OAC: Machine Readable Annotations

http://www.openannotation.org/

http://groups.google.com/group/oac-discuss

Robert Sanderson
Herbert Van de Sompel

Prototyping Team
Research Library
Los Alamos National Laboratory

OAC is funded by the Andrew W. Mellon Foundation
Outline

• OAC Features Reminder
• Human Interpretable Information
• Machine Interpretable Information
OAC Features Reminder

• The Body is somehow "about" the Target

• A ConstrainedTarget identifies the Segment of Interest in the Target

• Inline Content
  • Able to include data within Annotation document

• Body/Target of any Media Type
  • Able to have Body in any format
Human Interpretable Annotation

There is the night sky filled with swirling clouds, stars ablaze with their own luminescence, and a bright crescent moon. Although the features are exaggerated, this is a scene we can all relate to, and also one that most individuals feel comfortable and at ease with.

http://www.vangoghgallery.com/painting/starryindex.html

OAC: Machine Readable Annotations
Robert Sanderson, Herbert Van de Sompel
How Can We Translate for Machines?

• The Body consists of one or more Statements
  • Human understandable: Text, Image, Video, Symbols, …
  • Machine understandable: Data

• Humans can infer relationships and context, Machines cannot
  • Need to be explicit
  • Need structured information

• When would we do this?
  • Nano-Publications: publication of data for further processing
  • Named Entity Recognition
  • Entity Relationship Extraction
**Machine Interpretable Statements**

```
uu1 foaf:depicts <dbpedia.org/resource/Moon> ;
   _:shape <dbpedia.org/resource/Crescent>;
   _:bolometricLuminosity _:x .
   _:x _:value 0.5 ;
   _:units <dbpedia.org/resource/SolarLuminosity>
```

---

OAC: Machine Readable Annotations
Robert Sanderson, Herbert Van de Sompel
Remember This One?

We'll come back to these …
Simple Data Association

Equivalent to 12 people saying: "I Like This"

I = ex:User
Like = fb:likes
This = ex:Anno
Transcription for Humans: Image to Text
Transcription for Machines: Text to Semantic Data

OAC: Machine Readable Annotations
Robert Sanderson, Herbert Van de Sompel
Multiple Entity Relationship Extractions

OAC: Machine Readable Annotations
Robert Sanderson, Herbert Van de Sompel
OAC: Machine Readable Annotations
Robert Sanderson, Herbert Van de Sompel

Single Named Entity in Text


ex: Anno
oac:hasBody
oac:hasTarget
n3
c1t

oac:Text
Constraint

oac:constrainedBy

uu1

rdfs:define
lineStart
wordStart

lineEnd
wordEnd

5


Annotea and Semantic Web Supported Collaboration

Marja-Riitta Koivunen, Ph.D.
Annotea project

Abstract
Like any other technology, the Semantic Web cannot succeed if the applications using it do not serve the needs of the users. Annotea is a Semantic Web based project for which the inspiration came from users' collaboration problems in the Web. It examined what users did naturally and selected familiar metaphors for supporting better collaboration.

The selected metaphors were a good match also for demonstrating the use of the Semantic Web technologies. Metadata was generated in the form of Annotea objects. It enhanced collaboration by adding flexibility to the applications and easy creation of different views. Furthermore, Annotea objects also let users to make the metadata available beyond its original purpose for many other Semantic Web applications.
Combination of Extraction Types

```
_:Annotea rdf:type _:SemanticWebProject ;
   _:inspiredBy _:x .
_:x rdf:type _:Problem ;
   _:domain <dbpedia.org/resource/Web> ;
   _:domain <dbpedia.org/resource/Collaboration> .

cT2 dct:references vivo:Koivunen_M-R .
vivo:Koivunen_M-R dct:creator _:Annotea
```

OAC: Machine Readable Annotations
Robert Sanderson, Herbert Van de Sompel
Advantages

• Only uses existing OAC constructions

• Aligns very closely with human annotation practices
  • Hopefully easy to interpret!

• Consistent model that scales from resource to part of resource
  • Can annotate data extracted at most appropriate level
  • Could extract from sentence/paragraph/section/entire text

• Consistent model that allows association of any amount of data:
  • From Single Entity
  • To scholarly discourse extraction from entire document
open annotation collaboration

http://www.openannotation.org/