PROMISE

Participative Research labOratory for Multimedia and Multilingual Information Systems Evaluation

FP7 ICT 2009.4.3, Intelligent Information Management

Deliverable 5.3

Collaborative User Interface Prototype with Annotation Functionalities

Version 1.00, September 2012

Network of Excellence co-funded by the 7th Framework Program of the European Commission, grant agreement no. 258191

PROMISE
Participative Research labOratory for Multimedia and Multilingual Information Systems Evaluation

SEVENTH FRAMEWORK PROGRAMME

Document Information

Deliverable number	5.3
Deliverable title	Collaborative User Interface Prototype with Annotation Functionalities
Delivery date	September 2012
Lead contractor for this deliverable	ROMA1
Author(s)	Marco Angelini, Nicola Ferro, Guido Granato, Giuseppe Santucci
Participant(s)	ROMA1, UNIPD
Workpackage	WP5
Workpackage title	Collaboration and Knowledge Sharing
Workpackage leader	ROMA1
Dissemination Level	PU – Public
Version	1.00
Keywords	FAST Annotation Service, Visual Analytics Environment, Collaboration

History of Versions

Version	Date	Status	Author	Description
0.1	2012-08-03	Draft	UNIPD	First complete draft circulated to partners as documentation for the running annotation service prototype
0.2	2012-08-29	Draft	ROMA1	Integrated documentation about the visual analytics environment
0.3	2012-08-05	Draft	ROMA1	First draft circulated to partners for review
1.0	2012-08-25	Draft	ROMA1, UNIPD	Final version after partners feedback

Abstract

This documents describes the FAST annotation service: it presents the adopted annotation model and the search model offered by the system; it discusses the architecture of the system and the RESTful Web API which can be used to access it; finally, it presents the query language which implements the adopted search model. Furthermore, it discusses the advances for the visual analytics environment and its integration with annotation functionalities.

Contents

Executive Summary

As stated in D3.3 [Agosti et al., 2012], one of the major objective of PROMISE is to design and develop an innovative evaluation infrastructure which: (i) manages and provides access to the data produced during the experimental evaluation of multilingual and multimedia information access systems; (ii) allows for the development of rich applications on top of it. Among the components that are under development, the user interface plays a key role, allowing for setting up experiments (see, e.g., the successful usage of PROMISE in CHIC 2012, as reported in D4.3 [Berendsen et al., 2012]) and for accessing and analyzing the experimental results. Moreover, it has to provide different means for promoting collaboration and information sharing across users.

Work package 5 ("Collaboration and Knowledge Sharing") is responsible for designing, developing, and delivering the user interfaces and the annotation service needed to promote the collaboration among the stakeholders of the evaluation infrastructure and foster the knowledge sharing and reuse. Moreover, it is responsible for exploring how to apply information visualization and visual analytics techniques to information retrieval experimental data in order to improve their understanding and allow researchers to effectively cope with huge amount of data.

This deliverable focuses on the first implementation of the PROMISE user interface, concentrating on the the annotation component and its relationship with the Visual Analytics component. The actual implementation leverages on user requirements, PROMISE deliverables D5.1 and D5.2 [Croce et al., 2011; Granato et al., 2011], on the detailed description of the underlying infrastructure, deliverable D3.3 [Agosti et al., 2012], and on a clear comprehension of the experimental result data structure described on deliverable D3.2 [Agosti et al., 2011].

More in detail, the deliverable describes the annotation model structure and its underlying implementation, together with the formal language used for searching the annotation hypertext and the underlying architecture. Details about how to use this component through REST Web services are provided as well. That includes API, XML and JSON representation, and the relationship with the query language CQL.

Concerning the Visual Analytics component, the deliverable shows how the PROMISE experimental data structure (D3.2 [Agosti et al., 2011]) is used in the two analysis strategies currently implemented in the prototype, i.e., per topic analysis and per experiment analysis.

Moreover, the deliverable describes the main basic functionalities of the Visual Analytics component that are orthogonal with respect to the analysis strategy, i.e., interaction functionalities (manipulation, layout editing, reference lines) and annotation and saving, showing the connection between the Visual Analytics component and the annotation component.

1 Introduction

This document describes two main components of the PRIMISE user interfaces: the annotation service, one of the main means that is provided to promote the collaboration among the stakeholders of the evaluation infrastructure and the Visual Analytics component that allows for quick and automated support to the main analysis strategies that are currently adopted by PROMISE's stakeholders. The two components are tightly connected: annotating the visual analysis is one of the most effective means to spread knowledge and share findings. Moreover, the details presented on the deliverable D3.3 [Agosti et al., 2012] allows for a clear understanding of how the visualizations produced by the Visual Analytics component are managed by the infrastructure: they are treated as "first class objects" and the system saves and retrieves not simple images but all the details behind them: track, run, experimental data, analysis strategy, and image manipulation. Moreover, the FAST annotation service adopts the same architecture of the overall PROMISE infrastructure, described in Section 5 of D3.3.

The document is organized as follows: Section 2 introduces some general concepts on annotations, Section 3 discusses the model adopted by the *Flexible Annotation Semantic Tool (FAST)* annotation service for representing and managing annotations; the XML Schema model of the annotation is presented as well as examples of annotations expressed according to this model. Section 4 introduces the notion of document annotation hypertext and provides an example of it. Section 5 presents the search capabilities offered by FAST for searching and retrieving annotations, documents, and other resources such as users, groups, and so on. Section 6 presents the *REpresentational State Transfer (REST) Application Program Interface (API)* for accessing the different resources managed by the system while Section 7 discusses the *Contextual Query Language (CQL)* context set defined for the FASTannotation service, which provides us with a well-defined syntax for querying the system. Finally, Section 8 describes the Visual Analytics components pointing out the relationship that exists with the annotation component.

2 Annotations

Almost everybody is familiar with annotations and has his own intuitive idea about what they are, drawn from personal experience and the habit of dealing with some kind of annotation in every day life, which ranges from jottings for the shopping to taking notes during a lecture or even adding a commentary to a text. This intuitiveness makes annotations especially appealing for both researchers and users: the former propose annotations as an easy understandable way of performing user tasks, while the latter feel annotations to be a familiar tool for carrying out their own tasks. Therefore, annotations have been adopted in a variety of different contexts, such as content enrichment, data curation, collaborative and learning applications, and social networks, as well as in various information management systems, such as the Web (semantic and not), digital libraries, and databases. For a thorough review of the various viewpoints about annotations, please refer to [Agosti et al., 2007; Agosti and Ferro, 2008; Handschuh and Staab, 2003; Haslhofer et al., 2009]

In general, the final recipients of annotations can be computing devices and people. The former is mainly the case of metadata/semantic annotations which allow annotated objects to be auto-

matically processed, integrated and reused in different applications, even though these metadata annotations can be understandable and useful for people too, as in the case of bibliographic catalogue records. The latter is mainly the case of content annotations which elucidate and expound on an annotated object. Note that, also in this latter case, a computing device can become the recipient of such annotations, provided that some further step of processing is performed, e.g. indexing. However, in both cases, the semantics of the annotation itself needs to be taken into consideration and modeled. This can happen formally and precisely by agreeing on metadata standards which describe how annotations should be interpreted and used; alternatively, support can be provided for identifying different pre-defined annotation types, perhaps with varying levels of detail.

The medium of the annotation can vary a lot: it can range from textual annotations, to image, audio, and video annotations; in a general setting, we may need to deal with multimedia rich annotations, composed of different parts, each with its own medium. All of these different kinds of media have to be considered and properly modeled, in a uniform way where possible.

Moreover, both annotations and annotated objects need to be uniquely identified. Furthermore, annotations comprise a temporal dimension that is often not explicit, but which limits the creation of the annotation to the existence of another object. This temporal relationship between the annotation and the annotated object does not mean that the annotation cannot be considered a standÐalone intellectual work, but it does impose a temporal ordering between the existence of an annotated object and the annotation annotating it which cannot be overlooked. In addition, once we have identified both the annotation and the annotated object, we need to link and anchor the annotation to the part of the annotated object in question. This can happen in a way that mainly depends on the medium of the annotated object. On the whole, we need to model how annotations and annotated objects are uniquely identified and linked together, maybe with a varying degree of granularity in the anchoring, paying particular attention to the temporal dimension that regulates the relationships between annotations and annotated objects.

Finally, as far as co-operation is concerned, annotations have great potential for supporting and improving interaction among users, and even among computing devices. Therefore, there is a need for modeling and offering different scopes of annotations, e.g. private, shared, or public, and managing the access rights of various groups of users.

Figure 1 summarizes the above discussion and organizes the viewpoints about annotations along two dimensions: the structure of the annotations, which ranges from content (unstructured) to metadata/semantic (very structured) and the structure of the annotated content, which ranges from the Web (unstructured) to databases (very structured). The FAST service covers many of the uses and applications of annotations discussed above, since it is able to represent and manage annotations which range from metadata to full content; its flexible and modular architecture makes it suitable for annotating general Web resources as well as digital objects managed by different digital library systems; the annotation themselves can be complex multimedia compound objects, with varying degree of visibility which ranges from private to shared and public annotations and different access rights.

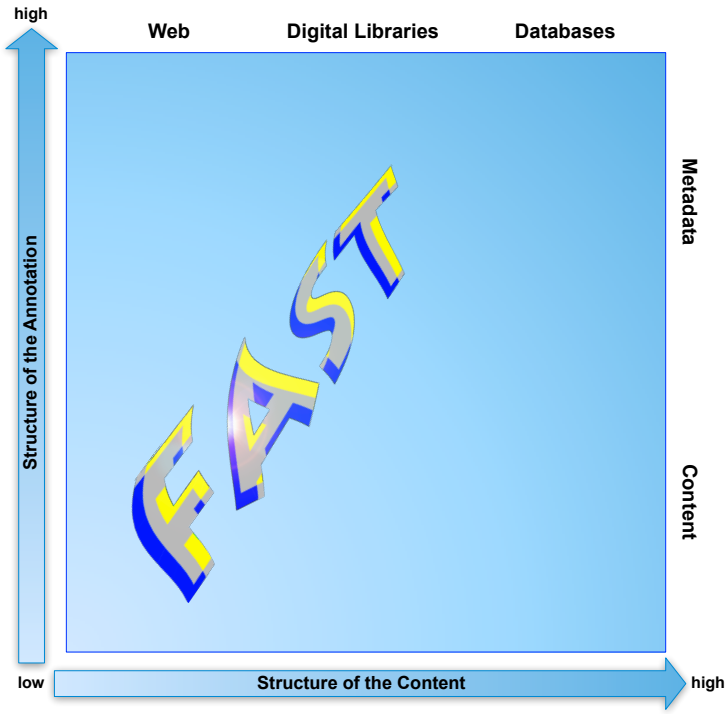

Figure 1: The FAST annotation service with respect to the various application contexts of the annotations.

3 Annotation Model

FAST adopts and implements the formal model for annotations proposed by [Agosti and Ferro, 2008] which has been also embedded in the reference model for digital libraries developed by DELOS , the European network of excellence on digital libraries [Candela et al., 2007].

According to this model, an annotation is a *compound multimedia object* which is constituted by different *signs* of annotation. Each sign materializes part of the annotation itself; for example, we can have textual signs, which contain the textual content of the annotation, image signs, if the annotation is made up of images, and so on. In turn, each sign is characterized by one or more *meanings* of annotation, which specify the semantics of the sign; for example, we can have a sign whose meaning corresponds to the title field in the *Dublin Core (DC)* metadata schema, in the case of a metadata annotation, or we can a sign carrying a question of the author about a document whose meaning may be “question" or similar.

An annotation is uniquely identified by an handle, which usually takes the form of a pair (namespace, identifier), where the namespace represents a space providing a logical grouping for the entities, so that occurrences of the same entity with the same identifier but belonging to different namespace can be distinguished, very much like the concept of namespace in *eXtensible Markup Language (XML)* [W3C, 2006, 2008].

An annotation has a scope which defines its visibility (public, shared, or private), and can be

shared with different groups of users. Public annotations can be read by everyone and modified only by their owner; shared annotations can be modified by their owner and accessed by the specified list of groups with the given access permissions, e.g. read only or read/write; private annotations can be read and modified only by their owner.

In the following, we need a terminology to distinguish between two kinds of digital objects: the generic ones managed by a digital library or available in the Web, which we call documents, and the ones that are annotations. Therefore, when we use the generic term digital object, we mean a digital object that can be either a document or an annotation.

Annotations can be linked to digital objects with two main types of links:

- *annotate link*: an annotation annotates a digital object, which can be either a document or another annotation.
  The “annotate link" is intended to allow an annotation only to annotate one or more parts of a given digital object. Therefore, this kind of link lets the annotation express intraÐdigital object relationships, meaning that the annotation creates a relationship between the different parts of the annotated digital object;

- *relate-to link*: an annotation relates to a digital object, which can be either a document or another annotation.
  The “relate-to link" is intended to allow an annotation only to relate to one or more parts of other digital objects, but not the annotated one. Therefore, this kind of link lets the annotation express interÐdigital object relationships, meaning that the annotation creates a relationship between the annotated digital object and the other digital objects related to it.

With respect to these two main types of link, we introduce the following constraint: an annotation must annotate one and only one digital object, which can be either a document or another annotation, i.e. an annotation must have one and only one “annotate link".

This constraint means that an annotation can be created only for the purpose of annotating a digital object and not exclusively for relating to a digital object. An annotation, then, can annotate one and only one digital object, because the “annotate link" expresses intraÐdigital object relationships and thus it cannot be mutual to multiple digital objects different from the annotated one. Finally, this constraint does not prevent the annotation from relating to more than one digital object, i.e. from having more than one “relate-to link".

This situation is very similar to what happens in the real world. When we deal with paper documents, we can annotate one or more parts of the document which we have at hand; this document also provides us with the physical medium for writing the content of the annotation. On the other hand, the content of the annotation can contain references to other documents; in other words, it can relate the document at hand to other documents that are currently being looked. Therefore, the act of annotating concerns one and only one document, to which the annotation is anchored, although there may be one or more references that relate the annotation to other documents. One could argue that in the digital world these limitations could be overcome and that an annotation could annotate multiple documents at the same time. Apart from being possible, what could we gain from this option? If we allow multiple annotate links, we are going to add some uncertainty because the

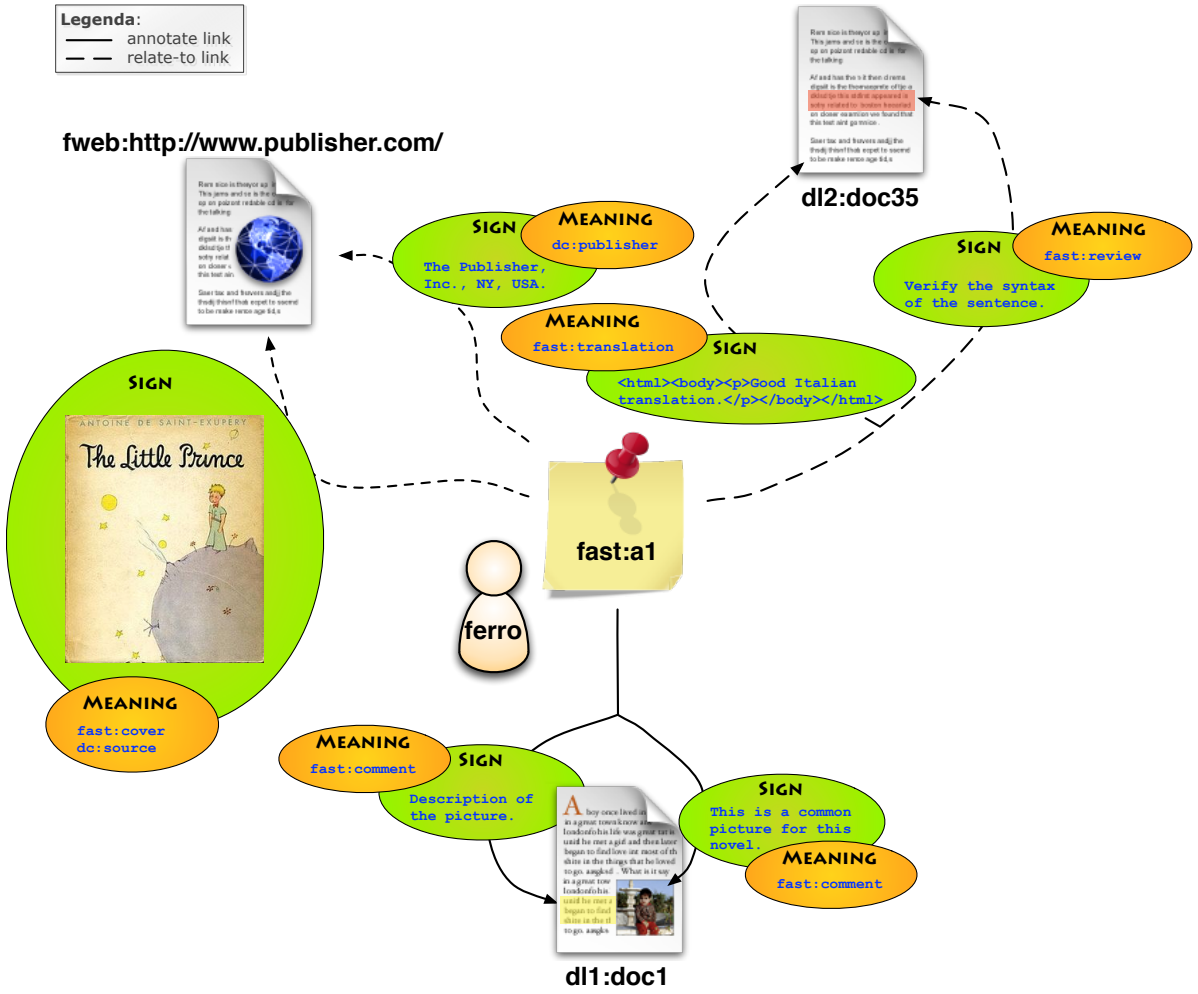

Figure 2: Example of annotation.

annotation would lose its strong relationships with only one object. In fact, this object represents its main purpose while linking the annotation to multiple objects would give us unclear semantics.

3.1 Example of Annotation

Figure 2 shows an example of annotation which summarizes the discussion so far. The annotation, with identifier `a1` and namespace `fast`, is authored by the user `ferro`[1]. It annotates a document containing a novel, whose identifier is `doc1` and which belongs to the namespace `dl1` of a digital library which manages it. The annotation relates to another document containing a translation of the

[1] Also the users of the system are identified by using a pair (identifier, namespace) but for space reasons the namespace for the user `ferro` is not reported in the figure.

novel, whose identifier is doc35 and which belongs to the namespace dl2 of a digital library different from the one which manages doc1; in addition, it relates also to the Web page of the publisher of the novel, whose identifier is http://www.publisher.com/ and which belongs to the namespace fweb, used for indicating Web resources.

In particular, a1 annotates two distinct parts of doc1. It annotates an image contained in the *Portable Document Format (PDF)* of the novel by using a textual sign whose content is “This is a common picture for this novel" and whose meaning is to be a comment in the fast namespace. It also annotates a sentence by using another textual sign whose content is “Description of the picture" and whose meaning is to be a comment in the fast namespace.

As previously discussed and made clear by the example, the annotate link is capable of expressing intra-digital object relationships, as discussed above, since it relates different parts of the same annotated digital object.

a1 relates the document doc1 to its Italian translation by linking to the whole document doc35 with a textual sign whose content is “Good Italian translation" and whose meaning is to be a translation in the fast namespace. It also relates to a specific sentence of the translation with a *HyperText Markup Language (HTML)* sign which asks to “Verify the syntax of the sentence" and whose meaning is to be a review in the fast namespace.

Finally, a1 also relates the document to the Web page of the publisher of the novel with a textual sign whose content is “The Publisher, Inc., NY, USA" and whose meaning is to be the publisher field of the DC metadata schema [ISO 15836:2009, 2009]. It also related the document to the Web page of the publisher via an image sign, containing the cover of the printed book of the novel by the publisher, and whose meaning is to be both a source in the DC metadata schema and a cover in the fast namespace.

As previously discussed and made clear by the example, the relate-to link is able to express inter-digital object relationships since, in this example, related a novel with both its Italian translation and its publisher’s Web page.

This example illustrates how annotations can range from metadata to content and they can also mix up metadata and content, as in the case of annotation a1. The example shows also another important feature of the annotations: they can take a part of a hypertext, as observed by [Agosti et al., 2004; Halasz et al., 1987], since they enable the creation of new relation ships among existing objects, by means of links that connect annotations together with existing objects, as we will see later in more detail. In addition, this hypertext can span and cross the boundaries of the single digital library, if users need to interact with the information resources managed by diverse digital libraries. In the example above, the annotation a1 links the document doc1 which belongs to the digital library dl1 to both the document doc35 which belongs to the digital library dl2 and the Web page http://www.publisher.com.

3.2 XML Schema Overview

The formal model for annotation discussed in the previous section has provided a sound basis for designing and developing an XML Schema [W3C, 2004b,c,d] for the FAST annotation service. The

FAST XSchema[2] allows annotations and related entities to be represented and exchanged into a well-defined XML format.

Figure 3 shows the fragment of the schema concerning the annotation, which offers the possibility to explore also the other entities related to it.

An annotation has the following attributes: `identifier` is a unique identifier for the annotation; `namespace`, typically a *Uniform Resource Identifier (URI)* [Berners-Lee et al., 2005]; `scope` specifies if the annotation is private, shared, or public; `created` and `last-modified` represent, respectively, the creation date and the last modified date of the annotation[3].

An annotation is authored by one `user` and can be shared among `groups` of users with the desired `access-permission`.

An annotation is materialized by one or more `signs`. A sign, in turn, is characterized by the following attributes: a unique `identifier`; `created` and `last-modified` represent, respectively, the creation date and the last modified date of the sign; `media-type` is the *Multipurpose Internet Mail Extensions (MIME)* media type [Freed and Borenstein, 1996a,b] of the actual content of the sign while `language`, if any, is the language of the content of the sign according to the ISO three-letters language codes [ISO 639-2, 1998]; the `content` element is the actual content of the sign and, in the case of a binary content is encoded Base64 [Josefsson, 2006] as specified in the `content-transfer-encoding` attribute.

Each sign holds an annotate `anchor`, which corresponds to the “annotate link" discussed above. The annotate `anchor` specifies the anchored digital object or annotation and may have a `pointer` to a specific part of the annotated digital object. It is up to the application using the FAST annotation service decide and interpret the format of the pointer in a meaningful way for the kind and MIME media type of the annotated digital object. To facilitate the task, the `pointer` element may have a `anchored-media-type` attribute which specifies the MIME media type of the anchored part: for example, you may anchor an image (`image/png`) within a PDF document (`application/pdf`). If no `pointer` element is specified, it means that the whole digital object is being annotated.

Similarly, a sign may hold a relate-to `anchor`, which corresponds to the “relate-to link" discussed above.

The semantics and meaning of each sign is expressed by one or more `concept` elements, viewed as an idea or notion, a unit of thought very much like the notion of concept in *Simple Knowledge Organization System (SKOS)* [W3C, 2009a,b].

3.3 Example of Annotation in XML

The following code listing reports how the annotation discussed in Section 3.1 and shown in Figure 2 is represented in XML according to the FAST Xschema.

```
1 <?xml version="1.0" encoding="UTF-8"?>
2 <ims:fast xmlns:ims="http://ims.dei.unipd.it/" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
3   xsi:schemaLocation="http://ims.dei.unipd.it/ http://ims.dei.unipd.it/data/xml/fast.3.00.xsd">
4   <ims:annotation ims:identifier="a1" ims:namespace="http://fast.dei.unipd.it/"
5     ims:scope="SHARED"
```

[2] http://ims.dei.unipd.it/data/xml/fast.3.00.xsd

[3] The last modification timestamp is used for adopting an optimistic locking strategy

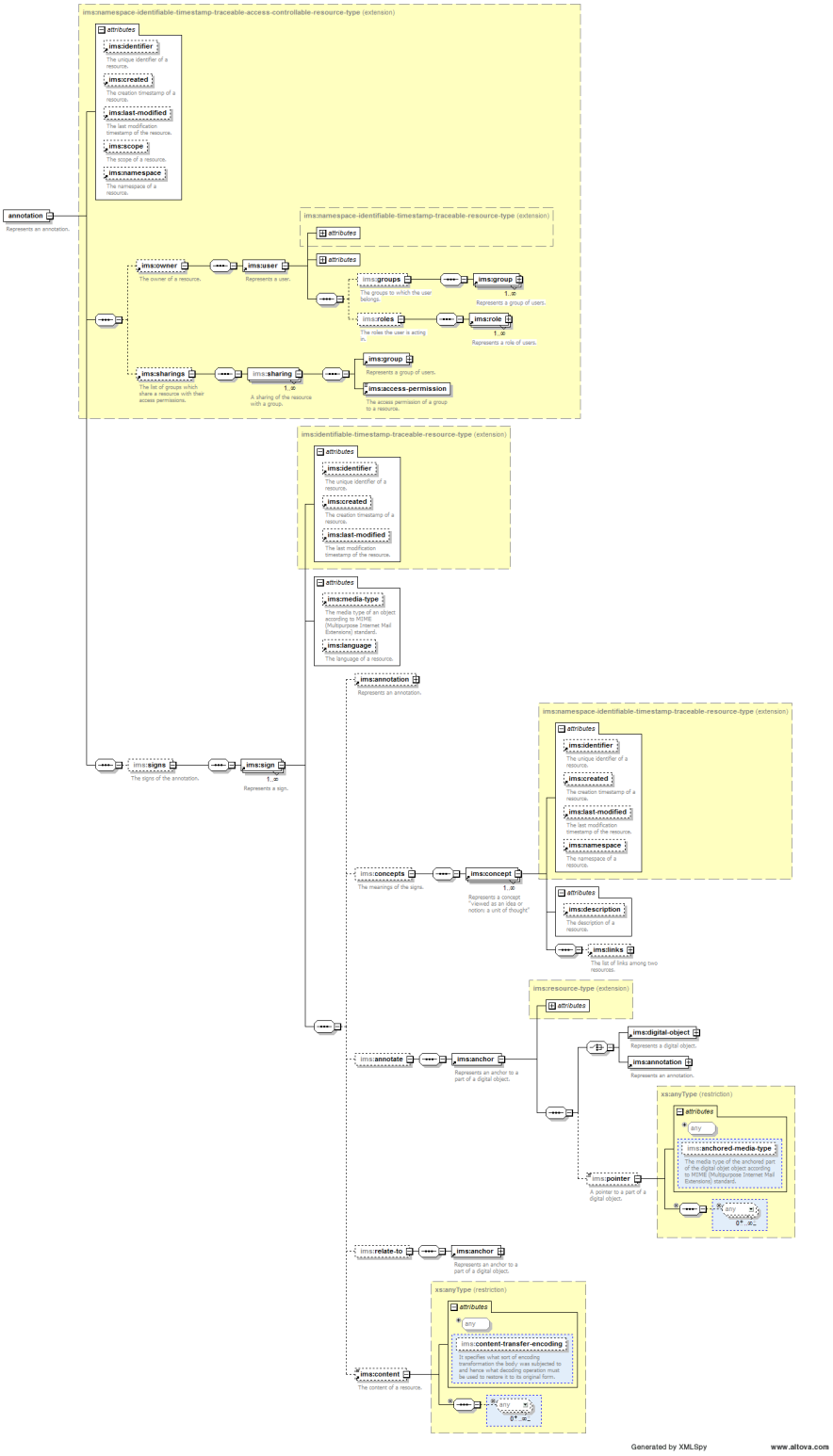

Figure 3: Overview of the XML schema for representing annotations.

```
6     ims:created="2012-07-31T18:42:20.563+02:00"
7     ims:last-modified="2012-07-31T18:42:20.563+02:00">
8     <ims:owner>
9       <ims:user ims:identifier="ferro" ims:namespace="http://fast.dei.unipd.it/" />
10    </ims:owner>
11    <ims:sharings>
12      <ims:sharing>
13        <ims:group ims:identifier="group-1" ims:namespace="http://fast.dei.unipd.it/" />
14        <ims:access-permission>
15          DENIED
16        </ims:access-permission>
17      </ims:sharing>
18      <ims:sharing>
19        <ims:group ims:identifier="group-2" ims:namespace="http://fast.dei.unipd.it/" />
20        <ims:access-permission>
21          READ_ONLY
22        </ims:access-permission>
23      </ims:sharing>
24      <ims:sharing>
25        <ims:group ims:identifier="group-3" ims:namespace="http://fast.dei.unipd.it/" />
26        <ims:access-permission>
27          READ_WRITE
28        </ims:access-permission>
29      </ims:sharing>
30    </ims:sharings>
31    <ims:signs>
32      <ims:sign ims:identifier="sign-1"
33        ims:media-type="text/plain" ims:language="eng"
34        ims:created="2012-07-31T18:42:20.556+02:00"
35        ims:last-modified="2012-07-31T18:42:20.556+02:00">
36        <ims:concepts>
37          <ims:concept ims:identifier="comment" ims:namespace="http://fast.dei.unipd.it/" />
38        </ims:concepts>
39        <ims:annotate>
40          <ims:anchor>
41            <ims:digital-object ims:identifier="doc1" ims:namespace="http://www.dl1.org/"
42              ims:media-type="application/pdf" />
43            <ims:pointer ims:anchored-media-type="image/jpeg">
44              <selection>
45                <path>
46                  /root/page1/image1
47                </path>
48                <area>
49                  10 5 20 10
50                </area>
51              </selection>
52            </ims:pointer>
53          </ims:anchor>
54        </ims:annotate>
55        <ims:content>
56          This is a common picture for this novel.
57        </ims:content>
58      </ims:sign>
59      <ims:sign ims:identifier="sign-2"
60        ims:media-type="text/plain" ims:language="eng"
61        ims:created="2012-07-31T18:42:20.560+02:00"
62        ims:last-modified="2012-07-31T18:42:20.560+02:00">
63        <ims:concepts>
64          <ims:concept ims:identifier="comment" ims:namespace="http://fast.dei.unipd.it/" />
65        </ims:concepts>
```

```
66          <ims:annotate>
67            <ims:anchor>
68              <ims:digital-object ims:identifier="doc1" ims:namespace="http://www.dl1.org/"
69                ims:media-type="application/pdf" />
70              <ims:pointer ims:anchored-media-type="text/plain">
71                /root/page1/sentence3[10-30]
72              </ims:pointer>
73            </ims:anchor>
74          </ims:annotate>
75          <ims:content>
76            Description of the picture.
77          </ims:content>
78        </ims:sign>
79        <ims:sign ims:identifier="sign-3"
80          ims:media-type="text/html" ims:language="eng"
81          ims:created="2012-07-31T18:42:20.560+02:00"
82          ims:last-modified="2012-07-31T18:42:20.560+02:00">
83          <ims:concepts>
84            <ims:concept ims:identifier="translation" ims:namespace="http://fast.dei.unipd.it/" />
85          </ims:concepts>
86          <ims:annotate>
87            <ims:anchor>
88              <ims:digital-object ims:identifier="doc1" ims:namespace="http://www.dl1.org/"
89                ims:media-type="application/pdf" />
90            </ims:anchor>
91          </ims:annotate>
92          <ims:relate-to>
93            <ims:anchor>
94              <ims:digital-object ims:identifier="doc35" ims:namespace="http://www.dl2.org/"
95                ims:media-type="text/rtf" />
96            </ims:anchor>
97          </ims:relate-to>
98          <ims:content>
99            <html>
100             <body>
101               <p>
102                 Good Italian translation.
103               </p>
104             </body>
105           </html>
106         </ims:content>
107       </ims:sign>
108       <ims:sign ims:identifier="sign-4"
109         ims:media-type="text/plain" ims:language="eng"
110         ims:created="2012-07-31T18:42:20.560+02:00"
111         ims:last-modified="2012-07-31T18:42:20.560+02:00">
112         <ims:concepts>
113           <ims:concept ims:identifier="review" ims:namespace="http://fast.dei.unipd.it/" />
114         </ims:concepts>
115         <ims:annotate>
116           <ims:anchor>
117             <ims:digital-object ims:identifier="doc1" ims:namespace="http://www.dl1.org/"
118               ims:media-type="application/pdf" />
119           </ims:anchor>
120         </ims:annotate>
121         <ims:relate-to>
122           <ims:anchor>
123             <ims:digital-object ims:identifier="doc35" ims:namespace="http://www.dl2.org/"
124               ims:media-type="text/rtf" />
125             <ims:pointer ims:anchored-media-type="text/rtf">
```

```
126                100 0 27
127              </ims:pointer>
128            </ims:anchor>
129          </ims:relate-to>
130          <ims:content>
131            Verify the syntax of the sentence.
132          </ims:content>
133        </ims:sign>
134        <ims:sign ims:identifier="sign-5"
135          ims:media-type="text/plain" ims:language="eng"
136          ims:created="2012-07-31T18:42:20.560+02:00"
137          ims:last-modified="2012-07-31T18:42:20.560+02:00">
138          <ims:concepts>
139            <ims:concept ims:identifier="publisher" ims:namespace="http://purl.org/dc/elements/1.1/" />
140          </ims:concepts>
141          <ims:annotate>
142            <ims:anchor>
143              <ims:digital-object ims:identifier="doc1" ims:namespace="http://www.dl1.org/"
144                ims:media-type="application/pdf" />
145            </ims:anchor>
146          </ims:annotate>
147          <ims:relate-to>
148            <ims:anchor>
149              <ims:digital-object ims:identifier="http://www.publisher.com/"
150                ims:namespace="http://fast.dei.unipd.it/web/" ims:media-type="text/html" />
151            </ims:anchor>
152          </ims:relate-to>
153          <ims:content>
154            The Publisher, Inc., NY, USA.
155          </ims:content>
156        </ims:sign>
157        <ims:sign ims:identifier="sign-6"
158          ims:media-type="image/jpeg"
159          ims:created="2012-07-31T18:42:20.560+02:00"
160          ims:last-modified="2012-07-31T18:42:20.560+02:00">
161          <ims:concepts>
162            <ims:concept ims:identifier="cover" ims:namespace="http://fast.dei.unipd.it/" />
163            <ims:concept ims:identifier="source" ims:namespace="http://purl.org/dc/elements/1.1/" />
164          </ims:concepts>
165          <ims:annotate>
166            <ims:anchor>
167              <ims:digital-object ims:identifier="doc1" ims:namespace="http://www.dl1.org/"
168                ims:media-type="application/pdf" />
169            </ims:anchor>
170          </ims:annotate>
171          <ims:relate-to>
172            <ims:anchor>
173              <ims:digital-object ims:identifier="http://www.publisher.com/"
174                ims:namespace="http://fast.dei.unipd.it/web/" ims:media-type="text/html" />
175            </ims:anchor>
176          </ims:relate-to>
177          <ims:content ims:content-transfer-encoding="base64">
178            c29tZSBKUEVHIHN0dWZmIGhlcmUu
179          </ims:content>
180        </ims:sign>
181      </ims:signs>
182    </ims:annotation>
183  </ims:fast>
```

Annotation `a1` belongs to the namespace `http://fast.dei.unipd.it/` (line 4) and it is a shared annotation (line 5). In particular, `group-1` cannot access the annotation (lines 12-17), `group-2` can only read the annotation (lines 18-23), and `group-3` can read and modify the annotation (lines 24-29). If annotation `a1` was public or private, the `sharings` elements (lines 11-30) would be not present.

Then the list of the signs materializing the annotation and its semantics follow.

`sign-1` (lines 32-58) is a textual sign written in English – MIME media type `text/plain` and language `eng` (line 33) – and it is a `comment` (line 37) on the annotated digital object `doc1` (lines 39-54) anchored to a specific part of it (lines 43-52): the `pointer` element uses an hypothetical XML-based syntax for selecting with its own elements (lines 44-51) an image within the PDF document which constitutes `doc1` and a specific area of that image; it also specifies that the anchored part has the `image/jpeg` MIME media type (line 43). Finally, the `content` element (lines 55-57) contains the actual content of the sign, as plain text since this is the MIME media type of the sign in this case.

`sign-2` (lines 59-78) is quite similar to `sign-1` with the difference that it shows how a plain text based syntax can be used for anchoring to a part of a digital object (lines 70-72).

`sign-3` (lines 79-107) represents the case of a sign with HTML content – see the MIME media type (line 80) and the HTML fragment embedded in the `content` element (lines 98-106). Note that if a sign carries a content which has an XML-like syntax, e.g. XML, HTML, *Resource Description Framework (RDF)* [W3C, 2004a], this content is represented as sub-elements of the `content` element. Moreover, `sign-3` shows the possibility of having a “relate-to link" (lines 92-97) in this case to the whole `doc35` since there is no `pointer` element.

`sign-4` (lines 108-133) and `sign-5` (lines 134-156) provide small variations and combinations of what has been already discussed in the case of the previous signs.

`sign-6` (lines 157-181) represents the case of an image sign whose binary content is Base64-encoded in the `content` elements (lines 177-179). Moreover, it shows how multiple meanings can be associated to a sign, as the `source` and `cover` concepts (lines 161-164) in this case.

3.4 Example of Annotation in JSON

The annotation model previously described can be represented also using *JavaScript Object Notation (JSON)* [Crockford, 2006].

Nevertheless, since the JSON schema [Zyp, 2010] has not been standardized yet and its current draft expired in May 2011, we have not provided a translation of the XSchema into a corresponding JSON schema to avoid relying on a tool that may vary considerably or never get standardized.

Therefore, to be consistent with the XML one, the JSON representation borrows element names from the XSchema with the main difference that attributes of XML elements have been mapped to sub-fields in JSON, since it does not support the notion of attribute of a field.

The following code listing reports how the annotation discussed in Section 3.1 and shown in Figure 2 is represented in JSON and the explanations and considerations made in the previous section hold also in this case.

```
{
    "fast":{
    "annotation":{
     "identifier":"a1",
     "namespace":"http://fast.dei.unipd.it/",
     "scope":"SHARED",
     "created":"2012-07-31T18:42:20.563+02:00",
     "last-modified":"2012-07-31T18:42:20.563+02:00",
     "owner":{
      "user":{
         "identifier":"ferro",
         "namespace":"http://fast.dei.unipd.it/"
      }
     },
     "sharings":[
      {
         "sharing":{
          "group":{
           "identifier":"group-1",
           "namespace":"http://fast.dei.unipd.it/"
          },
          "access-permission":"DENIED"
         }
      },
      {
         "sharing":{
          "group":{
           "identifier":"group-2",
           "namespace":"http://fast.dei.unipd.it/"
          },
          "access-permission":"READ_ONLY"
         }
      },
      {
         "sharing":{
          "group":{
           "identifier":"group-3",
           "namespace":"http://fast.dei.unipd.it/"
          },
          "access-permission":"READ_WRITE"
         }
      }
     ],
     "signs":[
      {
         "sign":{
          "identifier":"sign-1",
          "media-type":"text/plain",
          "language":"eng",
          "created":"2012-07-31T18:42:20.556+02:00",
          "last-modified":"2012-07-31T18:42:20.556+02:00",
          "concepts":[
           {
            "concept":{
               "identifier":"comment",
               "namespace":"http://fast.dei.unipd.it/"
            }
           }
          ],
          "annotate":{
```

```
            "anchor":{
             "digital-object":{
                "identifier":"doc1",
                "namespace":"http://www.dl1.org/",
                "media-type":"application/pdf"
             },
             "anchored-media-type":"image/jpeg",
             "pointer":"<selection><path>/root/page1/image1</path><area>10 5 20 10</area></selection>"
            }
           },
           "content":{
            "content":"This is a common picture for this novel."
           }
          }
       },
       {
          "sign":{
           "identifier":"sign-2",
           "media-type":"text/plain",
           "language":"eng",
           "created":"2012-07-31T18:42:20.560+02:00",
           "last-modified":"2012-07-31T18:42:20.560+02:00",
           "concepts":[
            {
             "concept":{
                "identifier":"comment",
                "namespace":"http://fast.dei.unipd.it/"
             }
            }
           ],
           "annotate":{
            "anchor":{
             "digital-object":{
                "identifier":"doc1",
                "namespace":"http://www.dl1.org/",
                "media-type":"application/pdf"
             },
             "anchored-media-type":"text/plain",
             "pointer":"/root/page1/sentence3[10-30]"
            }
           },
           "content":{
            "content":"Description of the picture."
           }
          }
       },
       {
          "sign":{
           "identifier":"sign-3",
           "media-type":"text/html",
           "language":"eng",
           "created":"2012-07-31T18:42:20.560+02:00",
           "last-modified":"2012-07-31T18:42:20.560+02:00",
           "concepts":[
            {
             "concept":{
                "identifier":"translation",
                "namespace":"http://fast.dei.unipd.it/"
             }
            }
```

```
        ],
        "annotate":{
         "anchor":{
          "digital-object":{
             "identifier":"doc1",
             "namespace":"http://www.dl1.org/",
             "media-type":"application/pdf"
          }
         }
        },
        "relate-to":{
         "anchor":{
          "digital-object":{
             "identifier":"doc35",
             "namespace":"http://www.dl2.org/",
             "media-type":"text/rtf"
          }
         }
        },
        "content":{
         "content":"<html><body><p>Good Italian translation.</p></body></html>"
        }
       }
    },
    {
       "sign":{
        "identifier":"sign-4",
        "media-type":"text/plain",
        "language":"eng",
        "created":"2012-07-31T18:42:20.560+02:00",
        "last-modified":"2012-07-31T18:42:20.560+02:00",
        "concepts":[
         {
          "concept":{
             "identifier":"review",
             "namespace":"http://fast.dei.unipd.it/"
          }
         }
        ],
        "annotate":{
         "anchor":{
          "digital-object":{
             "identifier":"doc1",
             "namespace":"http://www.dl1.org/",
             "media-type":"application/pdf"
          }
         }
        },
        "relate-to":{
         "anchor":{
          "digital-object":{
             "identifier":"doc35",
             "namespace":"http://www.dl2.org/",
             "media-type":"text/rtf"
          },
          "anchored-media-type":"text/rtf",
          "pointer":"100 0 27"
         }
        },
        "content":{
```

```
            "content":"Verify the syntax of the sentence."
           }
          }
       },
       {
          "sign":{
           "identifier":"sign-5",
           "media-type":"text/plain",
           "language":"eng",
           "created":"2012-07-31T18:42:20.560+02:00",
           "last-modified":"2012-07-31T18:42:20.560+02:00",
           "concepts":[
            {
             "concept":{
                "identifier":"publisher",
                "namespace":"http://purl.org/dc/elements/1.1/"
             }
            }
           ],
           "annotate":{
            "anchor":{
             "digital-object":{
                "identifier":"doc1",
                "namespace":"http://www.dl1.org/",
                "media-type":"application/pdf"
             }
            }
           },
           "relate-to":{
            "anchor":{
             "digital-object":{
                "identifier":"http://www.publisher.com/",
                "namespace":"http://fast.dei.unipd.it/web/",
                "media-type":"text/html"
             }
            }
           },
           "content":{
            "content":"The Publisher, Inc., NY, USA."
           }
          }
       },
       {
          "sign":{
           "identifier":"sign-6",
           "media-type":"image/jpeg",
           "created":"2012-07-31T18:42:20.560+02:00",
           "last-modified":"2012-07-31T18:42:20.560+02:00",
           "concepts":[
            {
             "concept":{
                "identifier":"cover",
                "namespace":"http://fast.dei.unipd.it/"
             }
            },
            {
             "concept":{
                "identifier":"source",
                "namespace":"http://purl.org/dc/elements/1.1/"
             }
```

```
241              }
242             ],
243             "annotate":{
244              "anchor":{
245               "digital-object":{
246                  "identifier":"doc1",
247                  "namespace":"http://www.dl1.org/",
248                  "media-type":"application/pdf"
249               }
250              }
251             },
252             "relate-to":{
253              "anchor":{
254               "digital-object":{
255                  "identifier":"http://www.publisher.com/",
256                  "namespace":"http://fast.dei.unipd.it/web/",
257                  "media-type":"text/html"
258               }
259              }
260             },
261             "content":{
262              "content-transfer-encoding":"base64",
263              "content":"c29tZSBKUEVHIHN0dWZmIGhlcmUu"
264             }
265            }
266         }
267        ]
268       }
269      }
270 }
```

4 The Document-Annotation Hypertext

As explained in Section 3, we consider that existing digital objects and annotations constitute a hypertext. The definition and the properties of this hypertext could be derived directly from the definition of annotation, as done in [Agosti and Ferro, 2008]. However, here we prefer to provide a definition of the document-annotation hypertext which does not need to rely on a previous formal definition of annotation still keeping all the properties which exhibits when formally defined. To this end, we adopt the definition introduced in [Agosti and Ferro, 2005].

Definition 1: Let A be the set of annotations, D the set of documents, and $DO = A \cup D$ the set of digital objects, which are either annotations or documents.

The **document–annotation hypertext** is a labeled directed graph

$$H_{da} = (DO, E_{da} \subseteq A \times DO)$$

where DO is the set of vertices and E_{da} is the set of edges. Let $\mathrm{l}_{da} : E_{da} \rightarrow LT$ be the labelling function. For each $e = (a, do) \in E_{da}$ there is a $\mathrm{l}_{da}(e)$-labeled edge from the annotation a to the generic digital object do, specifying the type of link, either *Annotate* or *Relate-to*. The following constraints must be satisfied:

1. each annotation a must annotate one and only one digital object[4]:

$$\forall a \in A \; \exists! \, e = (a, do) \in E_{da} \mid \mathrm{l}_{da}(e) = \textit{Annotate}$$

2. the graph does not contain loops:

$$\forall a \in A \; \nexists \, e = (a, do) \in E_{da} \mid a = do$$

3. the graph does not contain cycles:

$$\begin{aligned} &\nexists \, C = a_0 a_k a_{k-1} \cdots a_1 a_0 \mid \\ &\quad e_0 = (a_0, a_k), e_k = (a_k, a_{k-1}), \ldots, e_1 = (a_1, a_0) \in E_{da}, \\ &\quad \mathrm{l}_{da}(e_0) = \mathrm{l}_{da}(e_k) = \ldots = \mathrm{l}_{da}(e_1) = \textit{Annotate} \end{aligned}$$

Note that each $e \in E_{da}$ always starts from an annotation, while $e \in E_{da}$ that starts from a document does not exist. Each annotation is constrained to be incident with one and only one edge with link type “Annotate", thus formalizing the notion of link type mentioned above. The constraint related to loops prevent us from creating self-referencing annotations, which have no use for our purposes. Finally, annotations involve a temporal dimension, since each annotation has to annotate an already existing digital object. Thus, the last constraint about cycles of annotations prevents us from creating cycles where the oldest annotation a_0 annotates the newest annotation a_k; note that this is not an issue for document vertices, since “Annotate" links can start only from annotations.

Moreover, as demonstrated in [Agosti and Ferro, 2005, 2008], for each document there is a unique tree of annotations constituted by “Annotate" edges that can be rooted in the document. Remembering that in a tree any two given vertices are linked by a unique path, it follows that for each annotation we can determine the unique path to the root document of the tree to which the annotation belongs

Figure 4 shows an example of document-annotation hypertext. In the example, there are two digital libraries: one concerning illuminated manuscripts, whose namespace is `imDL`, and the other concerning computer science papers, whose namespace is csDL. The document-annotation hypertext allows us to express:

- *annotation sets concerning a document*: `fast:a1`, `fast:a2` is an annotation set concerning the document `imDL:doc10`;
- *annotation sets concerning an annotation*: `fast:a7`, `fast:a13` is an annotation set concerning the annotation `fast:a6`;
- *annotation threads concerning a document* : `fast:a1`, `fast:a4`, `fast:a5` is an annotation thread concerning the document `imDL:doc10`;

[4] $\exists!$ is the *unique existential quantifier*, and it is read “there exists a unique . . . such that . . .”.

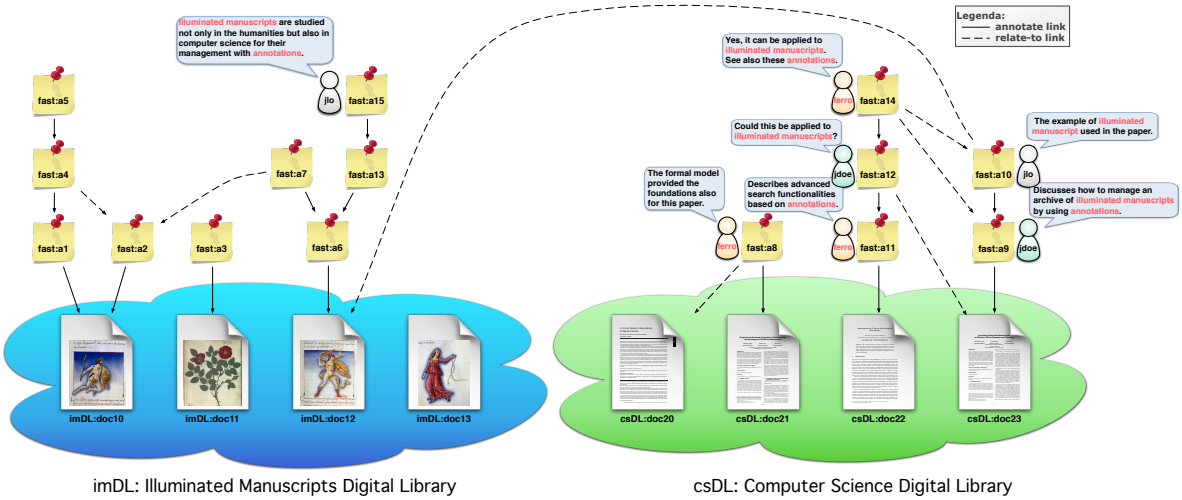

Figure 4: Example of document–annotation hypertext H_{da}.

- *annotation threads concerning an annotation*: `fast:a13`, `fast:a15` is an annotation thread concerning the annotation `fast:a6`.

These basic patterns can be mixed to obtain nested sets and threads of annotations.

As you can note, by following the “Annotate" links, each annotation results rooted in a unique document which is the root of the tree. On the other hand, “Relate-to" links allow annotations to provide alternative browsing and navigation paths which can connect different digital objects. Finally, annotations may involve documents that belong to different digital libraries, as in the case of annotation `fast:a10` which annotates the annotations `fast:a9`, it is rooted in the document `csDL:doc23` which belongs to the computer science digital library, and related to document `imDL:doc12`, which belongs to the illuminated manuscripts digital library.

5 Search Model

The FAST annotation service extends the overall search model of the PROMISE evaluation infrastructure, described in Section 4 of D3.3 [Agosti et al., 2012]. Here, only the model for dealing with the document-annotation hypertext is reported.

The hypertext that connects documents to annotations calls for a search strategy that takes it into consideration and allows us to modify the score of annotations and/or documents according to the paths in the hypertext. For example, we could consider that an annotation, retrieved in response to a user query, is more relevant if it is part of a thread where other annotations have also been retrieved in response to the same query rather than if it is part of a thread where it is the only annotation that matches the query. This is discussed in Section 5.1

5.1 Annotation Hypertext-driven Retrieval

Consider the document-annotation hypertext $H_{da} = (DO, E)$ where DO is a set of digital objects (either documents or annotations) and E is a set of edges indicating that an annotation is annotating a digital object, as introduced in [Agosti and Ferro, 2005, 2008] and previously discussed.

The **hypertext similarity score** between an annotation and a query is defined as:

$$\mathsf{sim}^{ht}_{\alpha}(a, q) = \frac{1}{\alpha}\mathsf{sim}(a, q) + \frac{\alpha - 1}{\alpha} \cdot \frac{1}{|\mathrm{succ}(a)|} \sum_{a_k \in \mathrm{succ}(a)} \frac{\mathsf{sim}(a_k, q) + \mathsf{sim}^{ht}_{\alpha}(a_k, q)}{2}$$

where $\mathsf{sim}(a, q) \in [0, 1]$ is a generic similarity function between an annotation and a query, $\mathrm{succ}(a)$ is a function that returns the set of successors of an annotation a_j and α is a real number called the *annotation thread damping* factor. We consider that $\mathsf{sim}(a_j, q) = 0$ for those annotations that do not match the query.

$\mathsf{sim}^{ht}_{\alpha}(a, q)$ computes the weighted average between $\mathsf{sim}(a, q)$, the similarity score of an annotation with respect to a query, and the similarity scores which come from the thread to which the annotation belongs. In particular, the thread similarity scores are given by the average between the similarity scores of the successors of a and the hypertext similarity scores of the successors of a; in other words, the hypertext similarity score recursively averages the similarity scores of the annotations that belong to the same thread of the given annotation a. Furthermore, $\mathsf{sim}^{ht}_{\alpha}(a, q)$ penalizes similarity scores which come from lengthy paths, because for a path $P = a_0 \ldots a_k$ the similarity score $\mathsf{sim}(a_k, q)$ of a_k is weighted $\frac{1}{2^k}$.

The hypertext similarity score resembles the CombANZ strategy of [Fox and Shaw, 1993], proposing a recursive version of this strategy, even though CombANZ averages only on non-zero similarity scores. Examples of functions similar to can be found in [Frommholz et al., 2003; Savoy, 1996], but [Frommholz et al., 2003] exploits a probabilistic framework and chooses the path with the maximum probability of the relevance of a document, while [Savoy, 1996] does not average the similarity scores and has an iterative approach to the problem.

By varying the value of α between 0 and ∞, it is possible to obtain a query processing intermediate between a traditional information retrieval model ($\alpha = 1$), when $\mathsf{sim}^{ht}_{1}(a, q) = \mathsf{sim}(a, q)$ and only the similarity between the annotation and the query is taken into account, and a pure hypertext driven retrieval model ($\alpha = \infty$), when $\mathsf{sim}^{ht}_{\infty}(a, q) = \frac{1}{|\mathrm{succ}(a)|} \sum_{a_k \in \mathrm{succ}(a)} \frac{\mathsf{sim}(a_k, q) + \mathsf{sim}^{ht}_{\infty}(a_k, q)}{2}$ and only the thread to which the annotation belongs is taken into account.

Note that the above definition of the hypertext similarity score provides us with an additional degree of freedom in the choice of the actual function to be used for computing the similarity between a term and the annotation, which allows us to plug in different retrieval models for annotations besides the one based on extended boolean retrieval proposed in Section 4 of D3.3 [Agosti et al., 2012].

Finally, the hypertext-driven retrieval model allows us to compute a similarity score also for the documents that have been annotated, so that it is possible to search and retrieve documents in response to a user query by means of their annotations. The **similarity score by annotation**

between the document and a query is defined as:

$$\text{sim}^{a}_{\alpha}(d,q) = \frac{1}{|\text{succ}(d)|} \sum_{a \in \text{succ}(d)} \text{sim}^{ht}_{\alpha}(a,q)$$

Basically, the similarity score by annotation of a document averages the hypertext similarity scores of the annotations that are annotating the document.

6 FAST RESTful WebService

The FAST annotation service extends the RESTful Web service of the PROMISE evaluation infrastructure, described in Section 7 of D3.3 [Agosti et al., 2012]. Here, only the resources concerned annotations are reported.

6.1 Annotation Resource

6.1.1 API

Action	HTTP Method	URI
CREATE_ANNOTATION	POST	`/annotation`
READ_ANNOTATION	GET	`/annotation/{id};{ns}`
READ_ANNOTATION	GET	`/annotation/{id};{ns}/content`
READ_ANNOTATION_TREE_ROOT	GET	`/annotation/{id};{ns}/tree-root`
UPDATE_ANNOTATION	PUT	`/annotation/{id};{ns}`
DELETE_ANNOTATION	DELETE	`/annotation/{id};{ns}`
LIST_ANNOTATIONS	GET	`/annotation`
LIST_ANNOTATION_PROVENANCE_EVENTS	GET	`/annotation/{id};{ns}/provenance`
SHARE_ANNOTATION	GET, PUT, POST	`/annotation/{id};{ns}/share/{sharer-id};{sharer-ns}/permission/{access-permission}`
UNSHARE_ANNOTATION	DELETE	`/annotation/{id};{ns}/share/{sharer-id};{sharer-ns}`
ADD_SIGN_TO_ANNOTATION	POST	`/annotation/{id};{ns}/sign`
READ_SIGN	GET	`/annotation/{id};{ns}/sign/{sid}`

Action	HTTP Method	URI
UPDATE_SIGN	PUT	`/annotation/{id};{ns}/sign/{sid}`
DELETE_SIGN	DELETE	`/annotation/{id};{ns}/sign/{sid}`
READ_SIGN	GET	`/annotation/{id};{ns}/sign/{sid}/content`
UPDATE_SIGN	POST, PUT	`/annotation/{id};{ns}/sign/{sid}/content`
UPDATE_SIGN	DELETE	`/annotation/{id};{ns}/sign/{sid}/content`

Table 1: API for accessing the annotation resource.

where `{id}` is the unique identifier of the annotation and `{ns}` is the namespace to which the annotation belongs; `{shared-id}` and `{sharer-ns}` are the identifier and namespace of the group which shares the annotation; `{access-permission}` is the access permission to the annotation for the group, namely `DENIED`, `READ_ONLY`, `READ_WRITE`; `{sid}` is the identifier of a sign.

6.1.2 XML Representation

See example in Section 3.3.

6.1.3 JSON Representation

See example in Section 3.4.

6.1.4 Basic Usage

The basic usage is intended for the cases in which the annotation consists mainly of textual signs or reasonably small size size, as in the case of the annotation discussed in Section 3.3. This means that the annotation can be processed as a whole.

Creation To create the annotation, you have to POST at the following URI:

`/annotation`

Note that with respect to the example of Section 3.4, you may decide to not specify the identifier for the annotation and for the signs contained in the annotation. In that case the system will automatically set a type 4 *Universally Unique IDentifier (UUID)* [ISO/IEC 9834-8, 2008; Leach et al., 2005] for the annotation and each of its signs.

Note that, even if you create a shared annotation as in the case of the example in Section 3.3, this is not automatically shared with the groups at creation time but, for each group you want to share the annotation with, you need a separate call to the following URI:

/annotation/{id};{ns}/share/{sharer-id};{sharer-ns}/permission/{access-permission}

Reading To read a whole annotation, you have to GET the following URI:

/annotation/{id};{ns}/content

Update To update the scope and/or the owner of an annotation, you have to PUT at the following URI:

/annotation/{id};{ns}

To update one of the signs of the annotation, you have to PUT at the following URI:

/annotation/{id};{ns}/sign/{sid}

Deletion To delete a whole annotation, you have to DELETE the following URI:

/annotation/{id};{ns}

6.1.5 Advanced Usage

The advanced use stems from the consideration that the payload of a sign can be of considerable size, e.g. in the case of images or sounds, and so a more efficient way of dealing with this use annotation is needed. The basic idea here is to allow to create, read, update, and delete the annotation and signs separately and, when possible, to avoid the XML or JSON wrapping.

Creation To create the annotation, without any sign, you have to POST at the following URI:

/annotation

an annotation where its representation does not contain any signs, as in the XML example below.

```
<?xml version="1.0" encoding="UTF-8"?>
<ims:fast xmlns:ims="http://ims.dei.unipd.it/" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://ims.dei.unipd.it/ http://ims.dei.unipd.it/data/xml/fast.3.00.xsd">
  <ims:annotation ims:identifier="a1" ims:namespace="http://fast.dei.unipd.it/"
    ims:scope="SHARED"
    ims:created="2012-07-31T18:42:20.563+02:00"
    ims:last-modified="2012-07-31T18:42:20.563+02:00">
    <ims:owner>
      <ims:user ims:identifier="ferro" ims:namespace="http://fast.dei.unipd.it/" />
    </ims:owner>
  </ims:annotation>
</ims:fast>
```

Then, to add each sign separately to the annotation, you have to POST at the following URI:

/annotation/{id};{ns}/sign

a representation of a sign, as in the XML example below.

```
<?xml version="1.0" encoding="UTF-8"?>
<ims:fast xmlns:ims="http://ims.dei.unipd.it/"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://ims.dei.unipd.it/ http://ims.dei.unipd.it/data/xml/fast.3.00.xsd">
  <ims:sign ims:identifier="sign-6"
    ims:media-type="image/jpeg"
    ims:created="2012-07-31T18:42:20.560+02:00"
    ims:last-modified="2012-07-31T18:42:20.560+02:00">
    <ims:concepts>
      <ims:concept ims:identifier="cover" ims:namespace="http://fast.dei.unipd.it/" />
      <ims:concept ims:identifier="source" ims:namespace="http://purl.org/dc/elements/1.1/" />
    </ims:concepts>
    <ims:annotate>
      <ims:anchor>
        <ims:digital-object ims:identifier="doc1"
          ims:namespace="http://www.dl1.org/"
            ims:media-type="application/pdf" />
      </ims:anchor>
    </ims:annotate>
    <ims:relate-to>
      <ims:anchor>
        <ims:digital-object ims:identifier="http://www.publisher.com/"
          ims:namespace="http://fast.dei.unipd.it/web/"
          ims:media-type="text/html" />
      </ims:anchor>
    </ims:relate-to>
    <ims:content ims:content-transfer-encoding="base64">
      c29tZSBKUEVHIHN0dWZmIGhlcmUu
    </ims:content>
  </ims:sign>
</ims:fast>
```

If the payload of your sign is really huge, you may decide to add the sign to the annotation via a POST at the following URI:

/annotation/{id};{ns}/sign

using representation of a sign without the `content` element, as in the XML example below.

```
<?xml version="1.0" encoding="UTF-8"?>
<ims:fast xmlns:ims="http://ims.dei.unipd.it/"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://ims.dei.unipd.it/ http://ims.dei.unipd.it/data/xml/fast.3.00.xsd">
  <ims:sign ims:identifier="sign-6"
    ims:media-type="image/jpeg"
    ims:created="2012-07-31T18:42:20.560+02:00"
    ims:last-modified="2012-07-31T18:42:20.560+02:00">
    <ims:concepts>
      <ims:concept ims:identifier="cover" ims:namespace="http://fast.dei.unipd.it/" />
      <ims:concept ims:identifier="source" ims:namespace="http://purl.org/dc/elements/1.1/" />
    </ims:concepts>
    <ims:annotate>
      <ims:anchor>
        <ims:digital-object ims:identifier="doc1"
          ims:namespace="http://www.dl1.org/"
          ims:media-type="application/pdf" />
      </ims:anchor>
    </ims:annotate>
    <ims:relate-to>
      <ims:anchor>
```

```
22          <ims:digital-object ims:identifier="http://www.publisher.com/"
23            ims:namespace="http://fast.dei.unipd.it/web/"
24            ims:media-type="text/html" />
25        </ims:anchor>
26      </ims:relate-to>
27    </ims:sign>
28 </ims:fast>
```

and then upload the actual payload of the sign, directly in binary format and specifying its MIME media type in the *HyperText Transfer Protocol (HTTP)* headers, via a PUT or POST at the following URI:

/annotation/{id};{ns}/sign/{sid}/content

Reading To read an annotation, you have to GET the following URI:

/annotation/{id};{ns}

and it will return a representation of the annotation where all its signs are listed without reporting their content, as in the XML example below.

```
1  <?xml version="1.0" encoding="UTF-8"?>
2  <ims:fast xmlns:ims="http://ims.dei.unipd.it/" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
3    xsi:schemaLocation="http://ims.dei.unipd.it/ http://ims.dei.unipd.it/data/xml/fast.3.00.xsd">
4    <ims:annotation ims:identifier="a1" ims:namespace="http://fast.dei.unipd.it/"
5      ims:scope="SHARED"
6      ims:created="2012-07-31T18:42:20.563+02:00"
7      ims:last-modified="2012-07-31T18:42:20.563+02:00">
8      <ims:owner>
9        <ims:user ims:identifier="ferro" ims:namespace="http://fast.dei.unipd.it/" />
10     </ims:owner>
11     <ims:sharings>
12       <ims:sharing>
13         <ims:group ims:identifier="group-1" ims:namespace="http://fast.dei.unipd.it/" />
14         <ims:access-permission>
15           DENIED
16         </ims:access-permission>
17       </ims:sharing>
18       <ims:sharing>
19         <ims:group ims:identifier="group-2" ims:namespace="http://fast.dei.unipd.it/" />
20         <ims:access-permission>
21           READ_ONLY
22         </ims:access-permission>
23       </ims:sharing>
24       <ims:sharing>
25         <ims:group ims:identifier="group-3" ims:namespace="http://fast.dei.unipd.it/" />
26         <ims:access-permission>
27           READ_WRITE
28         </ims:access-permission>
29       </ims:sharing>
30     </ims:sharings>
31     <ims:signs>
32       <ims:sign ims:identifier="sign-1"
33         ims:media-type="text/plain" ims:language="eng"
34         ims:created="2012-07-31T18:42:20.556+02:00"
35         ims:last-modified="2012-07-31T18:42:20.556+02:00">
36         <ims:concepts>
37           <ims:concept ims:identifier="comment" ims:namespace="http://fast.dei.unipd.it/" />
```

```
38          </ims:concepts>
39          <ims:annotate>
40            <ims:anchor>
41              <ims:digital-object ims:identifier="doc1" ims:namespace="http://www.dl1.org/"
42                ims:media-type="application/pdf" />
43              <ims:pointer ims:anchored-media-type="image/jpeg">
44                <selection>
45                  <path>
46                    /root/page1/image1
47                  </path>
48                  <area>
49                    10 5 20 10
50                  </area>
51                </selection>
52              </ims:pointer>
53            </ims:anchor>
54          </ims:annotate>
55        </ims:sign>
56        <ims:sign ims:identifier="sign-2"
57          ims:media-type="text/plain" ims:language="eng"
58          ims:created="2012-07-31T18:42:20.560+02:00"
59          ims:last-modified="2012-07-31T18:42:20.560+02:00">
60          <ims:concepts>
61            <ims:concept ims:identifier="comment" ims:namespace="http://fast.dei.unipd.it/" />
62          </ims:concepts>
63          <ims:annotate>
64            <ims:anchor>
65              <ims:digital-object ims:identifier="doc1" ims:namespace="http://www.dl1.org/"
66                ims:media-type="application/pdf" />
67              <ims:pointer ims:anchored-media-type="text/plain">
68                /root/page1/sentence3[10-30]
69              </ims:pointer>
70            </ims:anchor>
71          </ims:annotate>
72        </ims:sign>
73        <ims:sign ims:identifier="sign-3"
74          ims:media-type="text/html" ims:language="eng"
75          ims:created="2012-07-31T18:42:20.560+02:00"
76          ims:last-modified="2012-07-31T18:42:20.560+02:00">
77          <ims:concepts>
78            <ims:concept ims:identifier="translation" ims:namespace="http://fast.dei.unipd.it/" />
79          </ims:concepts>
80          <ims:annotate>
81            <ims:anchor>
82              <ims:digital-object ims:identifier="doc1" ims:namespace="http://www.dl1.org/"
83                ims:media-type="application/pdf" />
84            </ims:anchor>
85          </ims:annotate>
86          <ims:relate-to>
87            <ims:anchor>
88              <ims:digital-object ims:identifier="doc35" ims:namespace="http://www.dl2.org/"
89                ims:media-type="text/rtf" />
90            </ims:anchor>
91          </ims:relate-to>
92        </ims:sign>
93        <ims:sign ims:identifier="sign-4"
94          ims:media-type="text/plain" ims:language="eng"
95          ims:created="2012-07-31T18:42:20.560+02:00"
96          ims:last-modified="2012-07-31T18:42:20.560+02:00">
97          <ims:concepts>
```

```
98            <ims:concept ims:identifier="review" ims:namespace="http://fast.dei.unipd.it/" />
99          </ims:concepts>
100         <ims:annotate>
101           <ims:anchor>
102             <ims:digital-object ims:identifier="doc1" ims:namespace="http://www.dl1.org/"
103               ims:media-type="application/pdf" />
104           </ims:anchor>
105         </ims:annotate>
106         <ims:relate-to>
107           <ims:anchor>
108             <ims:digital-object ims:identifier="doc35" ims:namespace="http://www.dl2.org/"
109               ims:media-type="text/rtf" />
110             <ims:pointer ims:anchored-media-type="text/rtf">
111               100 0 27
112             </ims:pointer>
113           </ims:anchor>
114         </ims:relate-to>
115       </ims:sign>
116       <ims:sign ims:identifier="sign-5"
117         ims:media-type="text/plain" ims:language="eng"
118         ims:created="2012-07-31T18:42:20.560+02:00"
119         ims:last-modified="2012-07-31T18:42:20.560+02:00">
120         <ims:concepts>
121           <ims:concept ims:identifier="publisher" ims:namespace="http://purl.org/dc/elements/1.1/" />
122         </ims:concepts>
123         <ims:annotate>
124           <ims:anchor>
125             <ims:digital-object ims:identifier="doc1" ims:namespace="http://www.dl1.org/"
126               ims:media-type="application/pdf" />
127           </ims:anchor>
128         </ims:annotate>
129         <ims:relate-to>
130           <ims:anchor>
131             <ims:digital-object ims:identifier="http://www.publisher.com/"
132               ims:namespace="http://fast.dei.unipd.it/web/" ims:media-type="text/html" />
133           </ims:anchor>
134         </ims:relate-to>
135       </ims:sign>
136       <ims:sign ims:identifier="sign-6"
137         ims:media-type="image/jpeg"
138         ims:created="2012-07-31T18:42:20.560+02:00"
139         ims:last-modified="2012-07-31T18:42:20.560+02:00">
140         <ims:concepts>
141           <ims:concept ims:identifier="cover" ims:namespace="http://fast.dei.unipd.it/" />
142           <ims:concept ims:identifier="source" ims:namespace="http://purl.org/dc/elements/1.1/" />
143         </ims:concepts>
144         <ims:annotate>
145           <ims:anchor>
146             <ims:digital-object ims:identifier="doc1" ims:namespace="http://www.dl1.org/"
147               ims:media-type="application/pdf" />
148           </ims:anchor>
149         </ims:annotate>
150         <ims:relate-to>
151           <ims:anchor>
152             <ims:digital-object ims:identifier="http://www.publisher.com/"
153               ims:namespace="http://fast.dei.unipd.it/web/" ims:media-type="text/html" />
154           </ims:anchor>
155         </ims:relate-to>
156       </ims:sign>
157     </ims:signs>
```

```
158   </ims:annotation>
159 </ims:fast>
```

Then, you can read the actual content of each sign separately via a GET at the following URI:

`/annotation/{id};{ns}/sign/{sid}/content`

Update To update the scope and/or the owner of an annotation, you have to PUT at the following URI:

`/annotation/{id};{ns}`

To update one of the signs of the annotation but not its content, you have to PUT at the following URI:

`/annotation/{id};{ns}/sign/{sid}`

a representation of the sign where the `content` element is not present, as in the XML example below.

```
1  <?xml version="1.0" encoding="UTF-8"?>
2  <ims:fast xmlns:ims="http://ims.dei.unipd.it/"
3    xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
4    xsi:schemaLocation="http://ims.dei.unipd.it/ http://ims.dei.unipd.it/data/xml/fast.3.00.xsd">
5    <ims:sign ims:identifier="sign-6"
6      ims:media-type="image/jpeg"
7      ims:created="2012-07-31T18:42:20.560+02:00"
8      ims:last-modified="2012-07-31T18:42:20.560+02:00">
9      <ims:concepts>
10       <ims:concept ims:identifier="cover" ims:namespace="http://fast.dei.unipd.it/" />
11       <ims:concept ims:identifier="source" ims:namespace="http://purl.org/dc/elements/1.1/" />
12     </ims:concepts>
13     <ims:annotate>
14       <ims:anchor>
15         <ims:digital-object ims:identifier="doc1"
16           ims:namespace="http://www.dl1.org/"
17           ims:media-type="application/pdf" />
18       </ims:anchor>
19     </ims:annotate>
20     <ims:relate-to>
21       <ims:anchor>
22         <ims:digital-object ims:identifier="http://www.publisher.com/"
23           ims:namespace="http://fast.dei.unipd.it/web/"
24           ims:media-type="text/html" />
25       </ims:anchor>
26     </ims:relate-to>
27   </ims:sign>
28 </ims:fast>
```

To update the content of one of the signs of the annotation, you have to POST or PUT at the following URI:

`/annotation/{id};{ns}/sign/{sid}/content`

the actual payload of the sign, directly in binary format and specifying its MIME media type in the HTTP headers.

If you perform a DELETE on the same URI, you will empty the content of the sign but not delete the sign itself.

Deletion To delete a sign of an annotation, you have to DELETE the following URI:

`/annotation/{id};{ns}/sign/{sid}`

To delete a whole annotation, you have to DELETE the following URI:

`/annotation/{id};{ns}`

7 The CQL Context Set

The FAST annotation service extends the CQL Context Set of the PROMISE evaluation infrastructure, described in Section 8 of D3.3 [Agosti et al., 2012]. Here, only the indexes and modifiers concerned annotations are reported.

7.1 Indexes

This section describes the indexes available in the FAST context set for searching and accessing the different resources managed by the system.

7.1.1 Annotation Indexes

Index Name	Alias	Type	Relations	Wildcards	Description
fast.annotation.identifier	fast.r.id	Exact Match	=, ==, <>	* and ?	Matches annotations with respect to their identifier
fast.annotation.namespace.identifier	fast.r.ns.id	Exact Match	=, ==, <>	* and ?	Matches annotations with respect to the identifier of their namespace
fast.annotation.namespace.prefix	fast.r.ns.prefix	Exact Match	=, ==, <>	* and ?	Matches annotations with respect to the prefix of their namespace
fast.annotation.mediaType	fast.a.mt	Exact Match	=, ==, <>	* and ?	Matches annotations with respect to the MIME media type of their signs

Index Name	Alias	Type	Relations	Wildcards	Description
fast.annotation.language	fast.a.lang	Exact Match	=, ==, <>	* and ?	Matches annotations with respect to language of their signs expressed using ISO 639-2:1998 three letters codes
fast.annotation.scope	fast.a.scope	Exact Match	=, ==, <>, >, <, >=, <=	No	Matches annotations with respect to their scope
fast.annotation.general	fast.a.gen	Best Match	=, ==, <>	*	Matches annotations with respect the textual content of their signs
fast.annotation.created	fast.a.created	Exact Match	=, ==, <>, >, <, >=, <=	No	Matches annotations with respect to their creation timestamp
fast.annotation.lastModified	fast.a.lastModified	Exact Match	=, ==, <>, >, <, >=, <=	No	Matches annotations with respect to their last modification timestamp
fast.annotation.owner.identifier	fast.a.owner.id	Exact Match	=, ==, <>	* and ?	Matches annotations with respect to the identifier of their owner
fast.annotation.owner.namespace.identifier	fast.a.owner.ns.id	Exact Match	=, ==, <>	* and ?	Matches annotations with respect to the identifier of the namespace of their owner
fast.annotation.owner.namespace.prefix	fast.a.owner.ns.prefix	Exact Match	=, ==, <>	* and ?	Matches annotations with respect to the prefix of the namespace of their owner

Index Name	Alias	Type	Relations	Wildcards	Description
fast.annotation.accessPermission	fast.a.ap	Exact Match	=, ==, <>, >, <, >=, <=	No	Matches annotations with respect to their access permission
fast.annotation.sharingGroup.identifier	fast.a.sg.id	Exact Match	=, ==, <>	* and ?	Matches annotations with respect to the identifier of their sharing groups
fast.annotation.sharingGroup.namespace.identifier	fast.a.sg.ns.id	Exact Match	=, ==, <>	* and ?	Matches annotations with respect to the identifier of the namespace of their sharing groups
fast.annotation.sharingGroup.namespace.prefix	fast.a.sg.ns.prefix	Exact Match	=, ==, <>	* and ?	Matches annotations with respect to the prefix of the namespace of their sharing groups
fast.annotation.annotatedDigitalObject.identifier	fast.a.ado.id	Exact Match	=, ==, <>	* and ?	Matches annotations with respect to the identifier of their annotated digital object
fast.annotation.annotatedDigitalObject.namespace.identifier	fast.a.ado.ns.id	Exact Match	=, ==, <>	* and ?	Matches annotations with respect to the identifier of the namespace of their annotated digital object
fast.annotation.annotatedDigitalObject.namespace.prefix	fast.a.ado.ns.prefix	Exact Match	=, ==, <>	* and ?	Matches annotations with respect to the prefix of the namespace of their annotated digital object

Index Name	Alias	Type	Relations	Wildcards	Description
fast.annotation.annotatedDigitalObject.mediaType	fast.a.ado.mt	Exact Match	=, ==, <>	* and ?	Matches annotations with respect to the MIME media type of their annotated digital object
fast.annotation.annotatedDigitalObject.language	fast.a.ado.lang	Exact Match	=, ==, <>	* and ?	Matches annotations with respect to the language of their annotated digital object, expressed using ISO 639-2:1998 three letters codes
fast.annotation.relatedDigitalObject.identifier	fast.a.rdo.id	Exact Match	=, ==, <>	* and ?	Matches annotations with respect to the identifier of their related digital object
fast.annotation.relatedDigitalObject.namespace.identifier	fast.a.rdo.ns.id	Exact Match	=, ==, <>	* and ?	Matches annotations with respect to the identifier of the namespace of their related digital object
fast.annotation.relatedDigitalObject.namespace.prefix	fast.a.rdo.ns.prefix	Exact Match	=, ==, <>	* and ?	Matches annotations with respect to the prefix of the namespace of their related digital object
fast.annotation.relatedDigitalObject.mediaType	fast.a.rdo.mt	Exact Match	=, ==, <>	* and ?	Matches annotations with respect to the MIME media type of their related digital object

Index Name	Alias	Type	Relations	Wildcards	Description
fast.annotation.relatedDigitalObject.language	fast.a.rdo.lang	Exact Match	=, ==, <>	* and ?	Matches annotations with respect to the language of their annotated digital object, expressed using ISO 639-2:1998 three letters codes
fast.annotation.concept.identifier	fast.a.c.id	Exact Match	=, ==, <>	* and ?	Matches annotations with respect to the identifier of their concepts
fast.annotation.concept.namespace.identifier	fast.a.c.ns.id	Exact Match	=, ==, <>	* and ?	Matches annotations with respect to the identifier of the namespace of their concepts
fast.annotation.concept.namespace.prefix	fast.a.c.ns.prefix	Exact Match	=, ==, <>	* and ?	Matches annotations with respect to the prefix of the namespace of their concepts

Table 2: Indexes for searching the annotation resource.

7.1.2 Digital Object Indexes

Index Name	Alias	Type	Relations	Wildcards	Description
fast.digitalObject.byAnnotationContent	fast.do.byac	Best Match	=, ==, <>	*	Matches digital objects with respect to the content of their annotations

Table 3: Indexes for searching the digital object resource.

7.2 Relation Modifiers

The following Relation modifier is defined only for the `fast.annotation.general` and `fast.digitalObject.byAnnotationContent` indexes:

- `thread=value`
  The kind of matching to be applied when taking into account the document-annotation hypertext, according to the different match strategies discussed in Section 5.1.
  `value` can assume one of the following values:
  - `no`: No hypertext structure has to be taken into account but only the content of the digital object;
  - `half`: The hypertext structure influences as much as the content of the digital object;
  - `almost`: The hypertext structure influences much more than the content of the digital object;
  - `only`: Only the hypertext structure has to be taken into account an not the content of the digital object.

7.3 Examples

- `fast.annotation.general == giotto`
  Searches for annotations about Giotto.
- `fast.annotation.general ==/thread==half giotto`
  Searches for annotations about Giotto, taking into consideration also the annotations annotating them.
- `fast.annotation.general ==/limit==100 giotto`
  Searches for annotations about Giotto and returns only the first 100 items
- `fast.annotation.general ==/limit==100/offset==50 giotto`
  Searches for annotations about Giotto and returns only 100 items starting after the first 50 items
- `(fast.annotation.general == giotto) and/match==exact (fast.annotation.owner.identifier == ferro)`
  Searches for annotations about Giotto which are authored by the user with identifier `ferro`.
- `(fast.annotation.general == giotto) and/match==fuzzy (fast.annotation.owner.identifier == ferro)`
  Searches for metadata about Giotto which may be authored by the user with identifier `ferro`.
- `fast.digitalObject.byAnnotationContent ==/thread==half giotto`
  Searches for documents whose annotations are about Giotto.
- `fast.annotation.language == ita`
  Searches for annotations written in Italian.
- `ici.user.email == ‘‘ferro@dei.unipd.it"`
  Searches for users whose e-mail address is `ferro@dei.unipd.it`.

- `ici.user.email == *ferro*`
  Searches for users whose e-mail address contains the substring `ferro`, e.g. (`nicola.ferro`, `ferro.nicola`).

- `ici.user.group.identifier == admin*`
  Searches for users who belong to groups whose identifier starts with `admin`.

- `ici.user.country == ITA`
  Searches for Italian users.

- `ici.logEvent.created >/limit==100 2011-05-01`
  Searches for last 100 log events created after 1st May 2011.

- `ici.logEvent.identifier >/limit==100 1`
  Searches for last 100 log events.

8 The Visual Analytics component

This section reports the functionalities of the Visual Analytics component.

Figure 5 schematizes the visual analytics process that combines automatic and visual analysis methods with a tight coupling through human interaction in order to gain knowledge from data. The figure shows an abstract overview of the different stages (represented through ovals) and their transitions (arrows) in the visual analytics process.
The first step is often to preprocess and transform the data to derive different representations for further exploration (as indicated by the Transformation arrow). Other typical preprocessing tasks include data cleaning, normalization, grouping, or integration of heterogeneous data sources.
After the transformation, the analyst may choose between applying visual or automatic analysis methods. Alternating between visual and automatic methods is characteristic for the visual analytics process and leads to a continuous refinement and verification of preliminary results. User interaction with the visualization is needed to reveal insightful information, for instance by zooming in on different data areas or by considering different visual views on the data. In summary, in the visual analytics process, knowledge can be gained from visualization, automatic analysis, as well as the preceding interactions between visualizations, models, and the human analysts.

According to that, before introducing the analytical process, we quickly recall the structure of the PROMISE data.

8.1 PROMISE data

In order to understand how data are organized and displayed it is important to define a typical scenario in which these data are used. In particular, we focus on information retrieval evaluation campaigns. An evaluation campaign is an activity intended to support researchers in information retrieval by providing a large test collection and uniform scoring procedures. Within an evaluation

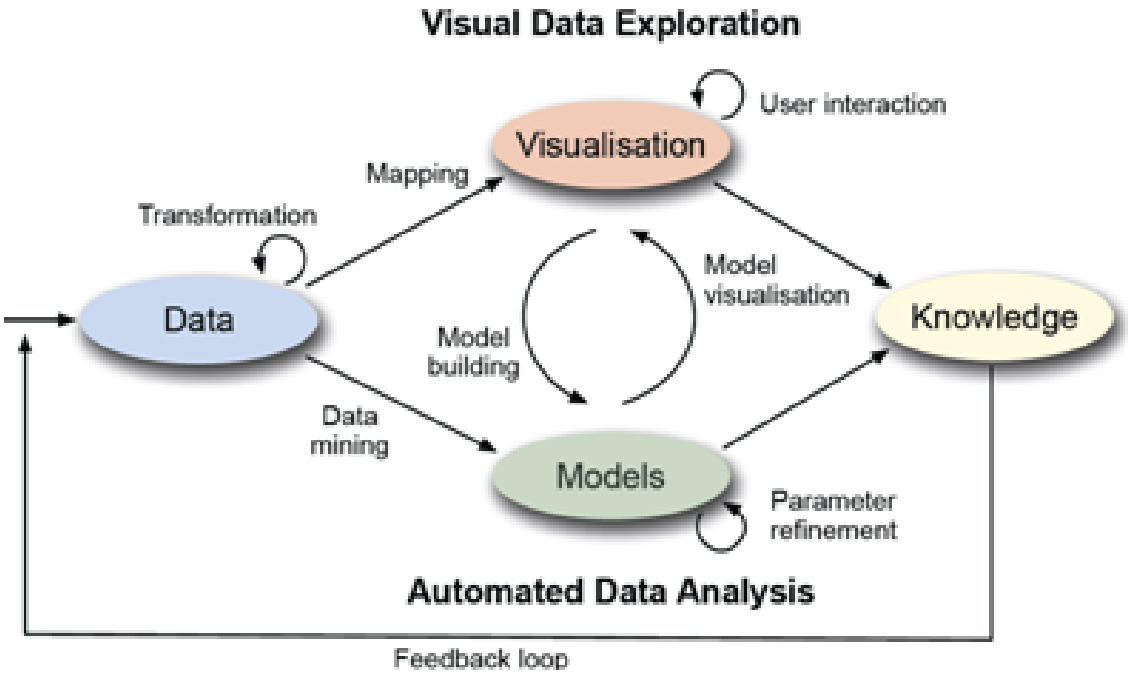

Figure 5: The Visual Analytics process

campaign there are many tracks like multimedia, multilingual, text, images, and so on. A track includes, in turn, several tasks. A task is used to define the experiment structure specifying a set of documents, a set of topics, and a relevance assessment. For each task the set of document can be structured defining for example a title, keywords, images, and so on. Some ad-hoc metadata allows for partitioning the set of documents. For example, in the same set we can have European or American documents and a mechanism that allows for choosing only one of these sets. Moreover, it is important to remark that very often in an evaluation campaign the so called closed world assumption holds, which means that the set of documents is finite and known a-priori.
A topic represents an information need. It is structured and its structure can change according to the task at hand. Documents can be assessed as being relevant or not (or more or less relevant) for a given information need (topic). The relevance of a document with respect to a specific topic is independent of the other documents in the collection, based solely on the qualities of that document. In some case we can have different sets of relevance assessment for a set of documents. The relevance assessment can be done manually, automatically, or using online approach like Amazon mechanical Turk.

Having introduced these basic notions we can analyze the PROMISE data. We will refer to data which are stored in the DIRECT system developed by Padua University (see PROMISE Deliverable 3.2 Specification of the evaluation infrastructure based on user requirements, Section 6, for details about accessing such data). These data can be represented by the TME (Topics-Metrics-

Experiment) cube shown on Figure 6:

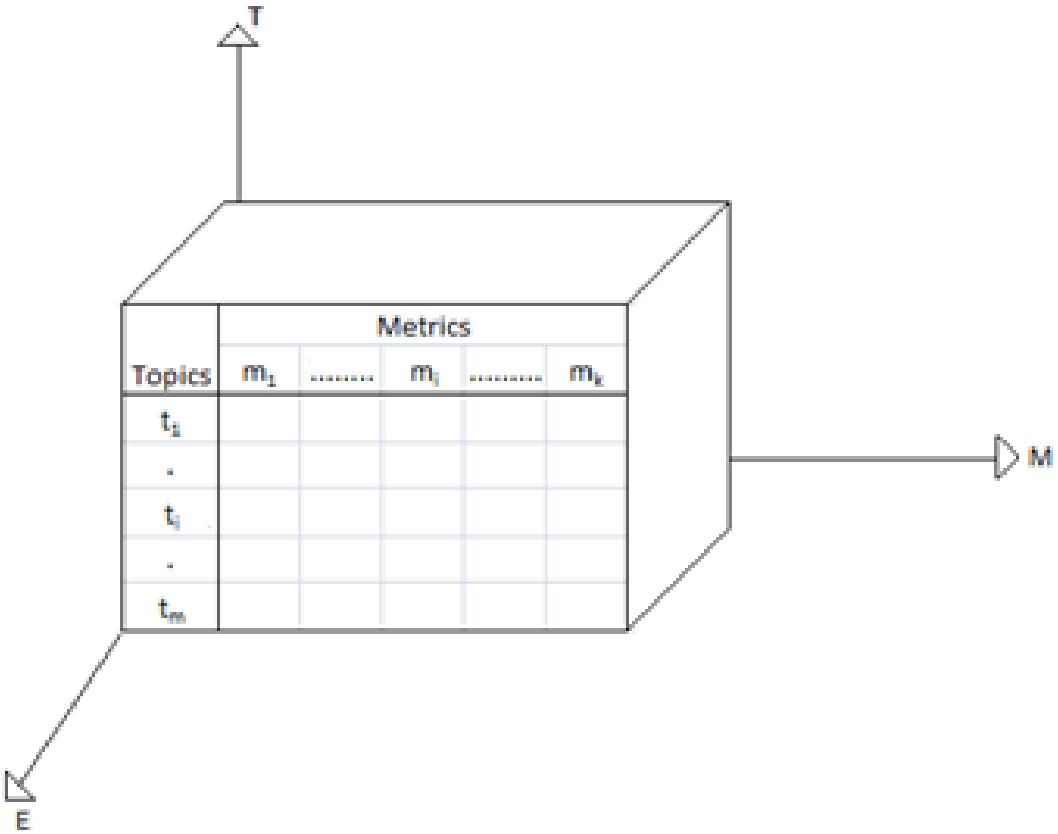

Figure 6: The PROMISE TME Data cube

Starting from this cube, we can aggregate or manipulate data in different ways, according to our needs. In particular we are recall the structure of a TE(m) table, useful in the next sections.

Considering the TME cube we can derive the table shown on Figure 7, useful to analyze a single metric m in terms of topics and experiments. In particular, this table is represented by a matrix T x E where T is the set of topics and E is the set of experiments. In the following we will refer to this kind of tables with the name TE(m) tables (topics x experiments table of metric m). Comparisons are made along rows, to evaluate the behavior of a single topic, or among columns to compare two or more experiments. For the number of topics the same considerations previously discussed hold. The number of experiments depends on how many algorithms are compared.

To complete our analysis we recall the concept of meta-attribute. A meta-attribute is a categorical attribute that is associated with a cube component (for example the experiments) and it is used to define a further classification of data, with respect to a category. Examples of meta-attributes are: reference track, year, and type of search. Meta-attributes are mainly associated to experiments and documents (see PROMISE Deliverable 3.2 Specification of the evaluation infrastructure based on

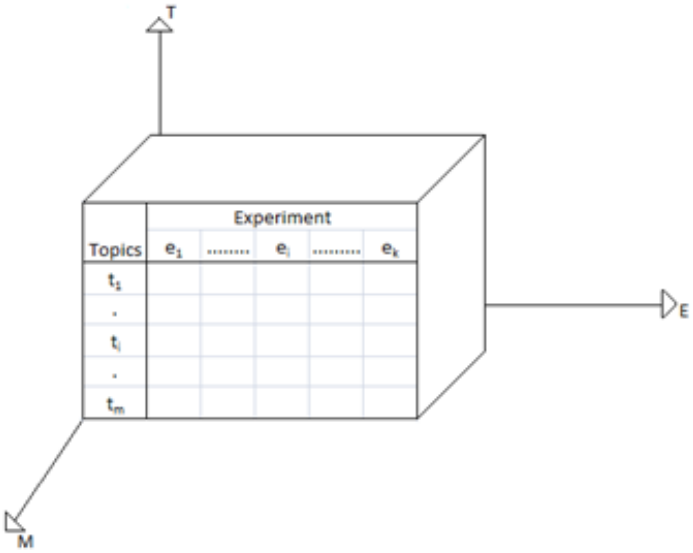

Figure 7: Projecting the TME data cube on the Topics-Experiments axes

user requirements), but also topics can have their own meta-attributes (for example the provenance data). Actually there are no meta-attributes defined on metrics. A possible meta-attribute for metrics is the scope of metric, but we will consider this categorization for future developments.

8.2 PROMISE Visual Analytics module architecture

The overall architecture is depicted on Figure 8 and its structure is totally parametric, without any assumption about the data structure (in the most general case it is contained in non-normalized table). Moreover, there are no assumptions about visualizations (it is possible to obtain any kind of visualization), about the mapping between data and visualizations, and about analytical components.
The most general situation is the one in which the system presents the user with multiple visualizations, each of them working on the same set of data. Visualizations are synchronized using two main interaction mechanisms: selection (it is just a way to focus the attention on a subset of data) and the highlighting (it allows for highlighting a part of the displayed data).
In order to produce a visualization, three main steps are, in principle, needed:

1. data extraction from PROMISE database.
2. data manipulation, i.e., deriving new attributes, applying some aggregation operations , applying some analytical algorithms, etc. During such a process the system adds some hidden attributes to the data, in order to support the selection and the highlighting mechanisms.
3. Mapping the data obtained from step two on one or more visualizations.

The first two steps are optional: in some cases the system will automatically perform them, allowing the user to focus only on the mapping and analysis activities.

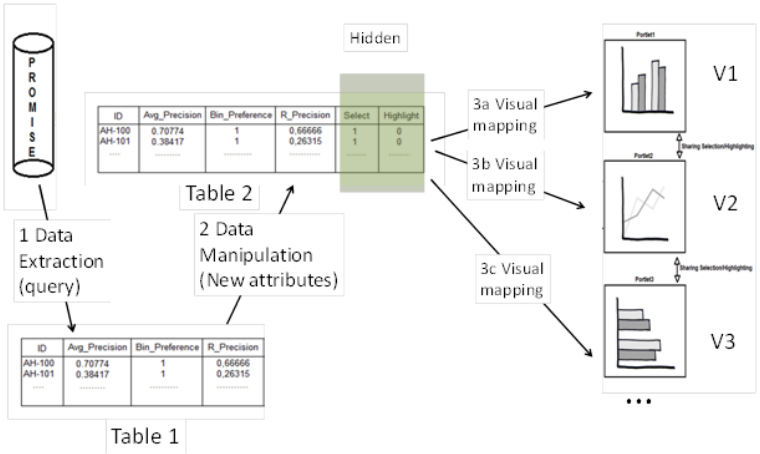

Figure 8: : The PROMISE Visual Analytics architecture

8.3 The User interface

The wizard home page allows for accessing a set of predefined ways to perform a data analysis similar to those currently used by users of PROMISE community. In particular we focus on two basic approaches: per topic analysis, and per experiment analysis.

According to the requirements, these two analysis allow for selecting six visualizations: bi-dimensional scatter-plot, bar charts, stacked bar-charts, box plots, table lens, and frequency distributions. Depending on the chosen approach, the system will present the user with different subset of these visualizations. In the following we will analyze the two basic approaches to data analysis and present a general description of their own visualizations. Each of them requires the preliminary selection of a task. This choice corresponds to a selection of an URL connected to the physical address of data. The URL can change depending on whether the user wants to load the entire data set related to a task or just a part of it.

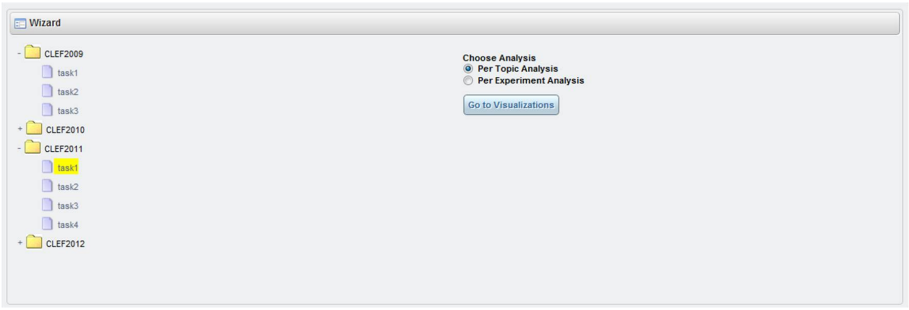

Figure 9: : Selection of the tasks

As shown in figure 9 the user can choose a task within a track of a campaign and the kind of analysis that he wants to perform. In particular he can choose between two kind of analysis

Topics	Experiment				
	e_1	.....	e_i	.....	e_k
t_1					
.					
t_i					
.					
t_m					

Figure 10: : The TE(m) table

- per topic analysis
- per experiment analysis

According to the typical PROMISE analysis tasks, we foresee a set of ad-hoc visualizations. These visualizations must support synchronization and interaction that are specific for each visualization. Moreover, for each visualization it is needed a mapping mechanism in order to support the user in the creation process. From a technical point of view, designing ad-hoc visualizations implies the design of a module for each visualization. If a user wants to use a visualization he has to select the suitable module and to map on it the desired data. The initial requirement analysis allows for selecting six visualizations: bi-dimensional scatter-plot, bar charts, stacked bar-charts, box plots, table lens, and frequency distributions. Depending on the chosen approach, the system will present the user with different subset of these visualizations. In the following we will analyze the three basic approaches to data analysis and present a general description of the foreseen visualizations.

8.3.1 Per topic analysis

In our model, per topic analysis means comparing a set of experiments on each topic with respect to a chosen metric. Therefore the first step for a user is to choose a metric m. Looking at the TME data cube described in the previous section we can note that choosing a metric is equivalent to fix an axis and reduce the set of data to the TE(m) table shown on Figure 10.

Per topic analysis implies a comparison on each topic, so we foresee to represent topics on x-axis in each available visualization.
Having chosen a metric m, the user has to choose the data to display. This choice corresponds to select a subset of the columns of the TE(m) table and that can be performed either by selection or using meta-attribute. For example, the user can select some participants or some experiments.

Moreover the user can decide to highlight some elements within the visualizations.
The per topic analysis is performed on TE(m) table and visualizations have the following meaning:

Figure 11: Per topic analysis

- table chart displays a TE(m) table.
- box-plot chart displays an aggregated visualization of values that each topic reached with respect to the chosen metric (rows of TE(m). In a per topic analysis a box plot chart is used to evaluate the trend of a topic among experiments with respect to a chosen metric. A box plot chart presents a box plot for each topic on x-axis and the chosen metric on y-axis.
- scatter-plot chart is used to compare two experiment on all topic with respect to the chosen metric.In a per topic analysis a bi-dimensional scatter plot presents on x-axis topics and on y-axis the chosen metric. Each metric value is represented by a point. For each topic there are as many points on y-axis as the selected TE(m) columns. To see the trend of a single experiment (a column) you can unify its points with a polyline. To highlight some point it is possible to use color or markers.

- bar/point/line/stacked-bar chart is used to compare topics.In a per topic analysis a bar chart can be used to compare two or more experiments on all the topics with respect to a chosen metric. Although it is possible to compare more than two experiments, as the number of experiments (topics) increases the chart representation loses clarity. Possible comparisons are two algorithm of the same participant or the best algorithm of two different participants

Figure 11 shows a screenshot of the actual prototype in a per topic analysis.

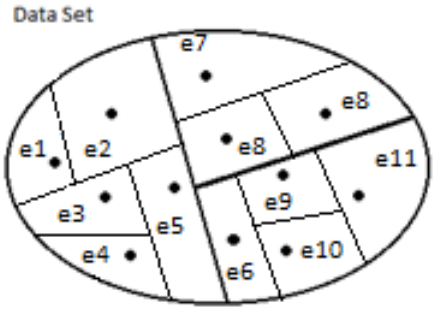

Figure 12: Partitioning data for per experiment analysis

8.3.2 Per experiment analysis

In our model per experiment analysis allows for comparing scores of a set of experiments with respect to a chosen descriptive statistics. Each visualization in a per experiment analysis presents experiments on x-axis and the chosen statistics on y-axis. It can be seen as a particular case of a participant analysis. In a per experiments analysis, we partition the set of data using the name of experiment as category. The resulting partitions are constituted by only one element. The Venn diagram on Figure 12 shows a partitioning of data set defined by the name of the experiments.

As consequence of the above, each visualization in a per experiment analysis has only one value for each point of x-axis.

Figure 13 shows an actual example of a per experiment analysis.

8.4 Main functionalities of the Visual Analytics component

The analysis environment, regardless the different analysis strategies described in the previous section, provide users with the following set of functionalities:

1. interaction functionalities
   - manipulation
   - layout editing

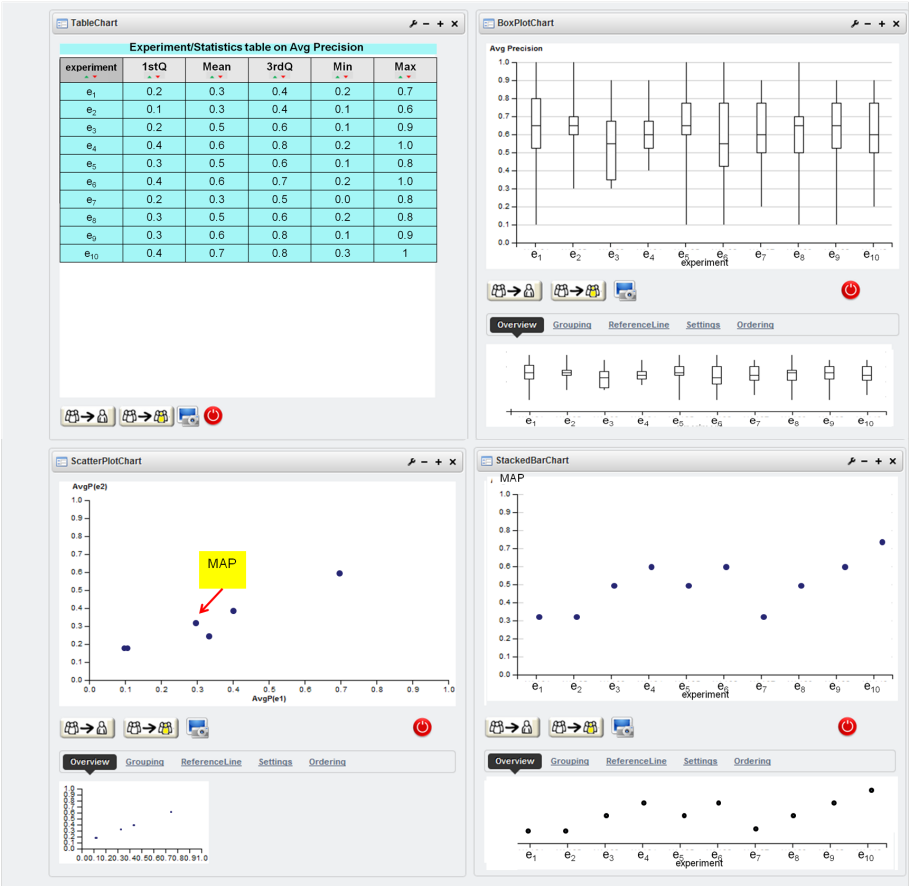

Figure 13: per experiment analysis

- reference lines

2. annotations and saving

Manipulation features are those interactions that changes the set of displayed data(i.e. semantic of visualization). In particular belong to manipulations selection, highlighting, and reset.

In Figure 14 is shown the implementation of the aforementioned manipulations. The user is capable of picking, from any of the different visualizations displayed, a set of relevant element for his goals: after the picking phase, the simple pressure of one of the contextual buttons will start one of the operations:

- Selection, that will display in the main focus area only the element picked.
- Highlight, that will give emphasis on the elements picked and will put in background the remaining, for pointing at particular facts of interest

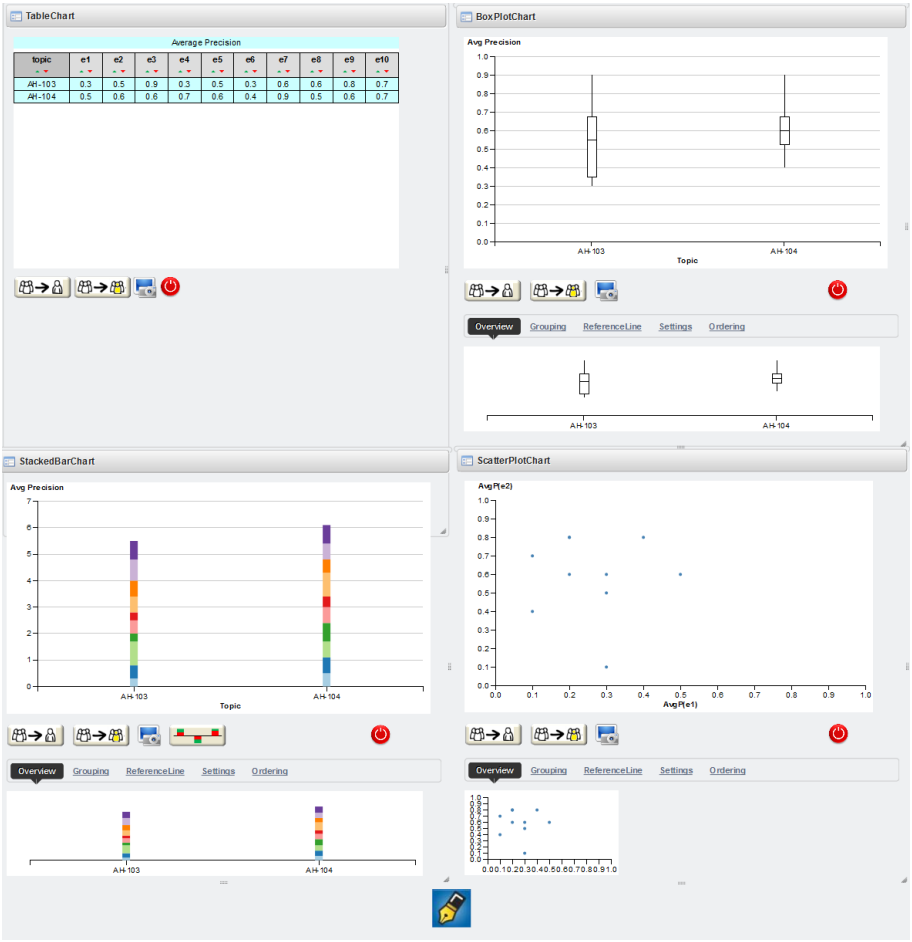

Figure 14: manipulation operation: Select

- Reset, that will return the visualizations to their original state, not taking into account any of the effects of previous manipulations

In any case, a practical overview of the whole set of data visualized will always be available just below the main one, in order for the user to don't become lost in his analysis.

Belong to **layout editing** those interactions that allow users to organize data according to their needs and to customize the final view (i.e. syntax of visualization). An example is the ordering feature.
In Figure 15 is shown one of the menu for selecting the type of ordering that can be chose for the data.

Belong to **reference lines** those interaction that emphasize interesting statistical values(mean, min, max). These statistical indicators can be really helpful for a quick analysis of the data: Figure 16 shows an example of their usage.

Figure 15: layout editing: ordering operation

Figure 16: reference Lines

Adding to the described operation a further one was conceived to quickly save screenshots of the manipulations applied to the analyzed data, in order to have the capability of quickly retrieve interesting insights. This operation is accessible by pressing the button representing a computer screen + camera near the reset button.

On a final note, we want to highlight two important features of this prototype: the first one is that all the visualizations supported are coordinated among them: that means that all the operations of manipulation,layout editing and reference lines will have effects on both the actual visualizations and all the other linked. The second one is that, having implemented each single visualizations like a portlet, we have obtained a very flexible and expandable environment where new visualizations can be easily added and synchronized with the existing one, and the layout can be easily adjusted to the needs of the user.

8.4.1 Annotations and saving

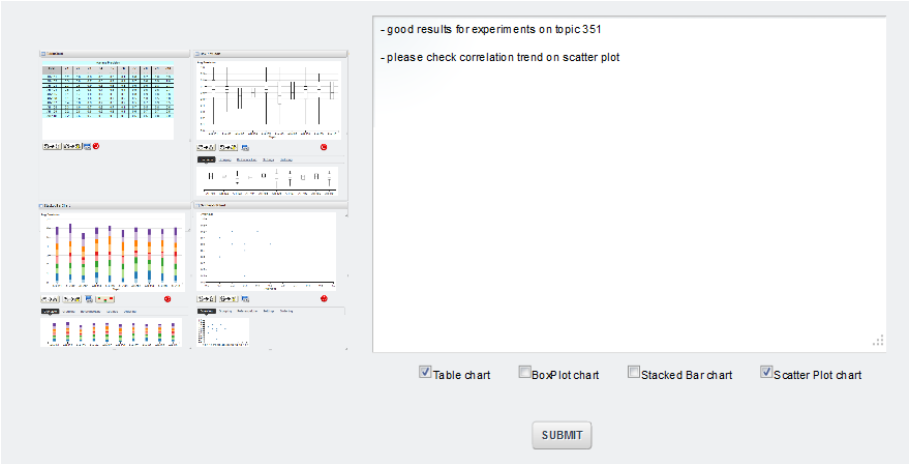

Figure 17: adding annotation to a visualization

One of the peculiar features of this version of the Collaborative User Interface prototype is the management of the annotations. After the user has made reason about the data that are visualized in the different charts available, he may want to add permanent annotations for future referencing and/or to involve other users in the particular insights that he has discovered.
In order to do so the user will have to simple press the button on the bottom of the screen, representing a pen, and it will led the system to a new tab for annotation shown in Figure 17.

In this tab will be present a preview screenshots of the particular set of visualizations under analysis: this preview will take into account all the different manipulations operations done by the user on the set of visualizations, and will serve as a reference for the textual and non-textual annotations that the user what to share.
Keeping on with the description of this tab, the second element is an hyper-textual box where the user can add all the relevant information he want to share with others or simply saving for future uses, and a check boxes system constructed with the title names of the involved visualizations for indexing and referencing to which kind of visualizations are referred the comments.
From a rationale point of view, this choice has been made in order to provide to the user the capability to referencing more than one visualizations at a a time in his annotations, and not force him to iterate in multiple annotations passages for each of the visualizations involved. Nonetheless, the system can be easily reverted to a single visualization-single annotation fashion by simply re-adjusting the general schema of annotation described above to a single one, and adding a command button for each visualizations in the same way as is managed the manipulation operations.
When the user will be sure about the comments he wants to add, he will simply press the SUBMIT button for letting the information be passed to the annotation management system, described deeper in the first part of this deliverable. After the phase of collecting the information added in the annotation panel, the system will present the user with the analysis environment tabs where are present the visualizations on which he was working on.

These pieces of information will be attached to the image(s) under analysis an will be saved together with them: the system treats images as "first class objects" and store and retrieve not simple images but all the details behind them (e.g.,track, run, experimental data, analysis strategy, and image manipulation), including annotations.

A XML Schemas

This appendix reports the XML schemas of the different managed resources.

The XML schema of FAST[5] is reported below. It relies on the XML schema of DIRECT[6] and the XML schema of the *IMS Component Integrator (ICI)*[7] library, reported in Appendix A of D3.3 [Agosti et al., 2012].

A.1 FAST XML Schema

```
<?xml version="1.0" encoding="UTF-8"?>
<!-- edited with XMLSpy v2012 rel. 2 sp1 (x64) (http://www.altova.com) by UniversitÃ
    degli Studi di Padova (UniversitÃ degli Studi di Padova) -->
<xs:schema xmlns:ims="http://ims.dei.unipd.it/" xmlns:xs="http://www.w3.org/2001/
    XMLSchema" xmlns:dc="http://purl.org/dc/elements/1.1/" xmlns:xcql="http://docs.oasis-
    open.org/ns/search-ws/xcql" targetNamespace="http://ims.dei.unipd.it/"
    elementFormDefault="qualified" attributeFormDefault="qualified" version="3.00" xml:
    lang="en">
  <xs:annotation>
    <xs:documentation xml:lang="en">This schema provides the base elements and types used
         by the Flexible Annotation Service Tool (FAST).</xs:documentation>
    <xs:documentation xml:lang="en">Version 3.00.</xs:documentation>
    <xs:documentation xml:lang="en">Created on 2005-10-05</xs:documentation>
    <xs:documentation xml:lang="en">Last modified on 2012-07-06</xs:documentation>
    <xs:documentation xml:lang="en">Authored by Nicola Ferro (ferro@dei.unipd.it)</xs:
        documentation>
    <xs:documentation xml:lang="en">Copyright (c) 2003-2012 - Information Management
        Systems (IMS) Research Group (http://ims.dei.unipd.it/) - Department of
        Information Engineering (http://www.dei.unipd.it/) - University of Padua (http://
        www.unipd.it/)</xs:documentation>
  </xs:annotation>
  <xs:include schemaLocation="http://ims.dei.unipd.it/data/xml/ici.3.00.xsd">
    <xs:annotation>
      <xs:documentation xml:lang="en">Imports the schema for the IMS Component Integrator
          (ICI) library.</xs:documentation>
    </xs:annotation>
  </xs:include>
  <xs:element name="annotation" substitutionGroup="ims:resource">
    <xs:annotation>
      <xs:documentation xml:lang="en">Represents an annotation.</xs:documentation>
    </xs:annotation>
    <xs:complexType>
      <xs:complexContent>
        <xs:extension base="ims:namespace-identifiable-timestamp-traceable-access-
            controllable-resource-type">
```

[5] http://ims.dei.unipd.it/data/xml/fast.3.00.xsd
[6] http://ims.dei.unipd.it/data/xml/direct.3.00.xsd
[7] http://ims.dei.unipd.it/data/xml/ici.3.00.xsd

```
24            <xs:sequence>
25              <xs:element name="signs" minOccurs="0">
26                <xs:annotation>
27                  <xs:documentation xml:lang="en">The signs of the annotation.</xs:
                        documentation>
28                </xs:annotation>
29                <xs:complexType>
30                  <xs:sequence>
31                    <xs:element ref="ims:sign" maxOccurs="unbounded"/>
32                  </xs:sequence>
33                </xs:complexType>
34              </xs:element>
35            </xs:sequence>
36          </xs:extension>
37        </xs:complexContent>
38      </xs:complexType>
39    </xs:element>
40    <xs:element name="anchor" substitutionGroup="ims:resource">
41      <xs:annotation>
42        <xs:documentation xml:lang="en">Represents an anchor to a part of a digital object.
              </xs:documentation>
43      </xs:annotation>
44      <xs:complexType>
45        <xs:complexContent>
46          <xs:extension base="ims:resource-type">
47            <xs:sequence>
48              <xs:choice>
49                <xs:element ref="ims:digital-object"/>
50                <xs:element ref="ims:annotation"/>
51              </xs:choice>
52              <xs:element name="pointer" minOccurs="0">
53                <xs:annotation>
54                  <xs:documentation xml:lang="en">A pointer to a part of a digital object.<
                        /xs:documentation>
55                  <xs:documentation xml:lang="en">The format of the pointer depends on the
                        MIME type of digital object and/or the MIME type of the anchored part
                        .</xs:documentation>
56                  <xs:documentation xml:lang="en">In version 2.0, extent was an attribute
                        of anchor. Since version 2.1, it is an element to allow for more
                        complex anchoring schemes, e.g. based on XML fragments.</xs:
                        documentation>
57                </xs:annotation>
58                <xs:complexType mixed="true">
59                  <xs:complexContent>
60                    <xs:restriction base="xs:anyType">
61                      <xs:sequence>
62                        <xs:any processContents="lax" minOccurs="0" maxOccurs="unbounded"/>
63                      </xs:sequence>
64                      <xs:attribute name="anchored-media-type">
65                        <xs:annotation>
66                          <xs:documentation xml:lang="en">The media type of the anchored
                                part of the digital objet object according to MIME (
                                Multipurpose Internet Mail Extensions) standard.</xs:
                                documentation>
67                        </xs:annotation>
68                        <xs:simpleType>
69                          <xs:restriction base="xs:string">
70                            <xs:pattern value="(text|image|audio|video|application|message|
                                  multipart)/(\p{L}|\.|\-)+(;.*)?"/>
71                          </xs:restriction>
```

```
                        </xs:simpleType>
                      </xs:attribute>
                    </xs:restriction>
                  </xs:complexContent>
                </xs:complexType>
              </xs:element>
            </xs:sequence>
          </xs:extension>
        </xs:complexContent>
      </xs:complexType>
    </xs:element>
    <xs:element name="sign" substitutionGroup="ims:resource">
      <xs:annotation>
        <xs:documentation xml:lang="en">Represents a sign.</xs:documentation>
      </xs:annotation>
      <xs:complexType>
        <xs:complexContent>
          <xs:extension base="ims:identifiable-timestamp-traceable-resource-type">
            <xs:sequence>
              <xs:element ref="ims:annotation" minOccurs="0"/>
              <xs:element name="concepts" minOccurs="0">
                <xs:annotation>
                  <xs:documentation xml:lang="en">The meanings of the signs.</xs:
                      documentation>
                </xs:annotation>
                <xs:complexType>
                  <xs:sequence>
                    <xs:element ref="ims:concept" maxOccurs="unbounded"/>
                  </xs:sequence>
                </xs:complexType>
              </xs:element>
              <xs:element name="annotate" minOccurs="0">
                <xs:complexType>
                  <xs:sequence>
                    <xs:element ref="ims:anchor"/>
                  </xs:sequence>
                </xs:complexType>
              </xs:element>
              <xs:element name="relate-to" minOccurs="0">
                <xs:complexType>
                  <xs:sequence>
                    <xs:element ref="ims:anchor"/>
                  </xs:sequence>
                </xs:complexType>
              </xs:element>
              <xs:element ref="ims:content" minOccurs="0"/>
            </xs:sequence>
            <xs:attribute ref="ims:media-type"/>
            <xs:attribute ref="ims:language"/>
          </xs:extension>
        </xs:complexContent>
      </xs:complexType>
    </xs:element>
    <xs:element name="fast" type="ims:ici-type">
      <xs:annotation>
        <xs:documentation xml:lang="en">Provides information about one or more resources of
            the FAST annotation service.</xs:documentation>
      </xs:annotation>
    </xs:element>
</xs:schema>
```

References

Agosti, M., Bonfiglio-Dosio, G., and Ferro, N. (2007). A Historical and Contemporary Study on Annotations to Derive Key Features for Systems Design. *International Journal on Digital Libraries (IJDL)*, 8(1):1–19.

Agosti, M., Braschler, M., Di Buccio, E., Dussin, M., Ferro, N., Granato, G. L., Masiero, I., Pianta, E., Santucci, G., Silvello, G., and Tino, G. (2011). Deliverable D3.2 – Specification of the evaluation infrastructure based on user requirements. PROMISE Network of Excellence, EU 7FP, Contract N. 258191. http://www.promise-noe.eu/documents/10156/fdf43394-0997-4638-9f99-38b2e9c63802.

Agosti, M., Di Buccio, E., Ferro, N., Masiero, I., Nicchio, M., Peruzzo, S., and Silvello, G. (2012). Deliverable D3.3 – Prototype of the Evaluation Infrastructure. PROMISE Network of Excellence, EU 7FP, Contract N. 258191. http://www.promise-noe.eu/documents/10156/3783730a-bce3-481b-83df-48e209c6286a.

Agosti, M. and Ferro, N. (2005). Annotations as Context for Searching Documents. In Crestani, F. and Ruthven, I., editors, *Proc. 5th International Conference on Conceptions of Library and Information Science – Context: nature, impact and role (CoLIS 5)*, pages 155–170. Lecture Notes in Computer Science (LNCS) 3507, Springer, Heidelberg, Germany.

Agosti, M. and Ferro, N. (2008). A Formal Model of Annotations of Digital Content. *ACM Transactions on Information Systems (TOIS)*, 26(1):3:1–3:57.

Agosti, M., Ferro, N., Frommholz, I., and Thiel, U. (2004). Annotations in Digital Libraries and Collaboratories – Facets, Models and Usage. In Heery, R. and Lyon, L., editors, *Proc. 8th European Conference on Research and Advanced Technology for Digital Libraries (ECDL 2004)*, pages 244–255. Lecture Notes in Computer Science (LNCS) 3232, Springer, Heidelberg, Germany.

Berendsen, R., Braschler, M., Gäde, M., Kleineberg, M., Lupu, M., Petras, V., and Reitberger, S. (2012). Deliverable D4.3 – Final Report on Alternative Evaluation Methodology. PROMISE Network of Excellence, EU 7FP, Contract N. 258191. http://www.promise-noe.eu/documents/10156/0092298d-892b-45c0-a534-b9a3d0c717b1.

Berners-Lee, T., Fielding, R., and Masinter, L. (2005). Uniform Resource Identifier (URI): Generic Syntax. RFC 3986.

Candela, L., Castelli, D., Ferro, N., Ioannidis, Y., Koutrika, G., Meghini, C., Pagano, P., Ross, S., Soergel, D., Agosti, M., Dobreva, M., Katifori, V., and Schuldt, H. (2007). *The DELOS Digital Library Reference Model. Foundations for Digital Libraries*. ISTI-CNR at Gruppo ALI, Pisa, Italy, http://www.delos.info/files/pdf/ReferenceModel/DELOS_DLReferenceModel_0.98.pdf.

Croce, M., Di Reto, E., Granato, G. L., Hansen, P., Sabetta, A., Santucci, G., and Veltri, F. (2011). Deliverable D5.1 – Collaborative user interface requirements. PROMISE Network of

Excellence, EU 7FP, Contract N. 258191. http://www.promise-noe.eu/documents/10156/50834686-2118-48f8-a57b-8553ec3d7981.

Crockford, D. (2006). The application/json Media Type for JavaScript Object Notation (JSON). RFC 4627.

Fox, E. A. and Shaw, J. (1993). Combination of Multiple Searches. In Harman, D. K., editor, *The Second Text REtrieval Conference (TREC-2)*, pages 243–252. National Institute of Standards and Technology (NIST), Special Pubblication 500-215, Washington, USA. http://trec.nist.gov/pubs/trec2/t2_proceedings.html.

Freed, N. and Borenstein, N. (1996a). Multipurpose Internet Mail Extensions (MIME) Part One: Format of Internet Message Bodies. RFC 2045.

Freed, N. and Borenstein, N. (1996b). Multipurpose Internet Mail Extensions (MIME) Part Two: Media Types. RFC 2046.

Frommholz, I., Brocks, H., Thiel, U., Neuhold, E., Iannone, L., Semeraro, G., Berardi, M., and Ceci, M. (2003). Document-Centered Collaboration for Scholars in the Humanities – The COLLATE System. In Koch, T. and Sølvberg, I. T., editors, *Proc. 7th European Conference on Research and Advanced Technology for Digital Libraries (ECDL 2003)*, pages 434–445. Lecture Notes in Computer Science (LNCS) 2769, Springer, Heidelberg, Germany.

Granato, G. L., Santucci, G., and Tino, G. (2011). Deliverable D5.2 – User interface and Visual analytics environment requirements. PROMISE Network of Excellence, EU 7FP, Contract N. 258191. http://www.promise-noe.eu/documents/10156/21f1512a-5b47-48ae-834a-89d6441d079e.

Halasz, F. G., Moran, T. P., and Trigg, R. H. (1987). Notecards in a Nutshell. In Carroll, J. M. and Tanner, P. P., editors, *Proc. Conference on Human Factors in Computing Systems and Graphics Interface (CHI 1987)*, pages 45–52. ACM Press, New York, USA.

Handschuh, S. and Staab, S., editors (2003). *Annotation for the Semantic Web*. IOS Press, Amsterdam, The Netherlands.

Haslhofer, B., Jochum, W., King, R., Sadilek, C., and Schellner, K. (2009). The LEMO Annotation Framework: Weaving Multimedia Annotations with the Web. *International Journal on Digital Libraries (IJDL)*, 10(1):15–32.

ISO 15836:2009 (2009). Information and documentation - The Dublin Core metadata element set. Recommendation ISO 15836:2009.

ISO 639-2 (1998). Codes for the representation of names of languages – Part 2: Alpha-3 code. Recommendation ISO 639-2:1998.

ISO/IEC 9834-8 (2008). Information technology – Open Systems Interconnection – Procedures for the operation of OSI Registration Authorities: Generation and registration of Universally Unique Identifiers (UUIDs) and their use as ASN.1 Object Identifier components. Recommendation ISO/IEC 9834-8:2008.

Josefsson, S. (2006). The Base16, Base32, and Base64 Data Encodings. RFC 4648.

Leach, P., Mealling, M., and Salz, R. (2005). A Universally Unique IDentifier (UUID) URN Namespace. RFC 4122.

Savoy, J. (1996). Citation Schemes in Hypertext Information Retrieval. In Agosti, M. and Smeaton, A., editors, *Information Retrieval and Hypertext*, pages 99–120. Kluwer Academic Publishers, Norwell (MA), USA.

W3C (2004a). Resource Description Framework (RDF): Concepts and Abstract Syntax – W3C Recommendation 10 February 2004. http://www.w3.org/TR/rdf-concepts/ [last visited 2007, March 23].

W3C (2004b). XML Schema Part 0: Primer Second Edition – W3C Recommendation 28 October 2004. http://www.w3.org/TR/xmlschema-0/.

W3C (2004c). XML Schema Part 1: Structures – W3C Recommendation 28 October 2004. http://www.w3.org/TR/xmlschema-1/.

W3C (2004d). XML Schema Part 2: Datatypes – W3C Recommendation 28 October 2004. http://www.w3.org/TR/xmlschema-2/.

W3C (2006). Extensible Markup Language (XML) 1.1 (Second Edition) – W3C Recommendation 16 August 2006, edited in place 29 September 2006. http://www.w3.org/TR/xml11/.

W3C (2008). Extensible Markup Language (XML) 1.0 (Fifth Edition) – W3C Recommendation 26 November 2008. http://www.w3.org/TR/xml/.

W3C (2009a). SKOS Simple Knowledge Organization System Primer – W3C Working Group Note 18 August 2009. http://www.w3.org/TR/skos-primer.

W3C (2009b). SKOS Simple Knowledge Organization System Reference – W3C Recommendation 18 August 2009. http://www.w3.org/TR/skos-reference.

Zyp, K. (2010). A JSON Media Type for Describing the Structure and Meaning of JSON Documents. RFC draft-zyp-json-schema-03.