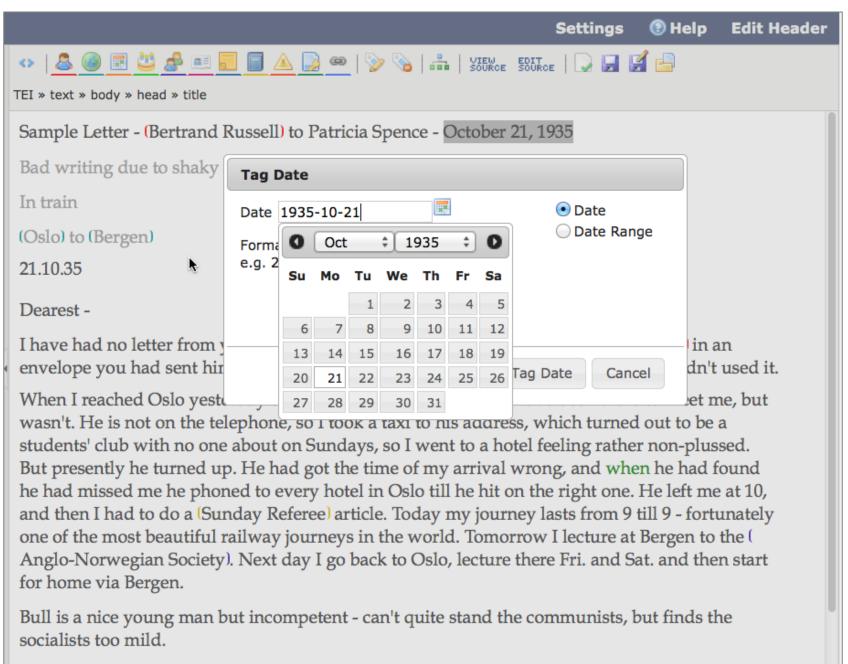
Entity in RDF

```
<rdf:Description rdf:ID="ent 1798">
      <w:type type="props">person</w:type>
      <w:content type="props">Bertrand Russell/
      w:content>
      <w:term type="info">Bertrand Russell, 1872-1970</
      w:term>
      <w:viafid type="info">36924137</w:viafid>
      <w:bnf type="info">11923140</w:bnf>
      <w:dnb type="info">118604287</w:dnb>
      <w:lc type="info">n79056054</w:lc>
      <w:certainty type="info">definite</w:certainty>
 </rdf:Description>
```





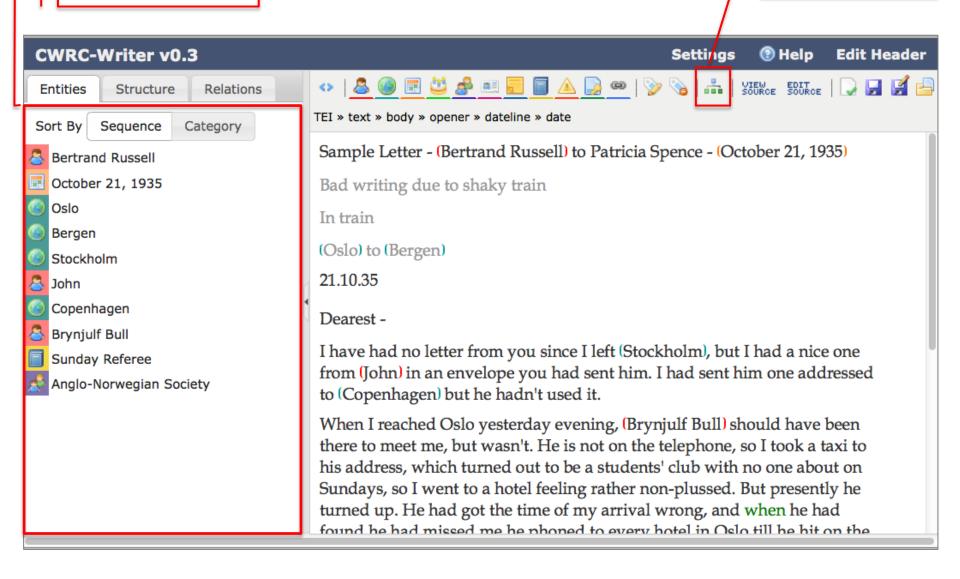
I am unhappily wondering what you are feeling about me.

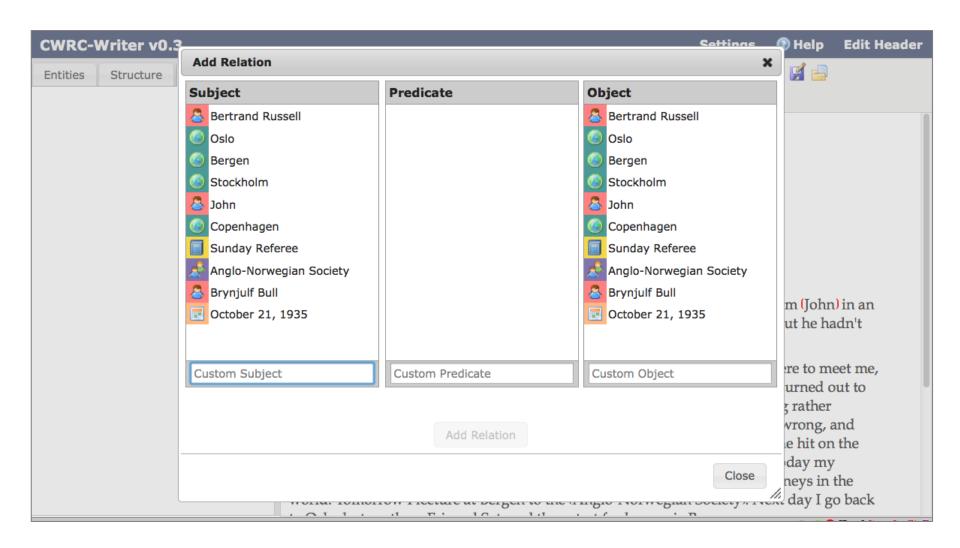


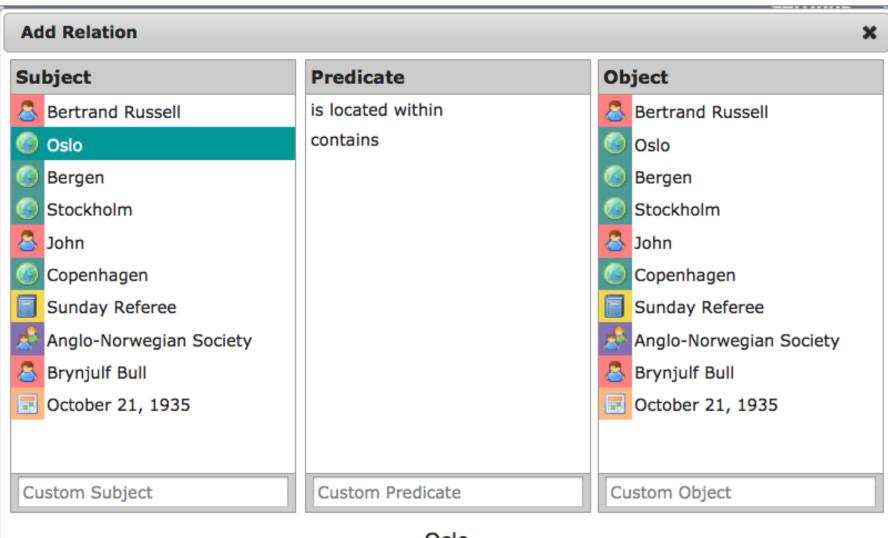
I am unhappily wondering what you are feeling about me.

List of named entities.

Add Relationship.

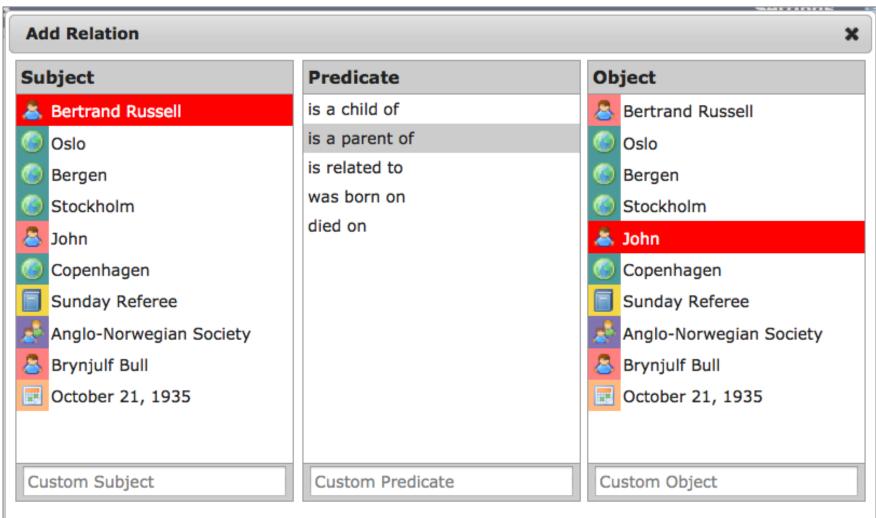






Oslo

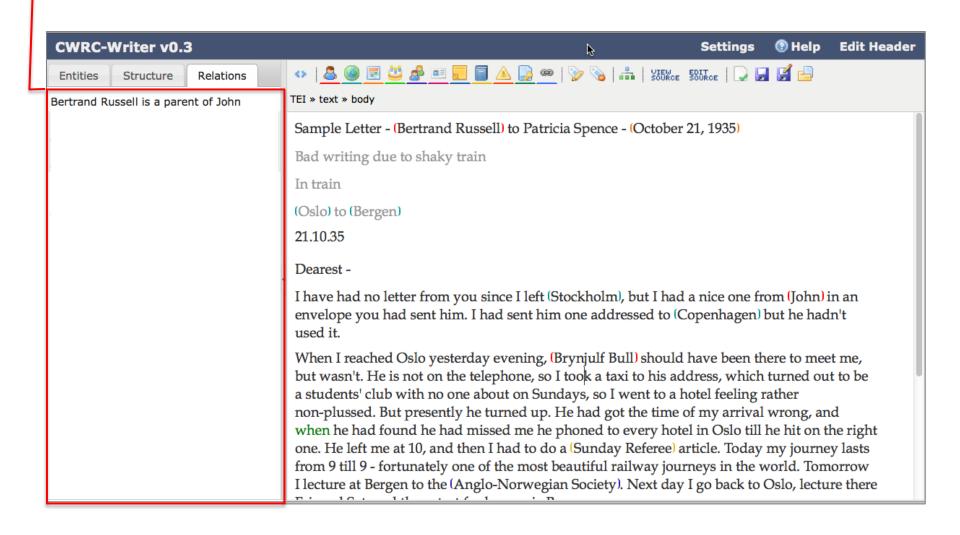
Add Relation



Bertrand Russell is a parent of John.

Add Relation

Relations Tab

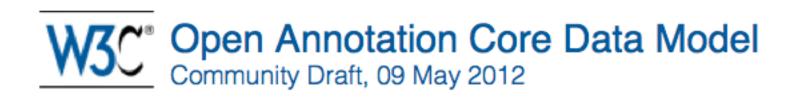


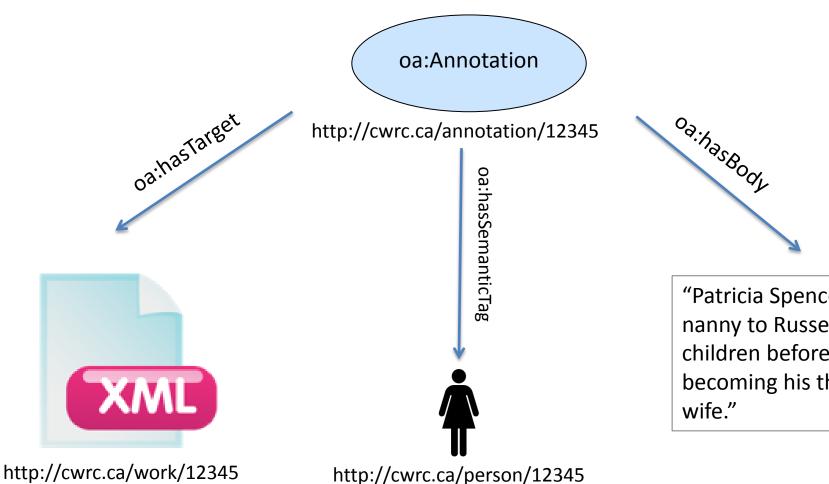
Relationship in RDF

```
<rdf:Description rdf:about="ent 1798"
  w:external="false">
   <w:isAParentOf w:text="is a parent of"
   w:external="false">
   <rdf:Description rdf:about="ent 1802"
   w:external="false" />
   </w:isAParentOf>
</rdf:Description>
```

Entity in RDF

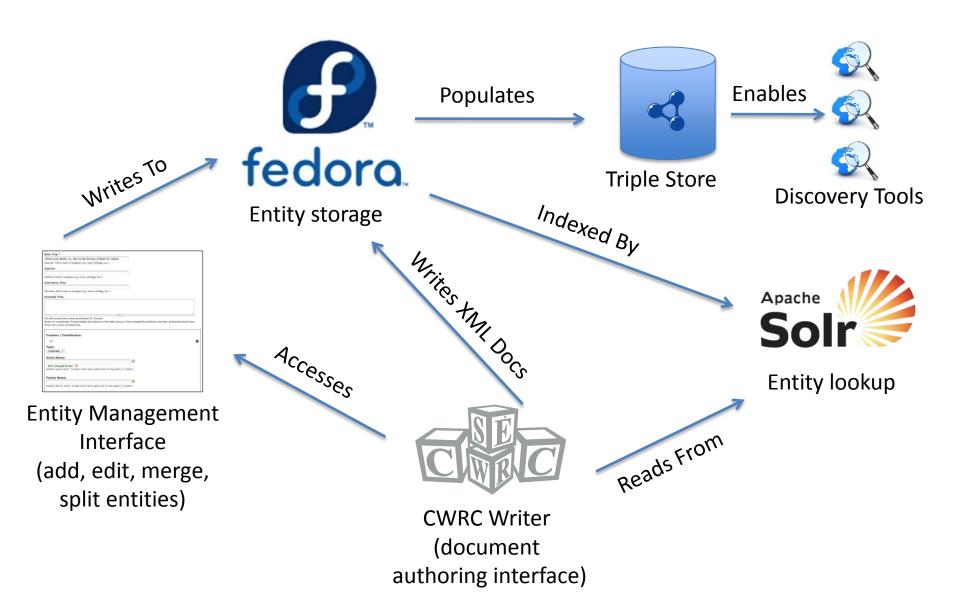
```
<rdf:Description rdf:ID="ent 1798">
     <w:id type="offset">ent 1798</w:id>
     <w:offset type="offset">117</w:offset>
     <w:length type="offset">16</w:length>
     <w:entity type="offset">true</w:entity>
     <w:type type="props">person</w:type>
     <w:content type="props">Bertrand Russell/w:content>
     <w:term type="info">Bertrand Russell, 1872-1970</w:term>
     <w:viafid type="info">36924137</w:viafid>
     <w:bnf type="info">11923140</w:bnf>
      <w:dnb type="info">118604287</w:dnb>
     <w:lc type="info">n79056054</w:lc>
     <w:certainty type="info">definite</w:certainty>
 </rdf:Description>
```





"Patricia Spence was nanny to Russell's children before becoming his third

CWRC Entity Management System



Credits

CWRC

<u>INKE</u>

Susan Brown
Mariana Parades-Olea
Jeffery Antoniuk
Denilson Barbosa
Ruth Knetchel
Omar Rodriguez-Arenas

Ray Siemens
Stan Ruecker
Harvey Quamen
John Simpson
Jentery Sayers

Related Links

CWRC Writer Beta:

http://cwrctc.artsrn.ualberta.ca/

CWRC Writer Overview:

http://www.cwrc.ca/projects/infrastructureprojects/technical-projects/cwrc-writer/

About CWRC:

http://www.cwrc.ca/about/