Department of
Information Engineering
and Computer Science DISI

UNIVERSITY OF TRENTO - Italy

DISI - Via Sommarive 14 - 38123 Povo - Trento (Italy)
http://www.disi.unitn.it

REPORT ON THE STATE-OF-THE-ART AND REQUIREMENTS FOR ANNOTATION REPRESENTATION MODELS

Tobias Burger, Ilya Zaihrayeu, Pierre Andrews, Denys Babenko, Juan Pane and Borislav Popov

January 2010

Technical Report # DISI-10-007

INSEMTIVES
FP7-ICT-2007-3
Contract no.: 231181
www.insemtives.eu

INSEMTIVES

Deliverable 2.1.1

Report on the State-of-the-Art and Requirements for Annotation Representation Models

Editor:	Tobias Bürger, STI, University of Innsbruck
Deliverable nature:	R
Dissemination level: (Confidentiality)	Public (PU)
Contractual delivery date:	30.06.2009
Actual delivery date:	30.06.2009
Version:	1.0
Total number of pages:	69
Keywords:	semantic annotation

Abstract

INSEMTIVES overall goals is to increase the amount of available semantic content by looking into incentives and by bridging the gap between human and computational intelligence. Annotations are one form of semantic content which are supposed to make implicit information explicit and which typically attach additional meaning to a resource. This deliverable provides a classification scheme for annotation models and summarizes the state of the art in the areas relevant for the INSEMTIVES project. It concludes with an analysis of the requirements of the use cases partners in the project with respect to annotation models and a summary of them.

Impressum

INSEMTIVES – Incentives for Semantics

INSEMTIVES

WP2: Models and Methods for the Creation and Usage of Lightweight, Structured Knowledge

Report on the state-of-the-art and requirements for annotation representation models

[Editor: Tobias Bürger, STI, University of Innsbruck]

[Work-package leader: Ilya Zaihrayeu, University of Trento]

[Estimation of PM spent on the Deliverable: 3 PM]

Acknowledgement

The project is co-funded by the European Union, through the ICT Cooperation programme http://cordis.europa.eu/fp7/cooperation/home_en.html

Executive Summary

INSEMTIVES' overall goal is to bridge the gap between human and computational intelligence in the current semantic content authoring R&D landscape in order to increase the amount and quality of annotations for a broad range of different resource types. As such it will develop methodologies for the creation of semantic content which intelligently combine automated approaches with human tasks. In order to foster user participation it will investigate incentive mechanisms to motivate users to contribute semantic annotations.

Work package 2, "Models and Methods for Creation of Lightweight, Structured Knowledge", develops models for representing semantic contents and their provenance, explicitly and implicitly generated by users. Apart from this, the work package develops a set of methods necessary to support the full life cycle of these semantic contents, starting from their generation and continuing with their maintenance and evolution in time.

This deliverable lays the foundation for the work in work package 2. It provides an overview of existing annotation representation models. Moreover, it describes requirements for annotation representation models developed in INSEMTIVES. It considers also those requirements from the case studies which deal with the annotation of textual content, media, and Web services.

As a starting point the deliverable sets the scene by motivating the work to be done in this work package and by defining the goals and scope of this deliverable. At the core, it presents the state of the art in annotation representation models: It first provides a classification schema which classifies models according to their most dominant dimensions which are: (1) structural complexity of annotations; (2) the type of vocabularies used; and (3) the collaboration model used to enable singe-user or community based annotation scenarios. Subsequently it outlines existing approaches in the application domains of the use case partners in INSEMTIVES, that is: document, Web service, and multimedia annotation.

The deliverable is concluded with a summary of requirements for annotation representation models which will provide the foundation for further work in this work package. The requirements summarize the annotation representation model needs in the respective use cases of INSEMTIVES by referring back to the classification scheme introduced earlier in the deliverable. The requirements were gathered from the use case partners using a questionnaire which is available in the Appendix of this document.

List of Authors

Company	Author
University of Innsbruck	Tobias Bürger
University of Trento	Ilya Zaihrayeu
University of Trento	Pierre Andrews
University of Trento	Denys Babenko
University of Trento	Juan Pane
Ontotext	Borislav Popov

Contents

List of Figures

List of Tables

Abbreviations

GATE General Architecture for Text Engineering

HTTP Hypertext Transfer Type Protocol

IR Information Retrieval

NLP Natural Language Processing

OWL Web Ontology Language

RDF Resource Description Framework

SAWSDL Semantic Annotations for WSDL

SOAP Simple Object Access Protocol

SWSF Semantic Web Services Framework

URI Unified Resource Identifier

WSDL Web Service Description Language

WSMO Web Service Modeling Ontology

XML Extensible Markup Language

Definitions

Annotation The term *annotation* is used both as a noun denoting a piece of additional information and as a verb referring to the process of creating this additional information.

Annotation Model An *annotation model* defines the actual form in which the annotation, that is, additional information is expressed, and how it is linked to the original content being annotated.

Attribute An *attribute annotation element* is a pair $\langle AN, AV \rangle$, where AN is the name of the attribute and AV is the value of the attribute [121].

Authority File An *authority file* is a kind of controlled vocabulary. In an authority file synonymous terms are grouped into concepts and one of the terms is selected as the concept name and used for visualization and navigation purposes (this term is called the *preferred term*) [33].

Controlled Vocabulary *Controlled vocabularies* are organized lists of words and phrases, or notation systems, that are used to initially tag content, and then to find it through navigation or search [117].

Folksonomies Folksonomies [116] are classification systems which are generated in a bottom-up fashion through collective and massive annotation made by a (large) group of users [40, 69].

Ontology As defined by Studer et al., "an ontology is an explicit specification of a (shared) conceptualization" [103]

Relation A relation annotation element is a pair $\langle Rel, Res \rangle$, where Rel is the name of the relation and Res is another resource (i.e., different from the one being annotated).

Resource A *resource* is described by an annotation within annotation models. A resource can be any identifiable content identified by an unique identifier such as URIs. Examples of resources in the context of the Web are electronic documents such as images or text documents or their parts (e.g., a reference to an entity from a text document or an area in an image).

Semantic Annotation *Semantic annotation* describes both the process and the resulting annotation or metadata consisting of aligning a resource or a part of it with a description of some of its properties and characteristics with respect to a formal conceptual model or ontology [80].

Tag A *tag* annotation element is a non-hierarchical keyword or free-from term assigned to a resource [123].

Taxonomy A *taxonomy* is an extension of the authority files controlled vocabulary with the broader term (BT) and narrower term (NT) relations which are defined between controlled vocabulary concepts [33, 58].

User A *user* in the context of this deliverable provides annotations for resources.

Web service *"A Web service is a software system designed to support interoperable machine-to-machine interaction over a network. It has an interface described in a machine-processable format (specifically WSDL [18]). Other systems interact with the Web service in a manner prescribed by its description using SOAP-messages [41], typically conveyed using HTTP with an XML serialization in conjunction with other Web-related standards."* [43]

1 Introduction

Motivation

Information retrieval today remains a challenge due to the vast amounts of information, the non-graspable characteristics of users and their needs, and the diversity of resources to be indexed. Due to this fact, annotations are being considered as a source for reliable metadata. Annotations are supposed to make implicit information explicit and to attach additional meaning to a resource. This additional information might be exploited in search and retrieval to increase the accessibility and visibility of resources. Due to the fact that for many resource types reliable annotations can not be provided fully automatically, a considerable amount of manual effort is needed to create them. Made popular on Web2.0 sites such as Flickr [27] or YouTube [125], annotations in the form of user-contributed tags gained a huge popularity as a means to effectively support organization and discovery of different types of resources on the Web. This type of annotation is considered to be informal as it bears no structure in it. The value of informal annotations in the form of syntactic tags is however restricted and commonly known problems occurring in subject-based indexing, such as ambiguity of terms and their meanings, remain. The goal of INSEMTIVES is thus to motivate end-users to contribute more formal annotations at a higher level of expressivity that, for instance, refer to concepts from ontologies or impose a higher structure.

Purpose

The purpose of this deliverable is to investigate existing models for representing annotations, to analyze their different characteristics, forms, and functions. It furthermore looks into media-type specific annotation models which are relevant for the use cases of the INSEMTIVES project. The deliverable summarizes the state of the art in this area and presents a comprehensive list of requirements to be fulfilled by an annotation representation model to be used in the INSEMTIVES project.

Scope

This deliverable has two main objectives: first to research means on how to represent annotations in general and for different media types in particular. These include multimedia objects, Web services, and corporate knowledge assets such as documents, employee profiles and others. The second objective is to gather requirements by analyzing the prospected use of annotations in the use cases of the INSEMTIVES project.

2 Annotation Representation Models: State of the Art

The term annotation is used both as a noun denoting a piece of additional information and as a verb referring to the process of creating this additional information. The first uses of annotation appear much before the digital representation of information, for example, notes manually added over a printed text, or circle added over an image. More complicated examples include annotations in a book, footnotes and others. Even in these non-digital examples, we can see patterns of the way annotations are expressed and linked to the original content. These patterns come as a natural result of intensive and collaborative work which intensifies the need of common standards and models, that can be called annotation models. Obviously annotations models are content and media dependent and can also be influenced for the purpose of the annotation process.

The remainder of this section is structured as follows: First we present a classification scheme for annotation representation models which classifies them according to their structural complexity used, the type of vocabularies, and their collaboration model. Second we provide an overview of the state of the art of annotation approaches in the targeted domains of the INSEMTIVES project.

2.1 A Classification of Annotation Representation Models

An annotation model defines the actual form in which the annotation, that is, additional information is expressed, and how it is linked to the original content being annotated. Further it may include various characteristics related to provenance, such as information about the author (be it a human or a machine) or a time stamp, links to external structured or unstructured content, or it may specify the part of the original content that is annotated.

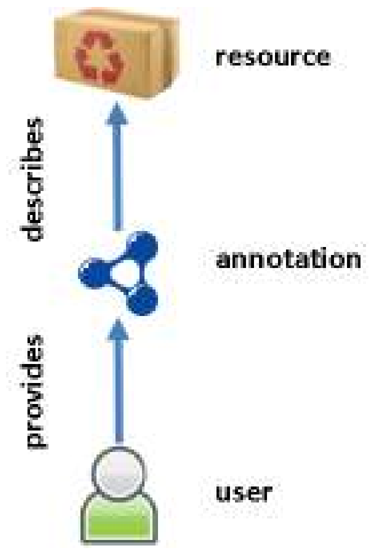


Figure 1: A Generic Annotation Model

A generic annotation model consists of a *resource* which is described with a piece of *annotation* provided by a *subject* (human or not) (see Figure 1). In the digital context, a resource is any artifact identifiable with an unique identifier, e.g. an Uniform Resource Identifier (URI). Examples of resources in the context of the Web are electronic documents such as images or text documents or their parts (e.g., a reference to an entity from a text document or an area in an image).

In this section we provide a classification of annotation representation models by describing a mutually independent set of features of the annotation and user elements in the generic annotation model. The categories were selected based on the intention to describe annotation representation models along two main dimensions:

1. the level of their formality and structural richness which largely conditions the extent to which semantic technologies can be applied,

2. the required level of user involvement in the annotation process which serves as an indicator of how much the user needs to be incentivised in order to support a given annotation model. Apart from this, we

discuss how much third parties such as, for example, domain experts, may need to be involved in order to support a particular annotation model.

Within each category we provide a specification of the most typical approaches by providing a description, comparative analysis of their advantages and disadvantages, and examples of systems. The goal of the comparative analysis is to show the differences in the trade off between the level of user involvement and provided services in each approach. We conclude this section by providing a summary of approaches which exhibit properties belonging to different categories.

2.1.1 Structural Complexity

The annotation element in the generic annotation model (see Figure 1) can be characterized by the different levels of structural complexity, discussed below. These are *tags*, *attributes*, *relations*, and *ontologies*.

Tags

DESCRIPTION a tag annotation element is a non-hierarchical keyword or free-from term assigned to a resource [123]. A tag *implicitly* describes a particular property of a resource as the computer and other users do not know the meaning that the annotator intended except if the natural language used is unambiguous. Normally, a tag is a single word or a sequence of characters without spaces, which typically serve as tag separators in the user input. Examples of tags include: the name of the person on a picture, the name of the place where a picture was taken, or a topic of a news article.

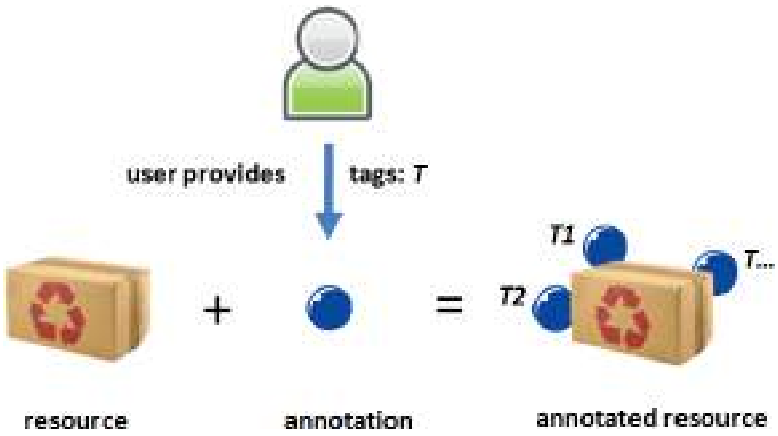


Figure 2: The Tag Annotation Model

PROS tags are a concept which is well-familiar for Web 2.0 users and, therefore, require nearly no learning curve for a typical user in order to start using them. They allow the user to easily annotate a Web resource with a tag and find other resources which were annotated with the same tag by browsing or searching. In fact, after four years of its existence Flickr reported to have “about 20 million unique tags” (Jan 08) and 3 billion images (Nov 08) [28].

CONS tags represent a minimal annotation model from the structural complexity point of view and, therefore, can enable only a limited number of services mainly focused on basic retrieval and browsing (e.g., retrieve resources that were assigned tag x). Because they only implicitly describe resource properties, tags are subject to ambiguity in the interpretation of these properties. For example, natural language tag “John” attached to a picture does not specify whether John is a person on the picture or if he is the photographer who took the picture.

APPLICATIONS Delicious [21] is a social bookmaking service that allows its users to memorize and share URLs of web resources such as blogs, articles, music, video, etc. It was founded in 2003 and now counts five million users and 150 million bookmarked URLs. The key idea of the service is that its users can access their bookmarks from any computer, for example at home, at work or while traveling. Delicious has been among the first popular systems that used tagging for organization and retrieval purposes. One of the important success factors of Delicious was its simplicity of use (see Figure 3). In a nutshell it works as follows: the user finds an interesting website and decides to bookmark it in Delicious; when adding the website the user assigns a set of terms that describe the website. Later, in order to retrieve a saved bookmark the system is queried with one or more of the previously assigned tags. If the answer set contains too many elements, it can be refined by adding more terms to the query. Delicious allows the user to browse not only a personal user space of resources, but also to find bookmarks saved and annotated by other users. Tagclouds (see Figure 4) and tag based search are starting points for navigation in the space of all (published) bookmarks of all users. While browsing bookmarks the system visualizes tags which were assigned to resources by other users, thus, allowing the user to use them as well as tags in a tagcloud in order to retrieve resources.

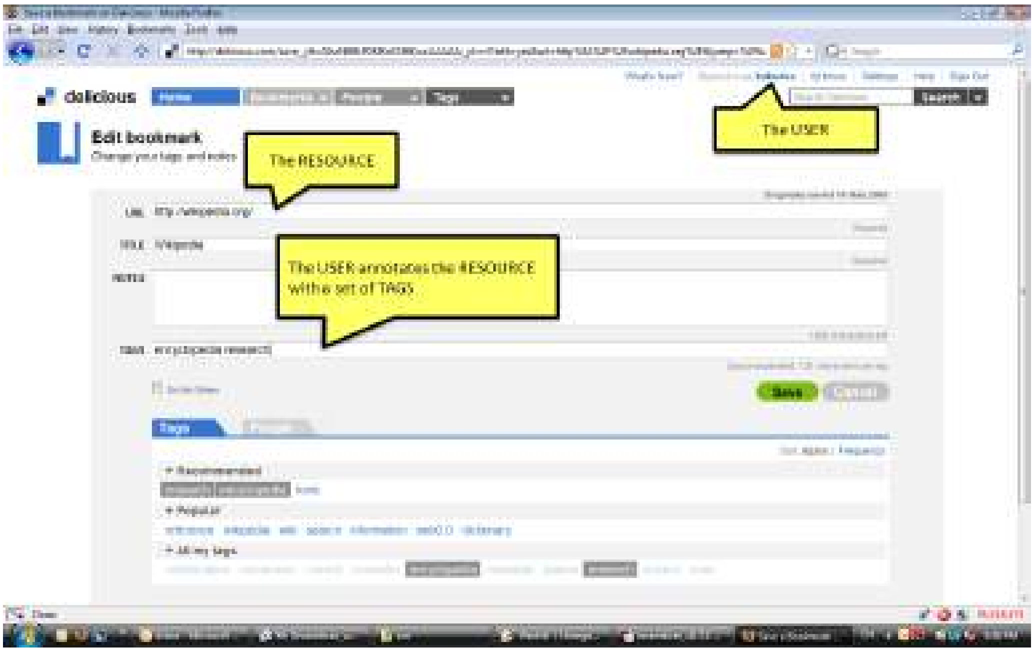


Figure 3: Annotating a Resource Using the Tag Annotation Model

Flickr is a free image hosting service that stores over three billion images [27]. It was launched in 2004 and popularized the concept of tagging together with Delicious. Flickr allows its users to upload their photos, organize and share them using tags. Even if, the users can establish relationships, form communities, comment and annotate photos of each other, the site is more used as a user's personal photo repository.

Another example of a popular social site which uses tags is Last.fm [57]. It is a UK-based Internet radio, founded in 2002, which has a thirty million active users in more than 200 countries. Last.fm allows its users to create their custom playlists and radio stations from audio tracks available from the service's library, and listen to them or any individual track on demand. It also offers numerous social features and can recommend tracks similar to the user's favorites. Last.fm users can tag resources such as a bands an artist, an album, a track and retrieve them using tag based search.

Examples of other systems that use tags for annotating their resources include: YouTube [125], CiteULike [19], LiveJournal [59]. All of these systems and the systems described above essentially use the tag annotation model with a slightly different modality of its application.

Attributes

DESCRIPTION an attribute annotation element is a pair $\langle AN, AV \rangle$, where AN is the name of the attribute and AV is the value of the attribute [121]. The attribute name defines the property of the annotated resource

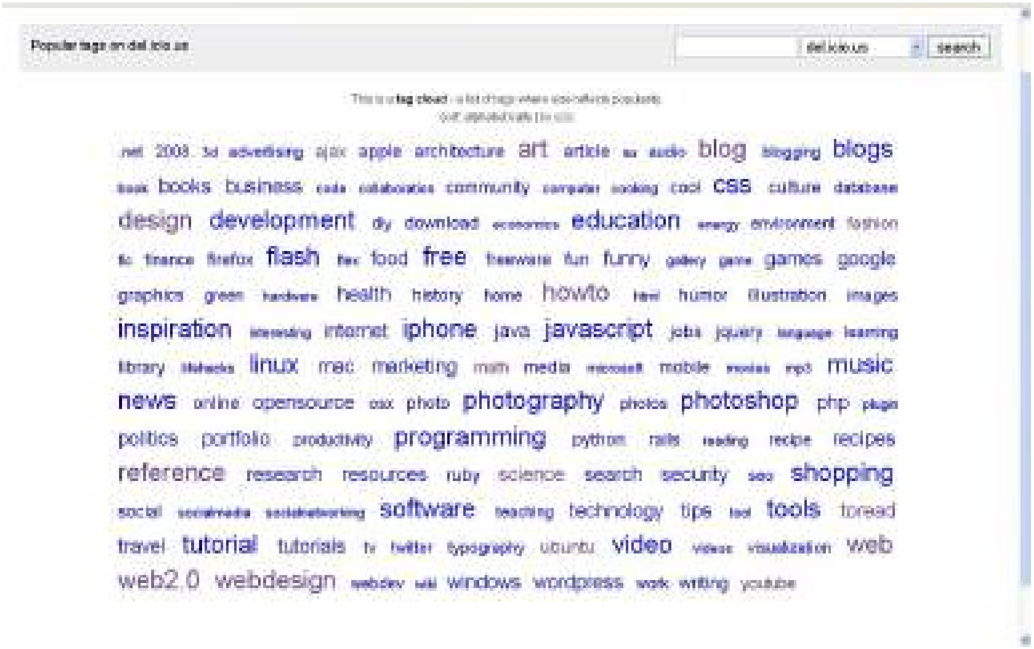


Figure 4: Delicious Cloud of Tags

(e.g., "location", "event", "starting date") and the attribute value specifies the corresponding value (e.g., "Trento", "birthday", "April 1, 2009"). Apart from this, the model allows us to define the data types for attributes and, therefore, enables type checking at query time. For instance, a "date" attribute defined on the STRING data type in a query will *not* be matched with a "date" attribute defined on the DATE data type in the annotation database.

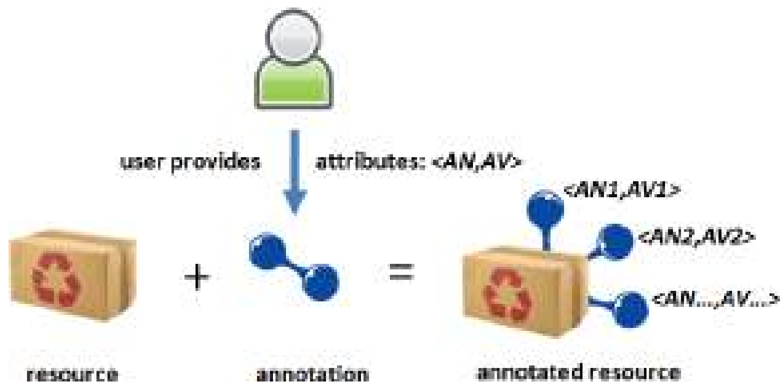


Figure 5: The Attribute Annotation Model

PROS attributes are pervasively used on the web and in desktop applications and, therefore, represent a well-familiar concept for end users. Differently from tags, attributes explicitly define the described resource properties and, therefore, enable a richer resource annotation and query language. For example, one could search for images of the Eiffel Tower taken between 1890 and 1900 by a specific photographer.

CONS while enabling more services than tags do, attributes are a still limited means of annotation because they refer to single resources and, therefore, cannot be used to effectively enable services which are based on interrelationships that exist between resources (e.g., search and navigation between related resources). On the other hand, attribute annotations require a more metadata-knowledgeable user than tag annotations do.

APPLICATIONS one of the earliest systems that used attributes for resource annotation was Semantic File Systems described by Gifford in 1991 [36]. The system allows the user to assign arbitrary number of name-value pairs to the user's files and then retrieve them by creating so called virtual directories. The user creates virtual directories at runtime specifying the list of attributes. According to the user input the virtual directory contains only those files whose attributes match attributes from the list. The implementation of ideas introduced in Semantic File systems can be found in operating systems search engines such as, for instance, Spotlight [122].

The Web has many examples of popular systems using attributes. Almost all social networks, (e.g. Facebook [25], Myspace [72]) consider the user as a resource and use the attribute annotation model to represent user profiles. The variety of attributes is very big and includes common attributes such as "Personal Info", "Contacts" as well as specific attributes, such as "Interests", "Traveling" (see Figure 6). Online markets such as Ebay [24] use the attribute annotation model to annotate resources which are items to sell. The seller can assign an item with attributes such as "item location: Trento", "item price: $100", "shipping to: Italy", etc.

Considering a user as a Web resource we can assume that users will be encouraged to annotate resources if they see the benefits of making them visible to others.

Figure 6: Resource Annotation Using the Attribute Annotation Model

Relations

DESCRIPTION a relation annotation element is a pair $\langle Rel, Res \rangle$, where Rel is the name of the relation and Res is another resource (i.e., different from the one being annotated). The relation name defines how the annotated resource is related with Res. In a sense, the relation annotation model is an extension of the attribute annotation model to the domain of resources, which allows the user to interlink these resources. For instance, in a scientific paper a citation referencing another paper is an example of a relation annotation which defines a relation between these documents.

PROS relation annotations provide a way to interlink various resources through typed links (i.e., relations). It allows the user to navigate from one resource to another and enable search and navigation based on these relation links.

CONS the user is expected to bear a higher mental load w.r.t. the previous models as, instead of describing one

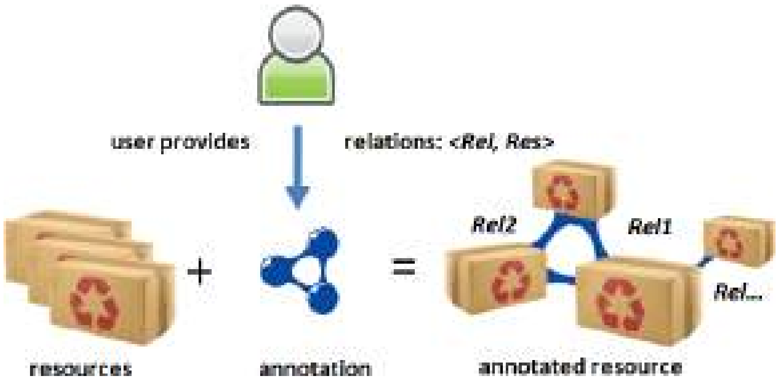
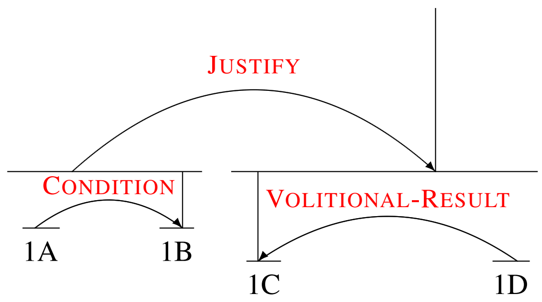


Figure 7: The Relation Annotation Model

resource, the user has to understand what the two resources are about and what kind of relationship holds between them.

APPLICATIONS Social network users annotate their profiles establishing relationships between each other, which allow them to find friends and to meet new people by navigating the network of relations between users. Apart from this, some social networks, for instance Facebook, allow its users to annotate photos with links to profiles of people depicted on them (See Figure 9). In 2007, Facebook had around 1.7 billion uploaded photos with around 2.2 billion relation annotations.

The relation annotation model can also be used to define relations within a resource. For example, the Araucaria project [85] annotates the rhetorical structure of a document using the RST relations annotation [61]. Figure 8 illustrates an example of intra document annotation of the RST relations; segments of text are given an identifier and links between each segment are annotated with a type of argumentation relationship (justify, condition, etc.).



[And if the truck driver's just don't want to stick to the speed limits,]^{1A} [noise and resentments are guaranteed.]^{1B} [It is therefore legitimate to ask for proper roads and speed checks.]^{1C} [And the city officials have signaled to support local citizens.]^{1D}

Figure 8: Rethorical Structure Theory Relationship

Upcoming.org is a social website for listing events that can be linked to Flickr photos by annotating them with so called triple or machine tag [124]. Triple tags use a special syntax to define extra information about the tag making it machine readable. For example, a tag like "upcoming:event=428084" assigned to a photo encodes that it is related to Upcoming.org event identified with "428084", therefore making these resources interlinked. Last.fm is another example of a system that use triple tags to link its tracks to Flickr photos.

Freebase is a large knowledge base containing around five million various facts about the world [30]. It is described as "an open shared database of the world's knowledge" and "a massive, collaboratively-

Figure 9: Resource annotation using the relation annotation model

edited database of cross-linked data."[1]. It allows its users to annotate resources (e.g., images, text, Web pages) using the Freebase annotation schema. The schema defines a resource annotation as a collection of attributes, its kind (e.g. person, event, location), and typed relations with other resources. It provides a convenient way to perform search and navigation in the space of resources allowing the user to find them using attributes, relations, and/or schema kinds.

Ontologies

DESCRIPTION This model is based on the notion of semantic annotation [80], the term coined in early 2000s, and applied ever since. It describes both the process and the resulting annotation or metadata consisting of aligning a resource or a part of it with a description of some of its properties and characteristics with respect to a formal conceptual model or ontology. As defined by Studer et al., "an ontology is an explicit specification of a (shared) conceptualization" [103]. In practice, ontologies are usually modeled with (a subset of) the following elements: concepts (e.g., `CAR`, `PERSON`), instances of these concepts (e.g., `bmw-2333 is-instance-of CAR`, `Marry is-instance-of Person`), properties of concepts and instances (e.g., `PERSON has-father`), restrictions on these properties (e.g., `MAX(PERSON has-father) = 1`[2]), relations between concepts (e.g., `PERSON is-a BEING`), relations between instances (e.g., `Marry has-father John`), etc [39]. The ontology annotation model allows the user to describe and interlink existing resources by qualifying resources as concepts or as instances and by defining relations, properties, and restrictions that hold between them.

For the purposes of this definition, semantic annotations can be divided in two classes with respect to expressivity: (i) reference or mention annotations and (ii) descriptive semantic annotations. The main difference is whether the annotation just points to a semantic description of a resource or entity (i) or the annotation itself contains semantic descriptors, relations and characteristics (ii). With respect to structural complexity we consider class (ii), whereas class (i) is described in Section 2.1.2.

[1]cf. http://www.crunchbase.com/company/metawebtechnologies
[2]The meaning of this rather informal notation is that any instance of the concept PERSON may have one father at most.

A good example is a mentioning of an entity in the text being annotated with its URI in a knowledge base with respect to a formal ontology, such as Spain annotated with prefixNS:CountrySpain. The same example can be extended to the descriptive model, where the annotation itself contains some characteristics of the annotation such as: partOf URI_Europe; type URI_Country; hasPopulationCount N; and so forth, where each of the properties used are defined in an ontology with further semantic descriptions, which are derived, used and can serve as basis for reasoning later on in the usage of the semantic annotations.

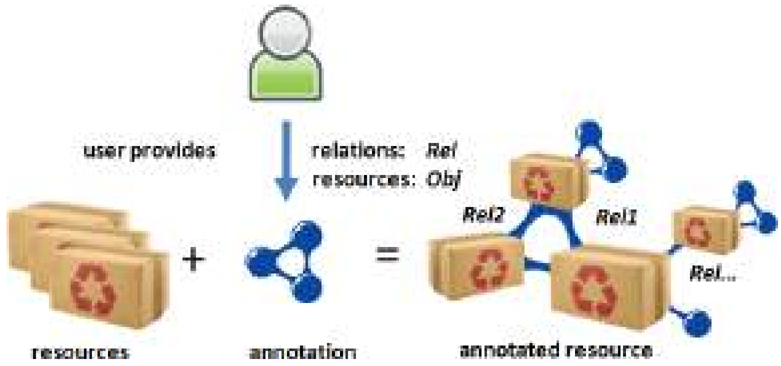


Figure 10: The Ontology Annotation Model

PROS ontology-based annotations or "semantic annotations" describe a resource with respect to a formal conceptual model, allowing meaning-bearing links between structured and unstructured data (such as an ontology and a text). This empowers a whole new range of retrieval techniques, which can be based on the knowledge schema expressed in the ontology, benefit from reasoning, co-occurrence of annotation or entities in the same resource or context, as well as combine this with unstructured data specific types of retrieval, such as full text search (FTS) in information retrieval (IR). The actual evidence metadata is encoded in the annotation and usually expresses metadata automatically or manually generated about the resource. The ontology and the corresponding instance bases capture background knowledge about a domain. The combination of the evidence based information about the resource and the background knowledge, allows indexing techniques, which are based on resource URIs as modeled in the ontology, ensuring retrieval and navigation through each of its characteristics (for example lexical representations such as NYC and New York City will be indexed as a single resource, despite their superficial differences, and this will lead to results containing the string "New York City", even though the user provides a query such as "NYC").

CONS Each of the annotation models described in order of increasing complexity, presents new challenges to human annotators, although disclosing richer potential for automatic processing. Semantic annotations, being the most sophisticated of this row, are no exception. The main challenges semantic annotation presents are in two major lines, namely (i) usability, and (ii) maintenance of the conceptual models. The usability aspect is key to human involvement in the generation of semantic metadata and is also going to be the main borderline that needs to be crossed to allow the approach weave in all forms of user interaction with software and data. One of the obvious reasons is the inherent richness and dynamics (as opposed to a predefined and fixed schema) of ontologies. There one can experience the challenge of presenting multiple taxonomical and subsumption structures over the same model, as well as multiple description facets of each resource. Empowering a human to find its way to the right class, entity or relationship he/she wants to cite, is a serious challenge to usability experts and visual interface designers. Even more, real world applications involve multiple instance and fact bases, and their corresponding ontologies, inter-aligned, resulting in millions or even billions of individual entity descriptions, exposing a challenge in terms of scalability, intuitive search and auto-suggest/complete methods, as well as summarization of entire knowledge bases to any useful level of granularity and with respect to any criteria. Another complicated task ontology provider needs to face is the maintenance and update of the knowledge, often coming from external sources, its syntactic and semantic alignment, and often, its challenging scale (e.g. bio-medical knowledge bases with billions of individual facts).

APPLICATIONS OntoWiki [8] is a free, open-source semantic wiki application, meant to serve as an ontology editor and a knowledge acquisition system. It allows its users to annotate Web resources representing them as concepts and instances of concepts. Different attributes can be assigned to a resource and interlinked with other resources, therefore, describing its characteristics. OntoWiki retrieval functionality allows to search and to generate different views and aggregations of resources based on concepts, attributes and relations.

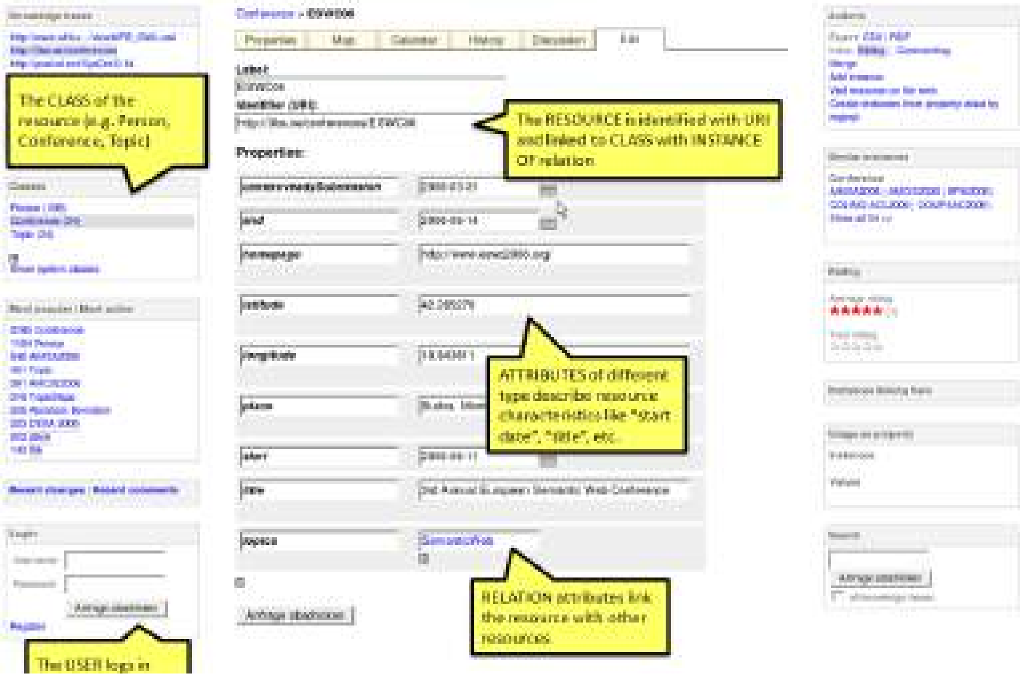


Figure 11: Resource Annotation Using the Ontology Annotation Model

Semantic Wikipedia [112] is an extension of Wikipedia [120] that allows its users to annotate wikipedia articles defining an article as a class like "person", "document" or an instance of previously declared class. The users can also interling articles by annotating them with typed links like "author" or "was born in". This brings semantic enhancement to Wikipedia that allows browsing and searching in a more specific ways, such as "Give me a table of all movies from the 1960s with Italian directors."

KIM is a semantic annotation, indexing and retrieval platform developed by Ontotext in 2002 [78]. It can mainly used o facilitate automatic semantic annotation on top of different content types, with built-in extended capabilities for semi- or non-structured text based on the GATE framework [29]. It is deployed on top of a native semantic database engine, currently OWLIM and/or Sesame [79, 1]. KIM has been applied in various domains, some of the most interesting ones being anti-corruption and asset recovery, and analysis of bio-medical content, such as clinical study reports or scientific papers. As a result of the semantic annotation and indexing, it provides multiple retrieval paradigms, using the content, its structure, and document-level metadata, along with entity co-occurrence and graph-like patterns on the ontologies and knowledge bases to return either sets of entities or documents, or track trends and calculate timelines using content's temporal aspect. It allows both of the above described kinds of semantic annotations, in the same time providing support for the simpler annotation models explained earlier and in the next section. Despite the fact it has matured as a platform since the early 2000s and is in active use, it has never, so far, been focused on usability or visualization aspects of the manual annotation process.

2.1.2 Vocabulary Type

When annotation elements (e.g., tags, attribute names and values, relation names) are provided by the user in the form of a free-form natural language text, these annotations unavoidably become subjects to the semantic heterogeneity problem because of the ambiguous nature of natural language [38]. Particularly, we identify the following three main issues related to the annotation process[3]:

1. **Polysemy:** Annotation elements may have ambiguous interpretation. For instance, the tag "Java" may be used to describe a resource about the *Java island* or a resource about the *Java programming language*; thus, users looking for resources related to the programming language may also get some irrelevant resources related to the Island (therefore, reducing the precision);

2. **Synonymy:** Syntactically different annotation elements may have the same meaning. For example, the attribute names "is-image-of" and "is-picture-of" may be used interchangeably by users but will be treated by the system as two different attribute names because of their different spelling; thus, retrieving resources using only one of the attribute names may yield incomplete results as the computer does not know the synonymy link;

3. **Specificity gap:** This problem comes from a difference in the specificity of terms used in annotation and searching. For example, the user searching with the tag "cheese" will not find resources tagged with "cheddar[4]" if no link connecting these two terms exists in the system.

Typically, the problems described above are addressed using so called *controlled vocabularies* – "*organized lists of words and phrases, or notation systems, that are used to initially tag content, and then to find it through navigation or search*" [117]. The basic underlying idea of how controlled vocabularies can be used for the annotation and search is as follows: by interpreting and linking user annotation terms to elements of a controlled vocabulary (e.g., words, concepts), the meaning of these terms is disambiguated (see Figure 12).

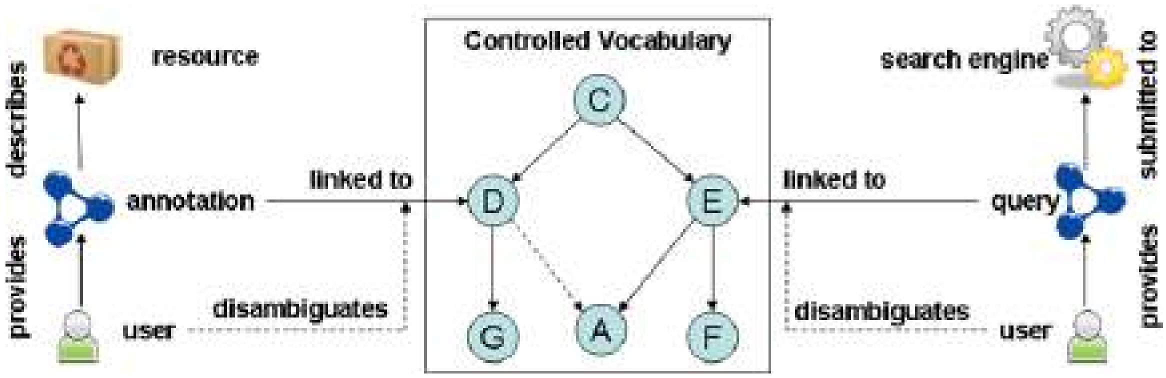


Figure 12: Controlled Vocabulary based Annotation and Search

Controlled vocabularies are normally built in a top-down fashion by a (small) group of experts and are often compared and contrasted with the so called *folksonomies* [116] – classification systems which are generated in a bottom-up fashion through collective and massive annotation made by a (large) group of users [40, 69]. Despite this difference, it has been argued that folksonomies largely conform to the requirements formulated for controlled vocabularies [101] and that both can co-exist and complement each other [2, 69, 40]. Given their complementary nature and the purpose of this section, we do *not* draw a difference between folksonomies and controlled vocabularies and we rather focus on differences in their models from the point of view of the resulting vocabulary structure and on their comparative advantages and disadvantages. In Section 2.1.3, we discuss models which differ in the way *how* the controlled vocabularies are actually built (e.g., top-down vs. bottom-up fashion).

Noteworthy, the term *controlled vocabulary* comes from the Library and Information Science community and in the Semantic Web community the term *ontology* is more used instead in order to describe similar kinds of knowledge organization systems. Hereinafter we will use both terms interchangeably.

Ontologies are often classified based on the level of their expressivity (or formality) which conditions the extent to which a certain form of ontology can be used in automated reasoning. Smith and Welty proposed

[3]Note that there exist also other problems related to the annotation process that are based on natural language input issues (e.g., plurals vs. singular forms, misspellings). Because they are not directly related to the annotation representation models, these problems are out of the scope of the current deliverable; interested readers are referred to [38] for details.

[4]which is a kind of cheese

such a classification shown in Figure 13 [98]. They note that there is a point on the scale (marked as a black bar) where automated reasoning becomes useful: this is where the ontology can be reasoned about using the subclass relations [69].

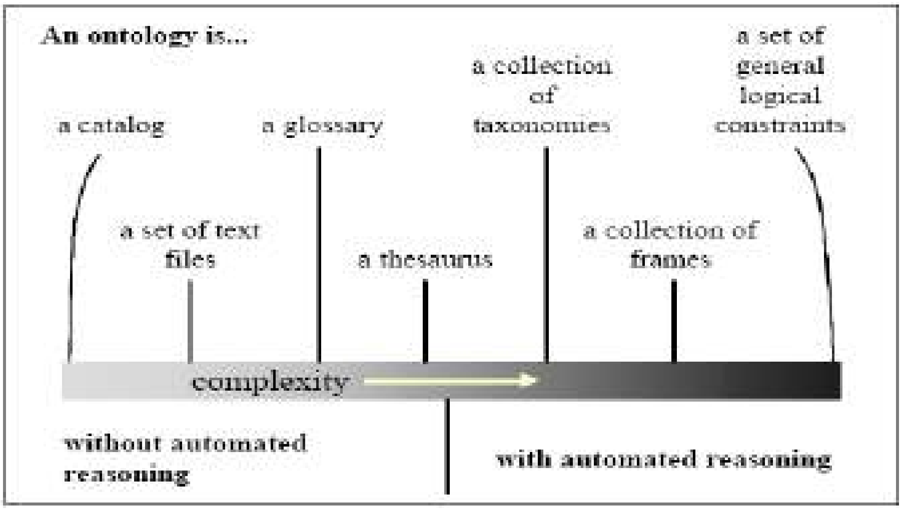


Figure 13: A Classification of Various Forms of Ontologies According to their Level of Expressivity (Formality) [98].

Note that the way how ontologies can be used to *annotate* resources (as described in Section 2.1.1) is different from the way how ontologies can be used to *support* the annotation process (as described in this section). In the former case, users build ontologies by providing pieces of it as annotation elements. For example, by linking a page about Napoleon to a page about people with ontological relation `is-instance-of`, the user annotates the page about Napoleon with an ontology annotation element. Such elements (possibly provided by different users) are then assembled into a bigger ontology which can be seen as a complex annotation structure used to describe the annotated resources. In the latter case, the users provide (simpler form of) annotation elements and (semi-automatically) map them to the background ontology (see Figure 12). For instance, the user might tag a web page with the term "dog" and link it to the concept `DOG` in the background ontology. It is worth mentioning that both approaches can potentially enable automated reasoning.

In the rest of this section we discuss three kinds of annotation representation models which differ in the kind of vocabulary they are based on. In the selection of the models we followed the principle of staying within the spectrum of so-called *lightweight ontologies* [37] in order to conform to the goal of the work package to use "lightweight, structured knowledge". The first model is not based on a controlled vocabulary, whereas the second and the third ones are. These last two models fall on the left and on the right side from the vertical bar shown in Figure 13 respectively. We show how the issues described above can be addressed in the last two approaches.

Uncontrolled Vocabulary

DESCRIPTION In this model, annotation elements are not linked to controlled vocabulary terms and, therefore, are subjects to at least the three problems described in the introduction to this section. In other words, no controlled vocabulary exists as part of the model and free text is used instead. In principle, search in this model is implemented as syntactic matching of strings representing query terms with those used in annotation elements.

PROS The advantage of this model is that the user does not need to know about the existence of a controlled vocabulary and does not need to be involved to help resolve ambiguities in linking annotation elements to controlled vocabulary terms.

The user involvement at annotation time is minimal as the simplest interface can be used: the user enters a list of free-text words in a text box.

CONS The main disadvantages of this model are the three problems described in the introduction to this section: polysemy, synonymy, and the specificity gap.

Moreover, the minimal involvement at annotation time translates into a higher involvement required at the search and navigation time. The problem of heterogeneity described earlier produces a higher level of noise in the results of search and the users need to filter out manually the results they are interested in. At navigation time, the lack of explicit structure of the annotation provides little help and only simple clustering can be performed (such as the one illustrated in the Delicious tag cloud in Figure 4).

APPLICATIONS Delicious [21] and Flickr [27] are both good examples of totally uncontrolled vocabulary models. The users can enter any free-text tag they want to annotate their bookmarks and photos (respectively). The systems do not check for the validity of the tag or if it maps to a known meaning.

Authority file

DESCRIPTION This model is based on a simple form of controlled vocabularies called the *authority file*. In this controlled vocabulary, synonymous terms are grouped into concepts and one of the terms is selected as the concept name and used for visualization and navigation purposes (this term is called the *preferred term*) [33]. Each concept may have one or more associated terms and each term may belong to one or more concepts. In this model, user annotation elements are mapped to the controlled vocabulary terms and, consequently, to concepts which uniquely codify the meaning of the annotation elements. For example the tag “automobile” could be mapped to a concept with the preferred term “car” in the controlled vocabulary. Note that when a term belongs to more than one concept, the user may need to be involved in disambiguating the concept selection at annotation or query time.

PROS the model allows us to resolve the polysemy and synonymy problems by linking annotation elements and the user’s search queries to a predefined meaning (i.e., to a concept) in the controlled vocabulary. For example, if the user searches for the term “java” that is mapped to “programming language” in the controlled vocabulary, the search will not return resources annotated with the term “java” that refers to the “island”.

The user will also be able to navigate easily through all the documents related to a “programming language”, for example, without having to know the particular terms used to annotate these documents.

CONS Both at annotation time and at search time, the user will have to get involved more. At annotation, if there is an ambiguity in the mapping from the term the user provides and multiple concepts in the controlled vocabulary, the user will need to choose which concept the term refers to. At search time, again, if there is an ambiguity in the terms the users search for, they will have to specify explicitly which concept they are referring to. Alternatively, the user may decide to leave the concept disambiguation task to the system which could use the default concept for an ambiguous term (e.g., the one which is used more often) or use all its concepts for the annotation and search.

A second issue with this type of controlled vocabulary is the creation of the *authority file*. Each term in the vocabulary has to be mapped to one or multiple *preferred term* for the vocabulary to be useful; this requires time from the vocabulary creator(s). In addition, the coverage of the vocabulary will rarely be as wide as the full human vocabulary and the user might want to use a term that is not present in the vocabulary. Thus, this type of vocabulary is usually limited to applications in specific domains.

APPLICATIONS Faviki [26] is a social bookmarking system similar to Delicious. However, when users enter a tag, they are asked to disambiguate its meaning to a known concept. These concepts are automatically extracted from the contents of Wikipedia. This disambiguation allows a better convergence between the terms used for the annotation by different users as well as it proves better results at search time.

Taxonomy

DESCRIPTION This model is based on a controlled vocabulary called a *taxonomy*. A taxonomy is an extension of the authority files controlled vocabulary with the broader term (BT) and narrower term (NT) relations which are defined between controlled vocabulary concepts [33, 58]. See Figure 15 for an example of a

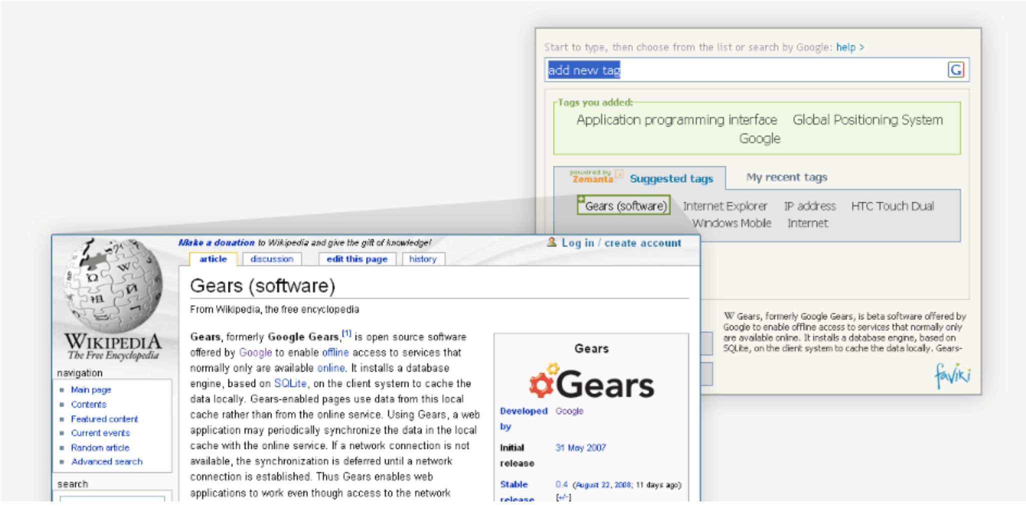


Figure 14: Faviki: Tagging with Concepts from Wikipedia (screenshot provided by faviki.com)

taxonomy. Intuitively, a BT/NT relation might correspond to the *is-a* or *part-of* relation. For example, the concepts "cat" and "dog" can be linked to a higher level concept "mammal" which is itself linked to a concept "animal". Another example is a geographical taxonomy, like the one in Geonames, where countries are linked to the regions composing them which are in turn linked to cities, etc. [35]

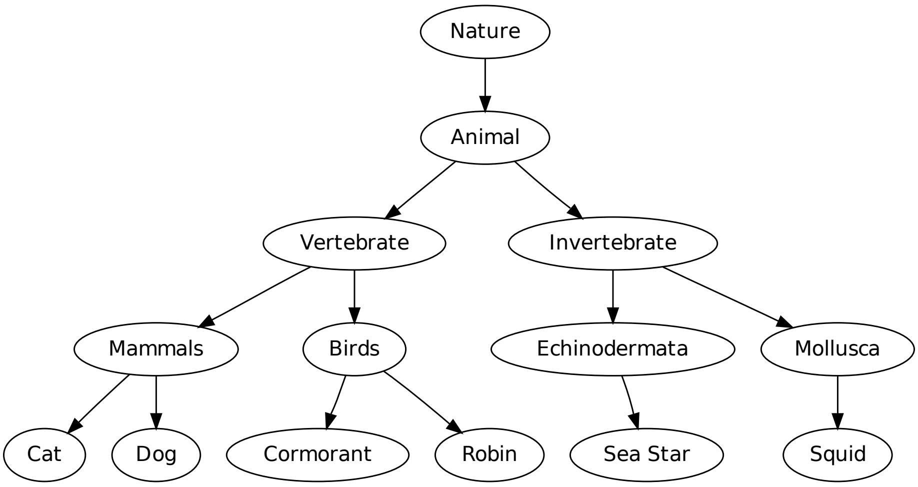


Figure 15: Example of a Taxonomy

PROS in principle, this model allows us to solve the specificity gap problem because query terms can be mapped to user annotation elements through BT and NT relations and, therefore, resources which are more specific in meaning than the user query can also be retrieved. For example, the user can annotate two different resources with "cat" and "dog" and then retrieve both resources by searching for "mammal".

CONS In terms of user involvement, the taxonomy annotation model has the similar drawbacks as the previous model. The user will have to make sure that the annotation and search terms correspond to the correct node in the taxonomy if there is an ambiguity.

The taxonomy creation time is also an issue as it requires specific time, from an expert or from the actual user to create a meaningful classification of terms that can then be used for annotation.

Another disadvantage of taxonomies is the problem of the vocabulary granularity. If a third party (an expert) produces the taxonomy to the user, there is no guarantee that this taxonomy will be detailed enough for the user. For example, in the biological classification example earlier, the expert might not include the "vertebrate" vs. "invertebrate" split. When the user wants to perform a generic search for all "vertebrate", only a search for the higher term "animal" will be possible. In opposition, if the taxonomy provided by the expert is too detailed, the disambiguation task at annotation and search time will be more complex for the user that might not know all the differences between the same meaning of terms or might have to search between many different – but very close in sense – meanings.

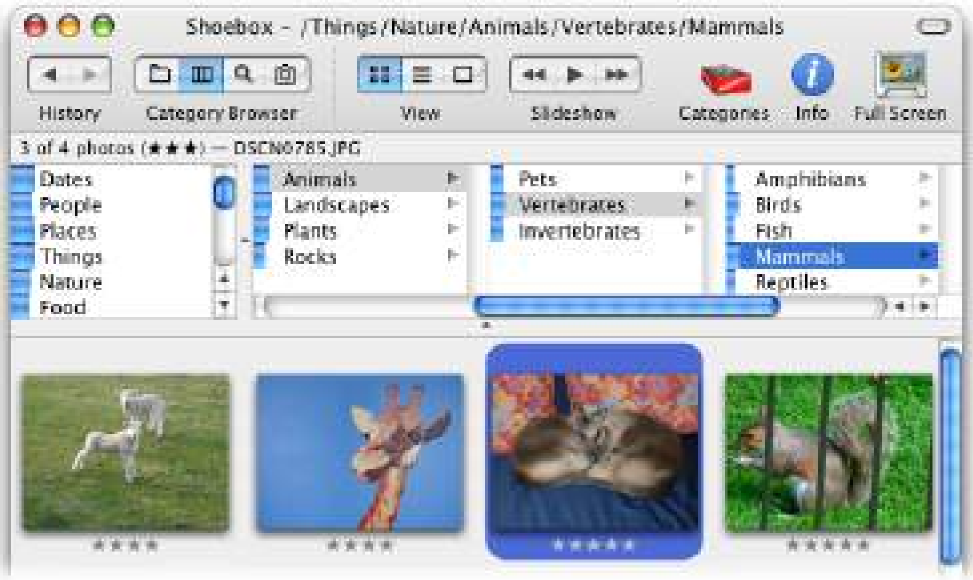


Figure 16: Searching of Photos by Categories in Shoebox: Here Things > Nature > Animal > Vertebrate > Mammals.

APPLICATIONS Shoebox is a photo management software that allows the users to annotate photos with annotations which are nodes in a taxonomy [94]. The software comes with predefined categories (such as countries/regions/cities, or animal kingdom) but the users can also freely extend the taxonomy by creating their own categories and sub-categories. The user can then search for all photos tagged with an annotation corresponding to a category or one of its sub-categories as illustrated in Figure 16.

2.1.3 Collaboration Type: Single User vs. Collaborative Annotation

In this section we will describe a dimension related to how users contribute to the creation of different types of annotations and the vocabulary used in the process (see Section 2.1.2). We can roughly distinguish between two approaches:

1. the single-user model, where a single user performs the task of either annotating resources or creating the vocabulary (or both),
2. the community model, where a set of users collaborate in the task of either annotating resources or creating the vocabulary (or both).

In this section we will focus on the interactions between users, and how these interactions affect, positively or negatively, the other two elements of our classification of annotations; these elements are the structural complexity of the annotation, i.e., how users interact with each other using some degree of structured annotation (see Section 2.1.1) to annotate resources, and the type of vocabulary used to annotate the resources (see Section 2.1.2).

Considering the annotation perspective, in the single user model we can identify two sub-dimensions by considering who uses the annotations (see Table 1). A single user can annotate a resource for personal use only, or can annotate a resource considering that these annotations will be later used by many other users. In the remainder of this section we will explain how this consideration affects the annotation process.

Feature \ Type of annotation	Single User		Collaborative
	Private use	Public use	
User interaction	none	none	encouraged
Type of vocabulary	personal	personal	shared

Table 1: Types of Collaboration for Annotating Resources and Building Vocabularies

Considering the vocabulary perspective we will focus on controlled vocabularies, since in the free-form annotation model (uncontrolled vocabulary), anyone is free to use any term (even terms not corresponding to a vocabulary) and hence there is little point on studying how interactions between users affect the vocabulary (as it is free). We will make clear how the interactions between users affect the construction of controlled vocabularies in the remainder of this Section.

Single-user (Private use)

DESCRIPTION In this model a user annotates resources for personal purposes, usually to organize these resources for future search or navigation. Single users could also build their personal controlled vocabularies, but as we will see in the CONS part, the necessary work does typically not justify this task.

PROS The advantage of the *single-user annotation model for private use* is that each user is in full control of the annotations and since there is no sharing, the semantic heterogeneity problem (see Section 2.1.2) is reduced to the scope of the single user. One can also argue that since the annotator and the user of the annotation are the same person, the annotations will reflect the personal taste and knowledge of the user, which in turn can be translated into more accurate results at search time.

CONS In this model, each single user has to annotate all the resources that he wants to have annotated for future use. These resources can be private resources (e.g., local files), resources shared by other users (e.g., shared photos), o publicly available resources such as web pages. This annotation process has the following disadvantages:

- annotating all of the resources takes time and requires motivation (building personal controlled vocabularies to annotate the resources requires even more time and motivation). If the user has no strong incentives, the quality and coverage of the annotations and the controlled vocabulary cannot be guaranteed;
- the single annotator might not be an expert in the annotation process (or vocabulary building) and might miss relevant annotations (or terms) for a resource;
- users have to remember which terms they normally use to annotate a particular concept in a resource and use it consistently across the whole set of resources. If this is not done properly the user will not be able to find the desired resources at search time (see Section 2.1.2).

APPLICATIONS Bookmarking is a well known example of annotation of Web resources for personal use; currently most Web browsers include this feature by default.

Other examples of personal annotation systems are Picasa [81] and Shoebox, where users tag their photos with free-form keywords. Picasa also adds the possibility of geo-tagging photos while Shoebox provides a controlled vocabulary in the form of taxonomy for the annotation.

Another system for the annotation of Web resources is Zotero, where users have the possibility to manage Web resources related to scientific research by classifying them and/or by adding free-form annotations [126].

Single-user (Public use)

DESCRIPTION In this model, a single user (normally an expert, or a small set of experts representing an institution) annotates resources with the goal of organizing the knowledge for a broader set of users. Library catalogs are perhaps the most well known example of this kind of model.

Considering the construction of the controlled vocabulary, in this model also a single user (also normally an expert or a small set of experts representing an institution) builds the controlled vocabulary to be used by a broader set of users in the annotation task and later for search. In this case, library classification schemes or thesauri are well known examples.

PRO Many people benefit from the work performed by a few experts of the field; the work (either the annotation or the resulting controlled vocabulary) is considered to be of good quality resulting in a good organization of the resources. In addition, experts in the annotation task are also able to use more complex and well structured vocabularies (see Section 2.1.2) as they are better trained in this task. Well studied and consistent controlled vocabularies developed by experts are very useful for solving the problem of semantic heterogeneity (see Section 2.1.2).

CONS It is normally costly to have dedicated experts to annotate and organize resources or to build controlled vocabularies. Moreover, in highly dynamic domains, the time lag between the publication of the resource and its annotation can be considerable [12]. This delay can be due to, either the scalability of the approach (work overload of the experts) making the classification of vast amounts of resources in short times impracticable, or to the fact that the field of knowledge is new or very dynamic and thorough studies have to be performed in order to reach consensus on the right vocabulary to be used for the annotations. Another issue with this approach is that the experts who annotate the resources are neither the authors nor the users of the annotations, therefore the annotations might not always reflect the users' perspective or the authors' intention [13]. Some possible solutions to these issues, still at research stage, are being proposed; for example, to let annotators define their classes and keywords that will later be approved by domain and annotation experts [64].

APPLICATIONS Library catalogs, such as the Library of Congress [76] and many other libraries fall into this model of annotation. Library classifications schemes used in library catalogs are examples of expert designed controlled vocabularies with the purpose of classification of books. There are also more general purpose controlled vocabularies designed by experts such as WordNet [70].

Collective annotation

DESCRIPTION In the collective annotation model, a set of users share their annotations about publicly available resources. These annotations are later also used by a set of users (possibly larger than the set of annotators) for navigation and search. In this model, the workload of annotating resources is distributed (in contrast to the previous 2 models). The users do not necessarily work together in the process of annotation, neither do they have to reach an agreement on the resource annotations (but obviously this could be the case).

Collaborative Tagging (or social tagging) is a well known model to annotate resources[5] with free-text tags [38]. When a critical mass of user-generated tags is available, these annotations can be used to generate folksonomies [116]. The main characteristic of folksonomies (folk + taxonomies) is the bottom-up approach (i.e., from individuals to groups) for reaching consensus on emerging concepts, without any a priory agreement or controlled vocabularies.

Considering the construction of controlled vocabularies, the basic idea is to let users collaboratively build a controlled vocabulary using a bottom-up approach. This approach is reflected in several proposals in the literature. These proposals differ in whether there should be a pre-defined controlled vocabulary which will be collaboratively extended [88], or the controlled vocabulary should emerge bottom-up from scratch [69] [12]. Depending on the model, users could work together in order to reach agreements in the construction process [12], or they could work independently and an automatic process could address the evolution of knowledge [69]. Aberer et al. [2] proposed several generic characteristics and issues that any system for *emergent semantics* should consider and address. The notions presented in these proposals allows us to address disadvantages present in centrally built controlled vocabularies.

[5] mainly on the web.

PRO The main advantage of the collaborative annotation model and construction of the controlled vocabulary is the distribution of the workload among all users of the system. This, in turn, lowers the barrier of user participation in the annotation process since one user can adopt the tags that were assigned to a resource by others — thus simplifying the annotation task —, or even, more importantly, can search and discover new resources that are annotated with the same tags as the user's tags [5]. Another advantage is that users can express their own views when annotating resources, which in contrast with annotations made by experts, could simplify search and discovery of resources for other users with the same interests, considering that these resources were annotated with many potentially different points of views and interests.

In addition, the mass of data provided through the annotations in a collaborative model as well as the users' relations with these annotations can be used for automatic extraction of new knowledge (new terms and relations in the controlled vocabulary); for example, a recommendation system can be built to propose relevant annotations for a new resource, based on existing annotations of similar resources [100] or to extract co-relations from unstructured annotation schemes [84, 89]. This model also provides the system with behavioural information about the users' interests through their interaction with other users' annotations and their own annotations. By using this information the system could infer a user model containing, for example, their education, background, interests or culture. Based on this user model the system can suggest and discover related resources or can discover hidden communities and sub-communities of users based not only on social relations (such as friends or coworkers) but also on interests and shared knowledge or background [69].

Building the controlled vocabulary in a bottom-up fashion can address the time lag problem for the inclusion of new concepts in the controlled vocabulary, therefore increasing steadily the precision and recall of any systems that is based in these kind of controlled vocabularies as these evolve dynamically and in direct response to the knowledge evolution of the involved users [2].

CONS Folksonomies and other collaborative tagging systems based on free-text annotations suffer from the lack of formal semantics that leads to the semantic heterogeneity problem, which makes indexing, search and navigation difficult [56]. This heterogeneity problem can be addressed by using controlled vocabularies or community-built controlled vocabularies that are better related to the domain of the resources. However, even with the use of these vocabularies, social annotation approaches could suffer from biases that express personal likes or dislikes of the users, and not the actual intention of the resource author, examples of these annotations are "best paper", "don't bother with this one", "to read". This type of annotations cannot be avoided since the primary goal of any annotation model is to allow users to find the resources they are interested in. These kind of biased annotations could be also important for ranking purposes of the resources as well as for the users. The system could try to limit such biased annotations by automatically separating the personal annotation from the general annotations or by avoiding the mix between tags that express biases and the ones that are not associated to the user's relative taste.

An important issue which remains open and needs to be treated in the community-built controlled vocabulary is the model to be used to allow knowledge to emerge from the contribution of each single user. The process of agreement (either automatic or user driven) in knowledge evolution is still an interesting research issue to be addressed [2, 12].

APPLICATIONS Diigo [22] and Delicious [21] are examples of applications implementing this model of annotation to build a shared collection of bookmarks. Flickr is a social website for sharing and organizing photos and videos, users can comment resources, tag or add them to their own personal favorite list. These tags are later used for search and navigation. Kolbitsch [56] proposed and extension for searching Flickr content where WordNet [70] is used to expand queries, using this approach users are able, for example, to find results about "automobiles" when searching for "cars".

The work presented by Gazan [34] incorporates the ideas of collaboration to library catalogs, where the users can become more than simple content consumers by annotating resources in digital libraries and become content creators. Facetag [83] is another initiative following this direction. The rationale in this system is to try to integrate top-down (i.e., from experts to individuals) and bottom-up classification in a semantic collaborative tagging system, incorporating the ideas of faceted classification, which helps users organize their resources for later search, navigation and discovery. Facetag incorporates the idea of collaborative annotation and collaborative controlled vocabularies in a single system.

Considering community-built controlled vocabularies, Braun et al. [12] proposed an ontology maturing process by which new knowledge is formalized in a bottom-up fashion going through several maturity stages: from informal emergent ideas using input from social annotation systems such as folksonomies, to formal lightweight ontologies via a learning process involving users. They also introduced SOBOLEO [99], a system for collaborative ontology engineering and annotation.

Another related work to community-built controlled vocabularies was presented by Mika [69] where a model for formalizing the elements in an ontology evolution scenario was proposed. The model consists of three elements, Actor-Concept-Instance. This model has a straightforward parallel to the model adopted in this section to explain the annotation process as seen in Figure 1; where Actor is equivalent to *users*, Concepts to *annotations* and Instances to *resources*.

Tagpedia [88] tries to build a general purpose controlled vocabulary to overcome the fundamental problems of Wordnet and Wikipedia as standard controlled vocabularies. The main issues with WordNet are the lack of knowledge about entities (specially people) and the lack of support for incorporating new knowledge. Wikipedia contains a lot of information about entities but suffers from the lack of a more formal/ontological structure [88]. The idea of Tagpedia is to initially mine Wikipedia to construct an initial set of *syntags* (as opposed to synsets of WordNet) and to allow users to extend this initial set of *syntags* dynamically in a bottom-up manner.

2.1.4 Summary

Note that each axis of the classification can be combined to create an annotation model, so for example, some system can use an attributes annotation model based on a taxonomy vocabulary for the attribute values. In Table 2 we provide examples of popular systems that mix the structural complexity with different vocabulary types.

Considering that the main motivation for a user for annotating a resource is to organize user's personal information for future use, we could state that distributing the workload of annotating shared resources between a community of users constitutes the primary incentive for users to participate in a collaborative model of annotation [5]. Nowadays (with few research exceptions [56]) most of the annotation models are based on free-form annotation approaches in the form of tags. We need to study the use cases of the project in order to assess whether a more structured form of annotation such as attribute-values could be more useful, and whether using controlled or community-built vocabularies are more suitable for the purposes of the use cases than just free text. Section 3 details the specific requirements of the use case partners with respect to the classification developed in this section.

	Tags	**Attributes**	**Relations**	**Ontologies**
Uncontrolled Vocabulary	*Flickr* annotates photos with free-text tags. *Delicious* annotates bookmarks with free-text tags.	*Facebook* allows users to describe their interests in books, movies, music with free-text fields.	*Freebase* allows its users to connect freebase topics with unlimited number of user defined relations.	*Semantic Wikipedia* users build ontology out of wikipedia articles by annotating them with categories, attributes and typed links that can be created on the fly.
Authority File	*Faviki* annotates bookmarks with free-text tags mapped to a controlled vocabulary built from wikipedia[a].	*The "Encyclopaedia of Life"*[b] provides a set of controlled attribute based annotations (species, genus, family, order, etc.) for users to tag photos and videos of flora and animals with their corresponding biological classification.	*Facebook* allows the user to annotate photos and videos with links to people present in these media. *Flickr* ↔ *Last.fm/Upcoming.org* by adding a special tag to a flickr photo, users can link it to the events (described on last.fm or upcoming.org) where it was taken.	
Taxonomy	*Shoebox* allows users to tag their photos with tags present in a taxonomy. The user can then do "generic" searches for all the photos tagged under a node in this taxonomy.		The *Rhetorical Structure Theory* provides an annotation model to describe the argumentation relation between segments of text. The relations are annotated with a type that can be chosen from a taxonomy of 25 different types of relations classified under two categories "nucleus ↔ satellite" or "multinuclear".	

Table 2: Examples of Systems in the Complexity and Vocabulary Dimensions.

[a] http://www.wikipedia.org
[b] http://www.eol.org

2.2 Annotation Approaches in the Application Domains of the Use Case Partners

A number of methodologies describing the process of annotation of various types of resources such as textual documents, images, graphics, or videos, were proposed in the last years (cf. [96]). Some of these methodologies are accompanied by software environments which allow the mechanization of specific annotation-related tasks; most notably for text or semi-structured data. Automatic means to create semantic annotations are seen as a viable alternative, in terms of costs, to human-driven approaches.

In this section we present examples of mostly human-driven approaches for the annotation of textual documents, Web services, and multimedia resources adhering to different categories of the categorization scheme presented earlier in this section.

2.2.1 Document Annotation

Manual annotation of text has been used by small professional communities (e.g. the linguistic community) for years, but it has not reached wider audiences. With the advance of Web 2.0 applications and the increasing dependency on metadata for various navigation and retrieval tasks, multiple approaches have been developed. These can be as simple as providing a keyword characterizing a textual resource, or as complex as associating a span of text with given start and end offsets with an annotation type, and a set of attribute - value pairs, such as in GATE [29], or even formulating the annotation as a reference to a named entity or a fact represented in an ontology or knowledge base, like the semantic annotation model of the KIM platform (cf. Section 2.1.1). Besides their expressive complexity, annotations may address different types of content: (i) document level annotations - usually, attribute - value pairs describing some characteristic of the document; (ii) inline or offset based annotations - usually denoting entity or relationship reference at a given offset; (iii) structural - annotations capturing the structure of a document - inherently hierarchical. In order to capture annotations coming from different annotating agents, or used for different purposes, a popular approach is to have named collections of annotations, usually called *annotation sets* (e.g. one capturing the original HTML markup, the other the annotations by agent A and another by agent A'). An interesting characteristic of inline annotations is that they can overlap and contain each other, which leads to representations of their co-positioning with graph models.

The first manual annotation frameworks were proposed around year 2000. Examples of such frameworks include Annotea [54] and CREAM [46], which were primarily implemented for the annotation of Web pages. Annotea specified an infrastructure for the annotation of Web-based resources, with a particular emphasis on the collaborative use of annotations. The format of the annotations is RDF. XML or HTML documents can be annotated using XPointer [114] for the location of parts of documents. CREAM is very similar to Annotea, while claiming to be more generic with respect to the supported types of documents that can be annotated. Putting aside the user-friendly interfaces for concept and resource selection, early annotation tools provided very limited process support. Amaya [82] or OntoMat [11] provide basic features such as drag-and-drop lists of ontology concepts which can be used to annotate text markups. Besides manual annotation approaches there are automated methods which are mainly applied to bootstrap the creation of annotations for large document collections. These methods are sometimes based on the availability of small user-generated training sets which are used to derive semantic annotations from resources. Popular approaches include the Armadillo system which can be used to annotate large Web document collections. It makes use of adaptive information extraction techniques by generalizing annotated examples provided by the user. A similar approach is followed by SemTag [23] which is known to have produced the largest annotated document collection up to date. Another popular approach is provided by the MnM application [110] (cf. Figure 17). MnM is based on manual annotations of a training corpus and uses wrapper induction and NLP techniques to generate rules that can be used to extract information from a text-based document collection. Another system which supports automated ontology-based annotation is Magpie [60].

Some annotation methods for documents are being integrated in the authoring process of the documents that are being generated. This includes early approaches such as the Amaya browser which implements the Annotea framework and allows is users to annotate Web pages during their creation and viewing. A newer approach for integrated Web page annotation is WEESA which includes RDF-based annotations in the Web-page creation process [86]. Another approach to annotate Web pages is implemented by SMORE (cf. Figure 18) which was developed in the Mindswap project at the University of Maryland [109]. It allows to annotate Web pages via

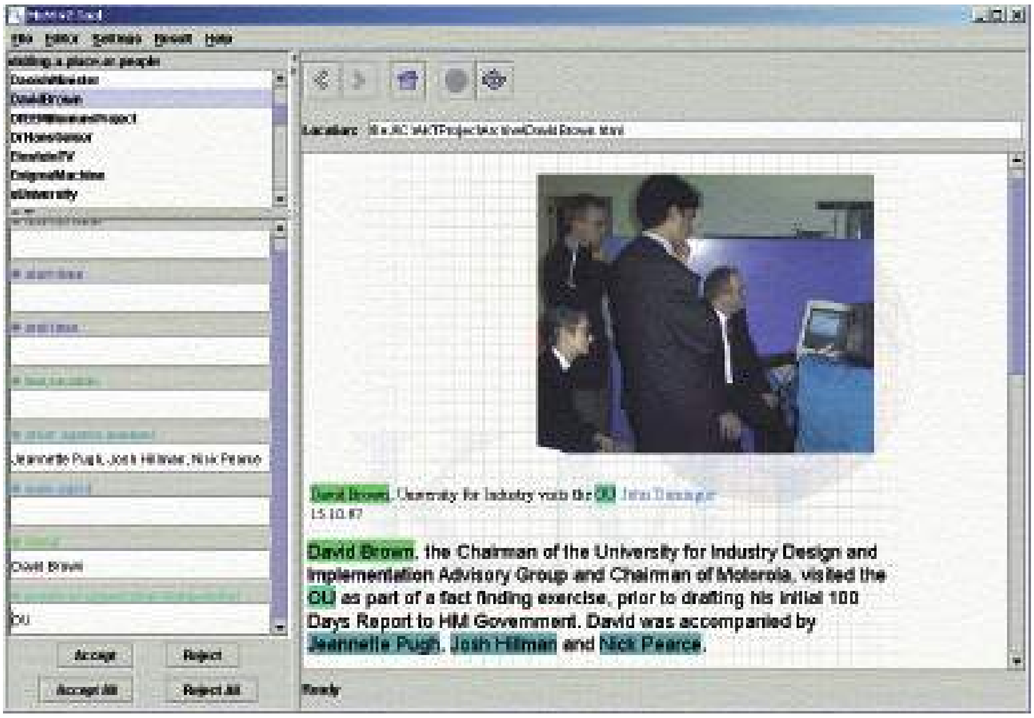


Figure 17: MnM: Results of the Automatic Extraction

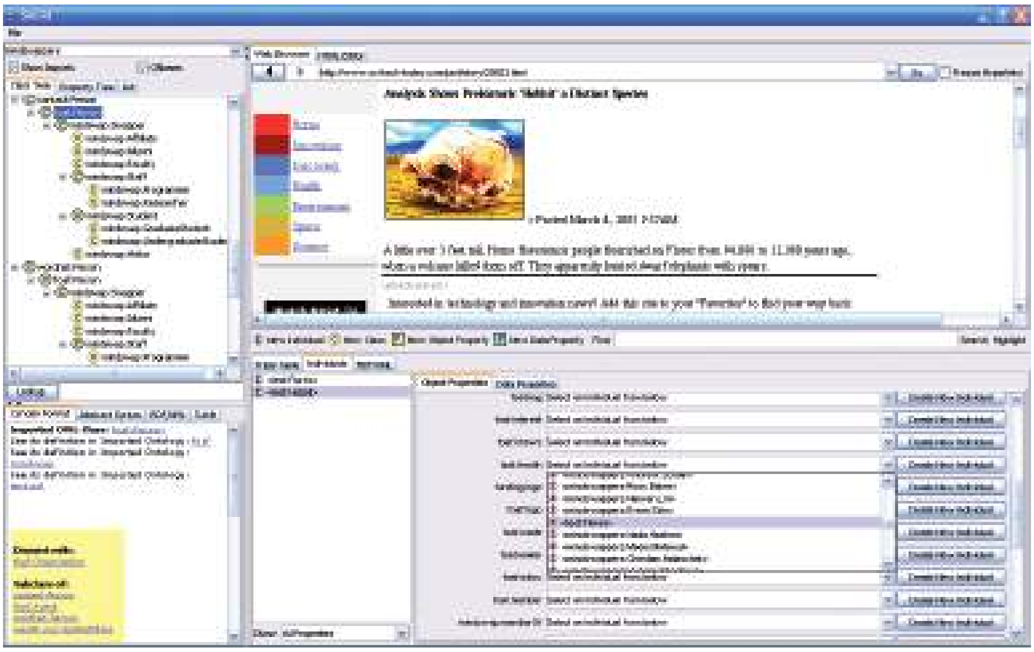


Figure 18: Creation and Markup of Web pages in SMORE

drag-and-drop and to create new HTML pages in the included HTML editor. Concepts, properties and relations can be assigned to resources via drag and drop.

Further examples include AktiveMedia which enables the creation and annotation of cross-media documents, the commercial tool OntoOffice [77] which can be used to annotate Microsoft office documents or the newer MyOntology Word plugin [71]. Semantic annotation is also offered in Semantic Wikis which typically allow the annotation and typing of Wiki pages using ontologies. A considerable amount of Semantic Wiki systems exist, the most popular being Semantic Wiki which is based on the popular MediaWiki software [93].

Further Semantic Wiki systems are listed online.[6]

Besides these approaches which provide formal annotations, there are at first tools which allow tagging of different types of textual content on the Web or tools which allow to attach informal notes to Web sites. WebNotes (cf. Figure 19) [118], Zotero [126], or MyStickies [73] allow to add notes in a post-it-like style to a

Figure 19: Annotation of Web Pages with Web Notes

Web page. Annozilla allows to add annotations to Web pages based on the Annotea model (cf. Figure 20) [6], a potential Google application which allows blog-based annotation of Web pages,[7] or Wikalong which enables collaborative annotation of Web pages [119]. It in principle aims at the creation of one Wiki page for each Web page on the Web (cf. Figure 21).

[6] http://www.semwiki.org/
[7] http://www.flickr.com/photos/bragadocchio/585792851/

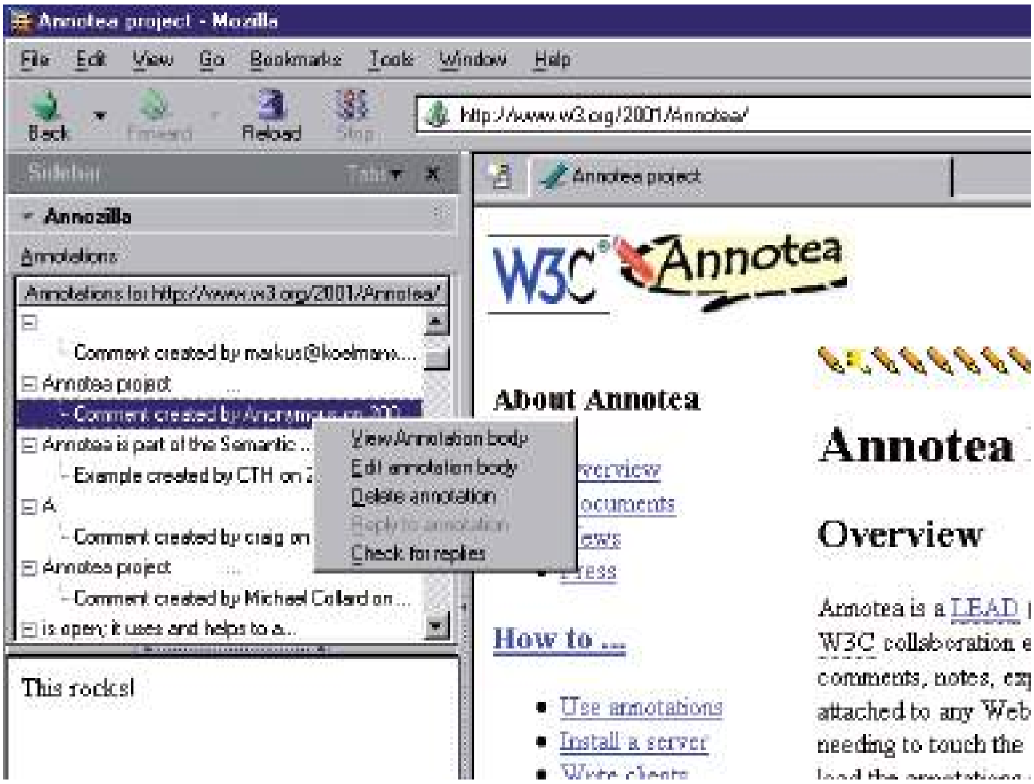


Figure 20: Annotating Web Pages with Annozilla

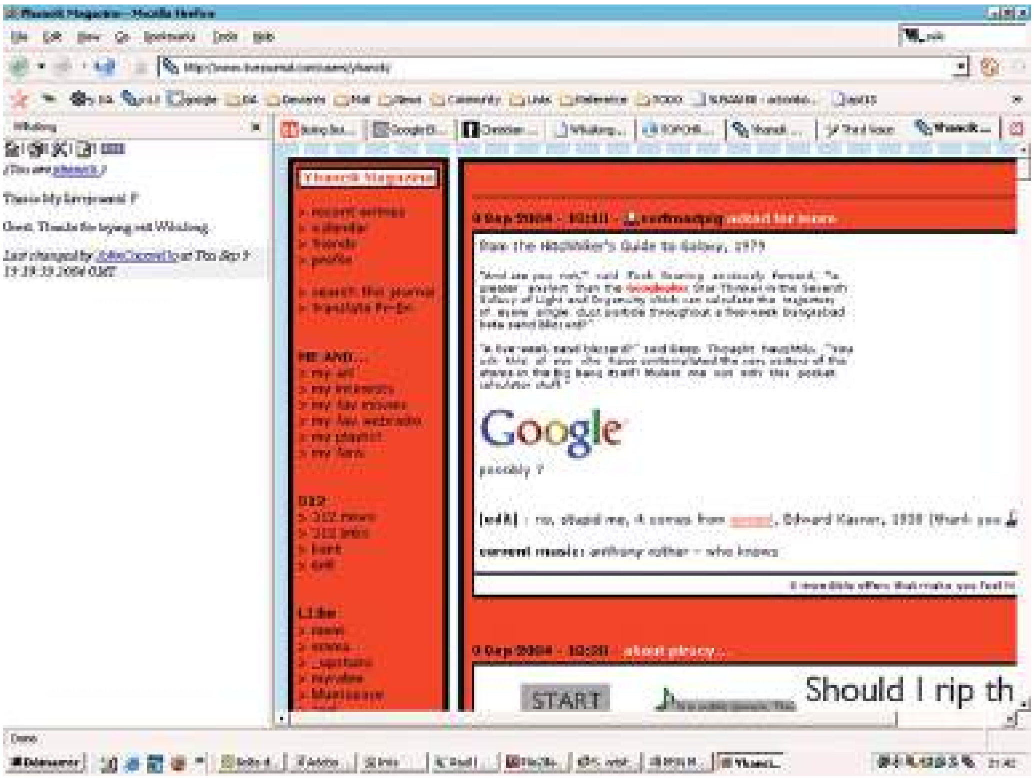


Figure 21: Wiki-based Annotation of Web Pages

2.2.2 Web Service Annotation

Web Service annotation was first approached in the area of Semantic Web services (SWS) [65]. The aim of SWS is the automation of specific Web service usage tasks such as discovery, selection, composition by using semantic descriptions to make them machine processable.

The most popular SWS approaches are WSMO [87], OWL-S [62], and SWSF [10] each of which puts a semantic layer on top of existing (technical) Web service descriptions. WSMO for instance allows the "annotation" of functionality and behavior of Web services. Further it allows the attchement of so-called non-functional properties whose intention is to describe non-functional aspects of Web services such as the creator, rights, or quality of service.

Approaches which directly integrate semantic annotations into the WSDL files are WSDL-S [4] and its successor SAWSDL.[8] Both provide means to point to semantic descriptions outside of the WSDL or to conversion services to obtain a semantic description of various aspects described in it.

More end-user based annotations are provided on platforms such as StrikeIron [102], the ProgrammableWeb, or the Seekda portals [92].

StrikeIron hosts a Web service marketplace which foremost groups services in pre-defined categories (cf. Figure 22) StrikeIron provides natural language descriptions of services in general, the features of the services,

Figure 22: StrikeIron Web Service Market Place

benefits or its use, and technical or pricing information on how to consume the service. Links to related services are given. There is no option for end users to provide annotations on StrikeIron.

Another portal which provides access to descriptions of Web services is ProgrammableWeb (PW). This site hosts a directory of different APIs and their descriptions which includes amongst others SOAP and REST-based Web services. PW offers detailed descriptions of APIs which are categorized and can be tagged using natural language. Besides attribute-based annotations on technical, licensing, and securtiy related issues, which partly use controlled vocabularies, PW allows end-user based comments. End user comments can be in form of natural language descriptions or can be ratings on a five-point scale (cf. Figure 23). Another kind of popularity mechanism is supported by proving links to Web pages or portals which use the respective services. The most sophisticated Web service descriptions are provided on a portal offered by Seekda (cf. Figure 24). Most notably the portal provides an attribute-based description of Web services as can be seen in Figure 24. Besides that, the

[8] http://www.w3.org/TR/sawsdl/

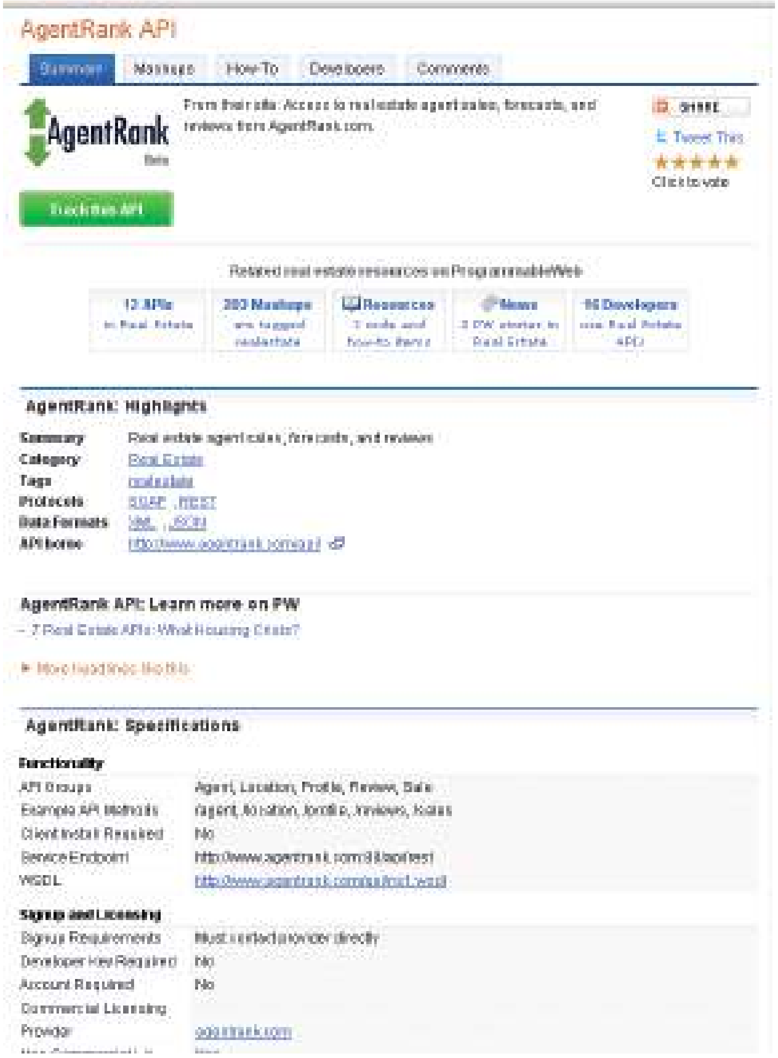


Figure 23: Web Service Description on ProgrammableWeb

current Seekda platform provides means to add Wiki-based comments, ratings, or free tags to Web services as such it realizes a collaborative approach to Web service annotation.

2.2.3 Multimedia Annotation

As multimedia content is a very broad concept and its types range from still images to complex 3D animations, different requirements for multimedia standards and description schemes are given which is why different standards emerged for different types, for different domains and and also for different usage scenarios as we summarized in [15]. Furthermore the shift from professionally to user generated content demands for different description facilities as the needs in professional environments are very different to the needs of home users in tagging scenarios on, for instance, social media sharing sites such as Flickr.

Multimedia annotation can be performed on different levels, i.e., on the metadata level (e.g., administrative or technical descriptions such as title, identifier, or format), the content level (e.g., depicted persons, locations, events), and the multimedia level (e.g., low-level descriptors such as color histograms). In the following we mainly present approaches in which multimedia resources are annotated on the content level with human involvement.

This section is organized as follows: At first we briefly introduce a set of relevant multimedia metadata standards which cover descriptions of multimedia resources on every level mentioned above and which can be used to describe media types in the realm of the PGP use case, that is, image, video, and 3D content. Second we introduce different multimedia annotation approaches which we classify according to the characteristics of

Figure 24: Web Service Description on the Seekda Marketplace

annotation representation models identified earlier in this section.

Metadata and Ontologies for Multimedia There are a multitude of widely deployed multimedia metadata formats which are used to describe pictures taken with digital cameras (cf. EXIF[9] or DIG35[10]) or images created and edited in image creation tools (cf. XMP [3] or NISO Z39.87[11]).

The most widely known and deployed standard for the description of audio visual material is MPEG-7 [63, 74, 75]. It standardizes multimedia description schemes to represent the specific media features, respectively the low level features such as visual features (e.g., texture, camera motion) or audio features (e.g., melody), and the semantic meaning of content (e.g., depicted agents or objects).

Recent research initiatives in the multimedia domain try to overcome commonly known drawbacks of existing multimedia metadata standards for the description of the semantics of multimedia content such as the ambiguities of the descriptions and the lack of interoperability among them. Different ontologies were created based on existing standards. This includes ontologies which are either based on existing standards (e.g., EXIF ontologies[12] or DIG35 ontology[13]) or which were more or less created from scratch (e.g., PhotoRDF [115] or Mindswap Digital Media [45, 44]).

The last years of multimedia semantics research brought up a considerable amount of approaches that deal with the augmentation or reformulation of MPEG-7 to increase interoperability of descriptions and to resolve ambiguities among them which are introduced by the use of natural language in content descriptions and naming for attributes and elements to describe resources (cf. [11, 51, 104, 31, 105, 7]).

Ontologies for 3D content include the “AIMSHAPE ontology for virtual humans” (VHO)[14] [42] which defines concepts, relations and axioms for the semantic descriptions of virtual humans, which are, as defined in [32], graphical representations of human beings that can be used in virtual environments. Tools which are based on this ontology are available online.[15] The SALERO Virtual Character Ontology (SVCO) extends this

[9] http://www.exif.org/
[10] http://www.i3a.org/resources/dig35/
[11] http://www.niso.org/standards/resources/Z39-87-2006.pdf
[12] http://www.kanzaki.com/ns/exif, http://nwalsh.com/java/jpegrdf
[13] http://multimedialab.elis.ugent.be/users/gmartens/Ontologies/DIG35
[14] http://www.aimatshape.net/resources/aas-ontologies/virtualhumansontology.owl/view
[15] http://www.aimatshape.net

ontology to annotate and describe virtual characters [16].

Lightweight approaches to describe visual media include syndication formats which can be used to describe video feeds such as MediaRSS [68] or TXFeed [108] and models to embed semantic descriptions of visual media into (X)HTML pages such as the hMedia microformat [67], Ramm.x [47], or the RICO ("Reusable Intelligent Content Objects") model [14].

More lightweight approaches to media annotation make use of tagging and folksonomies. Free tagging of multimedia is supported on many different Web sites that host multimedia content, such as Flickr [27], YouTube [125], or Slideshare [97]. A more structured tagging approach, which uses MPEG-7's basetypes to type tags, is described in [55].

A more detailed introduction to multimedia metadata standards, description schemes, and vocabularies is given in [15, 48].

Annotation Approaches for Multimedia Tagging-based approaches use single terms to annotate resources. Tags are mostly uncontrolled. In some cases, terms from a controlled vocabulary are used (cf. [49]). Prominent sites on which uncontrolled and free tagging of multimedia resources is possible, includes Flickr or Youtube. Figure 25 depicts an image available on Flickr. Flickr allows to freely tag images and additionally allows to annotate image regions using free text descriptions. Tagging on Flickr is performed in single and public mode. Another example is provided in Figure 9.

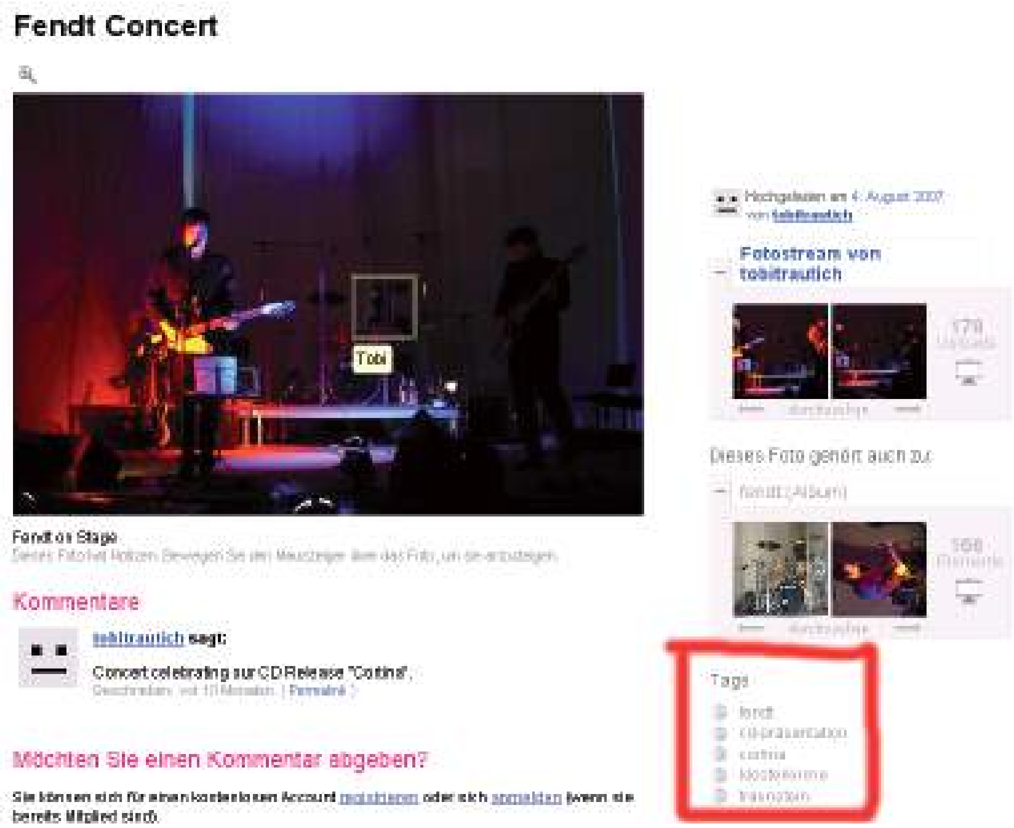


Figure 25: Tagging and Annotation of Images on Flickr

The work of Aurnhammer and others builds upon tags for Flickr images and uses visual features to disambiguate tags (cf. Figure 26) [9]. Sigurbjornsson and others exploit collective intelligence in order to provide tag recommendations for newly tagged images [95]. Rattenbury and others apply different techniques to extract event and location semantics from tags provided on Flickr [84]. Again others derive richer structure from tags which in turn can be exploited in retrieval (cf. [89]).

Tagging-based annotation in a collaborative form is applied in game-based approaches which mostly were inspired by Louis van Ahn's games with a purpose [113]. In games such as the ESP Game (cf. Figure 27) or ListenGame players are in the former case asked to classify images and in the latter to classify audio.

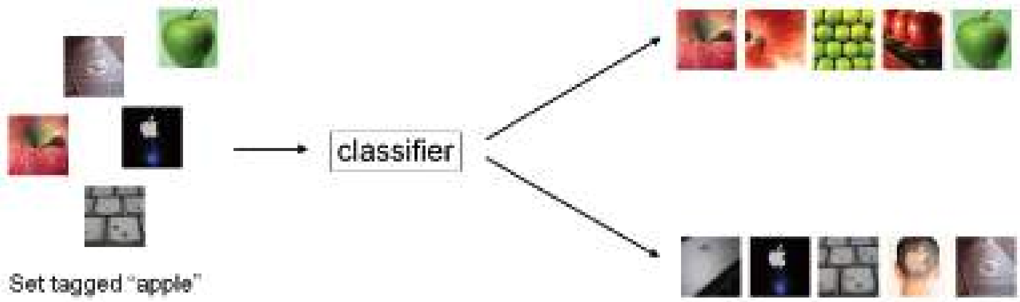


Figure 26: Tag-Disambiguation using Visual Analysis [9]

Figure 27: The ESPGame (http://www.gwap.com)

While the ESP Game allows the description of images using free tags, ListenGame uses a controlled-vocabulary-based approach [107]. A collaborative approach to tag streamed video is applied in Yahoo's video tag game (cf. Figure 28).[16]

The ANVIL video analysis and annotation tool supports an attribute-based annotation model and free-text comments (cf. top-right window in Figure 29). The values for the attributes can be freely defined by the user. Further the system supports a time-based annotation of specific properties of videos (cf. bottom of Figure 29).

Annotea is a generic annotation model which in its earliest version supported the annotation of various media using free text. Annotea has recently been conceptually extended to support collaborative tag- or ontology-based annotation of videos. This new version, called VAnnotea allows the annotation of spatio-temporal regions in videos using structured annotations. Structured annotations might include a number of attributes including hyperlinks, files, free text or controlled vocabularies as explained in [91]. As described in [90], it further can be used to describe media resources using lightweight ontologies such as the WordNet ontology and with simple formal statements (cf. Figure 30).

The Annotea model has been further extended to capture associations between resources and their parts. This extension, called Co-Annotea, is described in [52]. How the Co-Annotea system can be used to compare videos and images is depicted in Figure 31.

Hollink and others apply annotation templates which allow users to describe images using attribute-value pairs [50]. Their approach allows to use concepts from ontologies as values of the attributes. The annotation model thus can be considered as a mixture of attribute and the relation annotation model (cf. Figure 32)

[16] http://sandbox.yahoo.com/VideoTagGame/

Figure 28: Tagging Stream-based Media with the VideoTagGame

Figure 29: The ANVIL Video Annotation Tool (http://www.anvil-software.de)

Photocopain exploits context information to semantically annotate images using lightweight ontologies. It makes use of camera metadata, GPS data, automated low-level image analysis, calendar data, and finally Flickr tags are used to train the system [106]. In order to let user acknowledge or change the automatically generated annotations, Photocopain makes use of the Aktive Media Annotation tool to present the annotations to the user as shown in Figure 33.

More formal and structured annotations are used in tools such as the K-Space Annotation Tool (KAT) [53].

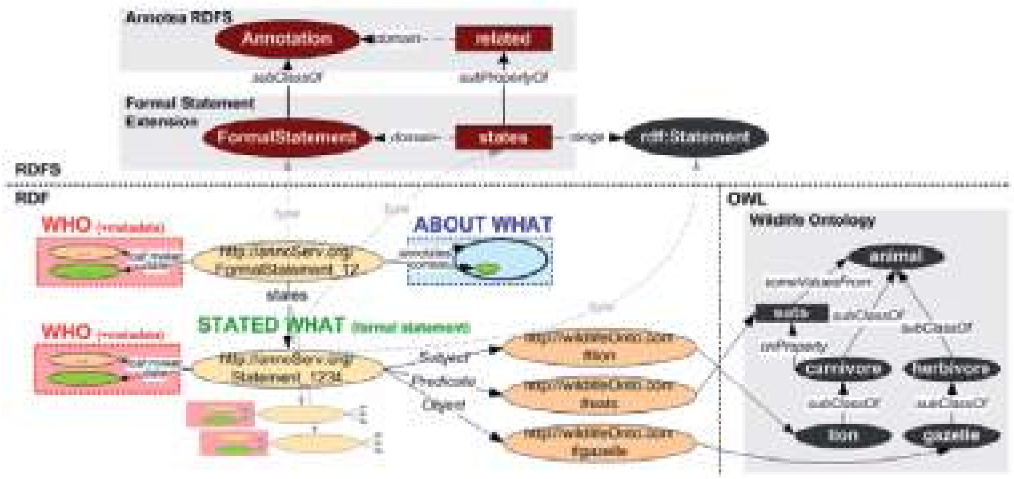


Figure 30: The VAnnotea Annotation Model [90]

Figure 31: Using CoAnnotea to Compare Regions of Medical Images [52]

The KAT tool supports semi-automatic low-level feature extraction and semantic annotation of images using the COMM ontology [7]. As such it allows to annotate resources on the multimedia but also on the content level (cf. Figure 34).

Furthermore, more complex approaches use ontologies to describe the structure of brain images [66] or to describe biological images [17]. Vembu and others use an MPEG-7 based ontology to describe sports events in videos [111]. Others use an MPEG-7 based ontology to apply multimedia reasoning to football videos [20].

2.2.4 Summary

The above brief overview of annotation approaches is supposed to give examples for different annotation models and to present ideas on annotation support for the human-driven creation of annotations. It shall demonstrate how different simple and complex annotation models can be supported on the end user side and it can probably support the INSEMTIVES use case partners in expressing their needs towards an annotation model. The introduced approaches range from tagging of resources over describing attributes of resources to expressing

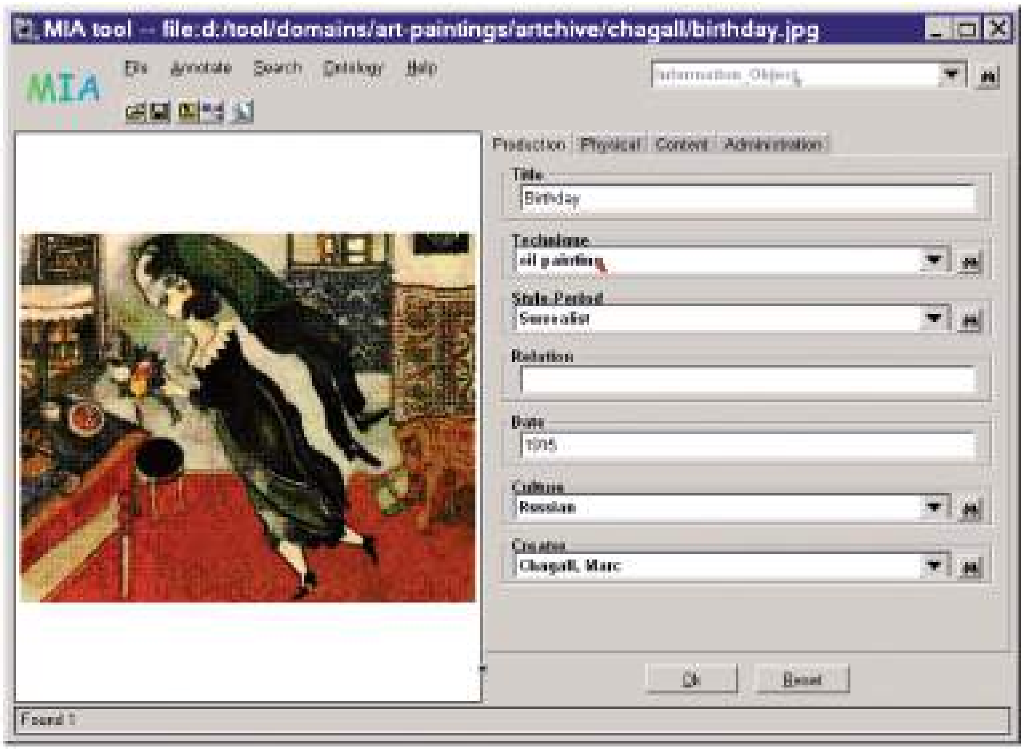


Figure 32: Using Annotation Templates and Ontologies for Annotation [50]

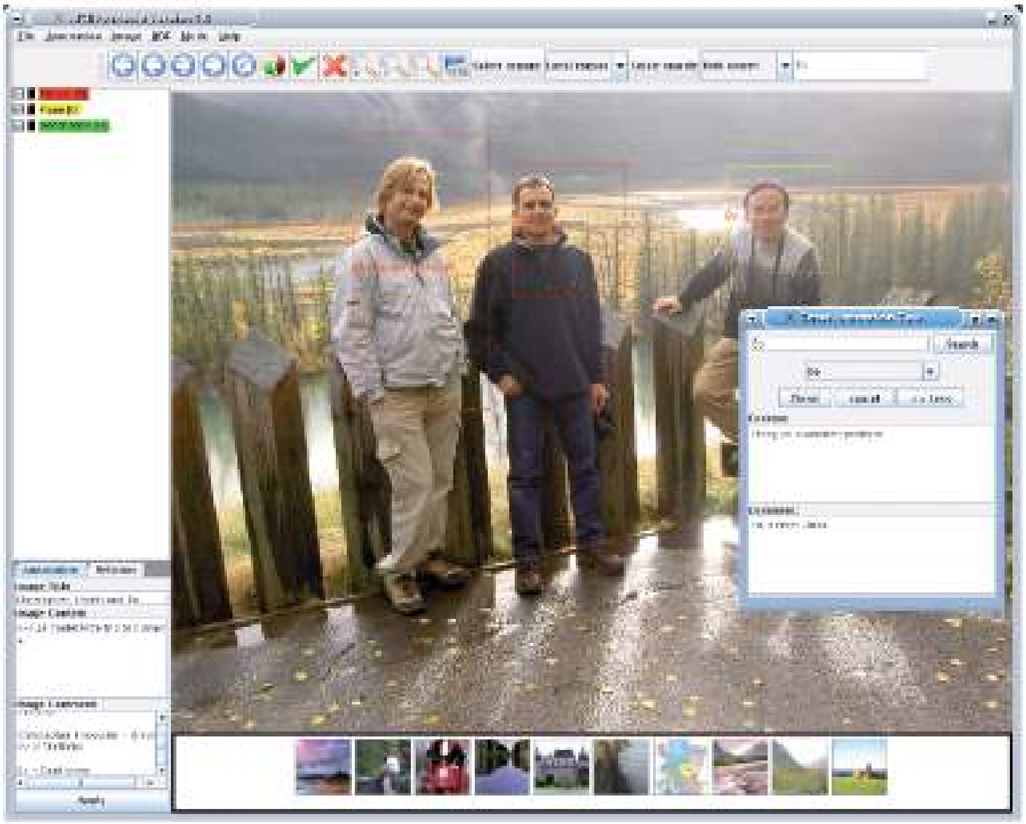


Figure 33: The Aktive Media Annotation Tool as Used by PhotoCopain [106]

more complex statements about resources. Depending on the granularity of the annotations, user interfaces may become arbitrarily complex, however their usefulness is in most of the cases given, as shown above. A

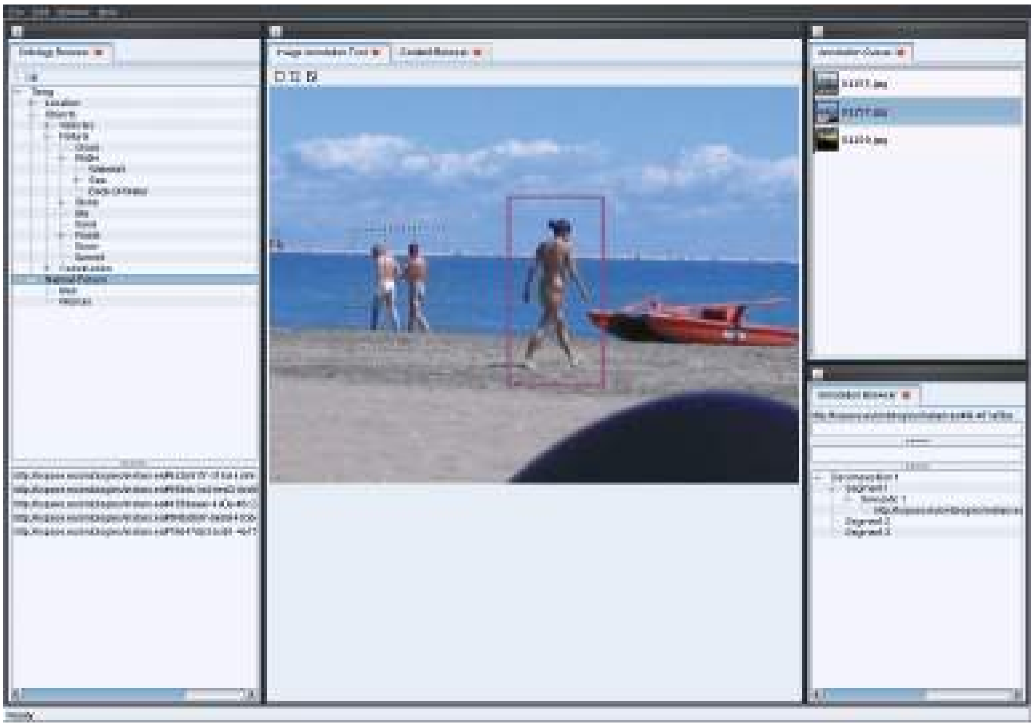


Figure 34: The K-Space Annotation Tool

recommendation about which approach to follow can at this stage not be made as it strongly depends on the specific needs of the partners. In general we however expect to follow the direction of Louis van Ahn's games with a purpose on the interface and incentive level while on the annotation level a model has to be developed which supports the annotation of different media types, structural annotations of parts and whole resources, collaboration, and the storage of provenance information. A model which matches these requirements is, for instance, the VAnnotea model.

3 Requirements for Annotation Representation Models

This section presents the requirements for the annotation model to be implemented in the INSEMTIVES platform which have been elaborated together with the INSEMTIVES use case partners, namely PGP, Seekda, and Telefonica. A summary of the requirements extracted from each use case is provided in Table 3. In the following we provide a summarization of the interviews whose transcripts are available in Appendix , , and

3.1 Requirements from SEEKDA

In this section we report on the requirements for the annotation model which is supposed to support the needs of the SEEKDA use case. These requirements were developed from the results of the interview reported in Appendix C and are structured according to the annotation features presented in Section 2: the required structural complexity, the vocabulary type, and the user collaboration model.

The SEEKDA use case will provide annotation tools for the (end user) annotation of Web services. It demands for a solution to annotate Web services, their properties, and different aspects of stakeholders engaged in the Web services such as their providers (see question 1). The annotation should, but not exclusively, point to aspects already described in WSDL files. As such annotations might point to functional aspects or to descriptive elements (see question 2).

3.1.1 Structural Complexity

The SEEKDA use case requires a mixed complexity in the annotation model. Due to considerable differences in user characteristics, an annotation procedure should support quick and easy annotations but also very detailed annotations (see questions 15 and 16). The annotators are already used to a *simple annotation mechanism* which supports tagging, rating, and commenting of resources (see questions 36 and 37).

The annotators should be able to provide annotations about the whole resources as well as parts of them (see question 3). Furthermore users should be able to annotate specific attributes (see question 4) which are consequently used in search for filtering (see questions 8 and 27). Furthermore, semantic relations between different resources should be expressible (see question 32).

The annotation model should thus be able to support *tags* which are not associated to any particular property of the resource, but also *attributes* to describe specific metadata and semantic *relations* between resources for navigation.

3.1.2 Vocabulary Type

In the SEEKDA use case different types of vocabularies shall be employed, ranging from free-text to ontology-based annotations (see questions 11 and 14). Different user groups might either enter free-text comments or tags, while others provide structured and detailed knowledge about properties of a resource.

The controlled vocabularies are predefined (see question 9) and will solely be provided by engineers inside of SEEKDA (see questions 10 and 20). Current experiments within SEEKDA use an ontology for annotation (see question 9).

SEEKDA expressed interest in resolving of synonymy and polysemy and to expand queries (see questions 25 and 26). This demands at least for the representation of an *authority file*. Furthermore, users might be interested in seeing detailed classification information or clustering of results (see question 33), which will require a *taxonomy*.

3.1.3 Collaboration Type

SEEKDA provides shared resources. Access to them is restricted on the level of annotations which means that every user can view a resource, but he should not be able to see every annotation available for it (question 30). Search for resources is consequently not limited (see question 29).

Resources can be annotated by many users (see questions 17-19). Some properties can however only be described by the owner of a resource (see question 6).

The prospected collaborative annotation model should thus further embed provenance information (who, when; see question 12) as well as versioning information (see question 13).

3.2 Requirements from TID

In this section we report on the requirements for the annotation model to be used in the Telefonica use case. These requirements were developed from the results of the interview reported in Appendix D and are structured according to the annotation features presented in Section 2: the required structural complexity, the vocabulary type, and the user collaboration model.

Telefonica's use case will provide an annotation platform for users of an Intranet. The users will be able to annotate blog entries and forum posts in HTML format but will also be able to annotate news videos (see question 1 and 2). For all type of resources, Telefonica requires to be able to annotate parts of the resource (i.e., sentences, paragraphs for text and segments for videos; see question 3).

3.2.1 Structural Complexity

The Telefonica use case requires a mixed complexity in the annotation model. The users are only interested in an *easy and quick* annotation procedure (see questions 4, 15, and 16) and the annotators are already used to a *simple tagging interface* (see questions 36 and 37).

However, in the use of the annotation for searching and navigating, the users should be able to *filter out* searches by using particular attributes (e.g., the author, a date range, etc.; see question 27). Telefonica also expressed an interest in having an annotation model that supports relational annotations to link resources together, but also parts of resources (see question 3) for navigating between resources (see questions 14, 28, and 32).

The annotation model should thus be able to support *tags* which are not associated to any particular property of the resource, but also *attributes* to describe specific metadata and *relations* between (and within) resources for navigation.

3.2.2 Vocabulary Type

During the interview, Telefonica made clear that the annotation should not be limited to a specific controlled vocabulary and that users will enter free-text annotations (see questions 9, 10, and 11). However, after further discussion and by looking at the intended usage, it appears that the free-text annotation will be combined with a controlled vocabulary of some sort, provided by experts (see questions 7, 9, 20, 31). The controlled vocabulary might take the form of a taxonomy to ease the navigation and search, and the users will be ready to aligne their free-text annotations to the concepts in the controlled vocabulary (see question 33).

A set of expert users, with a good knowledge of the content as well as of the annotation system, will provide an initial controlled vocabulary that can be used to bootstrap the recommendation system. This vocabulary is then extended by the users when providing free-text annotations.

Telefonica has set as a strong requirement that the issues of synonymy and polysemy should be resolved for improving the search accuracy (see questions 25 and 26). This will either require an *authority file* or a *taxonomy* and resources that are annotated with unambiguous metadata. Hence, this use case will require a controlled vocabulary where users will help disambiguating the annotation to map *free-text tags* to the corresponding concept in the vocabulary (see question 33).

3.2.3 Collaboration Type

From the use case description and the interview, it is clear that Telefonica will have many shared resources that every user can access, annotate, and search for (see questions 6, 10, 17, 29, and 35). In the exception of some of the resources where there is a need for access control, all the resources and annotations will be available to everyone.

Moderators will be controlling the user's annotations to make sure they are correct. Telefonica further expressed a strong need for traceability of the annotations for auditing. The collaborative annotation model thus should embed provenance information (when, who; see question 12) as well as versioning information (see question 13).

In addition, the controlled vocabulary required will be built both in a *top-down* fashion, with a small set of experts creating the initial vocabulary but also in a *bottom-up* fashion, with users and moderators extending the existing vocabulary and annotating resources on their own.

3.3 Requirements from PGP

In this section we report on the requirements for the annotation model of the PGP use case. These requirements were developed from the results of the interview reported in Appendix B and are structured according to the annotation features presented in section 2: the required structural complexity, the vocabulary type, and the user collaboration model.

The PGP use case will provide an annotation platform for users of an online virtual world called MyTiny-Planets. An aspect to be highlighted for the PGP use case is that it consists of three sub-use cases, each covering different types of content and target groups. This aspect increases the complexity of the use case and thus puts additional demands on the annotation model. The annotators are supposed to annotate (astrological) images (use case 1), videos (use case 2), and multi-layer Flash content representing virtual landscapes (use case 3) (see question 1). The content formats used are Flash and various image formats (see question 2). PGP especially requires to annotate parts of all content types which means spatial regions in images, spatio-temporal regions in videos, and regions in layers of the Flash content (see question 3).

3.3.1 Structural Complexity

The PGP use case requires a mixed complexity in the annotation model. The user characteristics of possible annotators demand for an *easy to use* annotation procedure which hides most of the complexity behind the annotation interface (see questions 15 and 16). There is no indication that the prospected annotators are used to annotation interfaces.

The annotators should in some cases be able to annotate specific aspects or attributes of the content (see questions 4 and 27) and annotations shall be attached to both resources and parts of it (see question 3).

PGP furthermore expressed an interest in a model that supports relationships to be able to identify semantically related resources (questions 28 and 32)

The annotation model should thus be able to support *tags* which are associated to the *whole resource*, *but also parts with it*. Further it should be possible to describe specific *attributes* of resources which refers to the *attribute*-model. Finally, *relationships* between resources shall be described (cf. the *relation*-model).

3.3.2 Vocabulary Type

As expressed in the interview, the annotations in the PGP use case should be solely based on *controlled vocabularies* (questions 9,10) which can only be defined by administrators of the platform (question 20).

PGP considers offering thematical browsing based on taxonomies (see question 33) or query expansion strategies (see question 25) which requires a *taxonomy* with which resources are annotated with.

The third part of the use cases requires annotation based on more structured knowledge (see question 14) as narrative structures shall be represented (see question 7).

Hence this use case will require a *controlled vocabulary* which might be organized in a *taxonomy* or richer knowledge structures.

3.3.3 Collaboration Type

PGP will have different types of resources that any user of the PGP platform can annotate (see questions 6, 17, and 19). The created annotations will always be available to everyone and access will not be restricted (see questions 29 and 30). Annotators and searchers might be different depending on the sub-use case (see question 35).

There is no need to keep track of a version history for the annotations, but the username and a timestamp have to be stored to allow their traceability (see questions 12 and 13).

The PGP use case thus demands for a *collaborative annotation* model in which the vocabulary used will be provided by PGP.

3.4 Generic Requirements for Annotation Representation Models

This section presents the minimal set of requirements for the INSEMTIVES platform based on the study of the state-of-the-art as well as the specific requirements for each use case partner. The specific requirements that are

	Structural Complexity					Vocabulary Type						Access	
	Tag	**Attributes**	**Relations**	**Ontologies**	**Granularity**	**Free-Text**	**AF**[a]	**Taxonomy**	**Provider**	**Provenance**	**Versioning**	**Read**	**Write**
TID	√	√	√		parts	√[b]		√	experts+users	√	√	controlled	controlled
SEEKDA	√	√	√		whole	√[c]		√	experts	√	√	controlled	controlled
PGP	√	√	√	√	parts			√	experts	√		everyone	everyone
Generic	√	√	√		whole			√	experts	√		everyone	everyone

Table 3: Summary of the Requirements for each Use Case

[a] Authority File
[b] in the Telefonica use case, the free-text annotations will be mapped to a controlled taxonomy.
[c] for SEEKDA, the free-text annotations are independent from the controlled taxonomy.

not covered in the platform with the generic requirements should be added as use case specific extensions of the platform.

3.4.1 Structural Complexity

The platform should support a mix of structural annotation complexity. The resources can be annotated with *tags* which are not linked to an explicit property of the resource but the model should also provide ways to annotate specific attributes of a resource (e.g., author, contributors, etc.) as well as define relations between resources.

The generic annotation model should allow the annotation of whole resources. Annotations within a resource are use case specific as they will depend on the type of media (text, video, image) to annotate.

3.4.2 Vocabulary Type

The generic model should support controlled vocabularies structured in *taxonomies*.

The Telefonica and SEEKDA use cases also require free-text annotation. The way they are used and linked to the controlled taxonomy is however use case specific and should be left to the specific model.

3.4.3 Collaboration Type

The platform should support a collective annotation model which supports the annotation of resources with a controlled vocabulary provided by experts.

The Telefonica use case is the only one which requires a special case of controlled vocabulary that can be dynamically extended by users.

Telefonica and SEEKDA both require some sort of access control on who can annotate and view annotations but this is very specific to each use case and should thus be left outside of the generic model.

3.4.4 Provenance and Versioning

The concept of incentives for semantic annotation will require to keep track of who provides which annotation. The generic model should thus integrate provenance information (i.e., who and when) for every annotation produced on the plattform.

The versioning of annotation is not required by all use case partners and can thus be left to the specific partner requirements.

4 Conclusions and Outlook

This deliverable investigated existing models for representing annotations. It analyzed their different characteristics, forms, and functions. It furthermore looked into media-type specific annotation models which are relevant for the use cases of the INSEMTIVES project. Based on this analysis, a classification scheme for annotation models was developed which distinguishes between structural complexity of annotations, the type of the used vocabularies, and the collaboration type supported. Furthermore a comprehensive overview of existing annotation approaches is given which provides examples of annotation models and their different characteristics.

The deliverable further reports on structured interviews with the use case partners in the project from which requirements for annotation models have been extracted. The requirements which were gathered based on a questionnaire available in the deliverable, are summarized: This includes requirements for annotation models, but also a set of further functional and non-functional requirement about, for instance, incentive mechanisms, storage, or scalability and performance of an annotation platform supporting the thought model. The requirements will provide the foundation for further work in this work package. Our investigations revealed, that the prospected annotation model to be developed has to support a wide range of requirements summarizing many characteristics of existing approaches. These requirements are consolidated in the deliverable. In the course of our work we further abstract from these requirements in order to be able to address needs of a wider spectrum of applications. Despite having many commonalities, the use cases are different enough to ensure a wide applicability of the prospected annotation model.

Based on the results presented in this deliverable, the next deliverable (D2.1.2), will specify models for representing single- and multi-user based annotations to be supported in INSEMTIVES.

References

[1] Sesame. http://openrdf.org. (last accessed on 01.06.2009).

[2] Karl Aberer, Philippe C. Mauroux, Aris M. Ouksel, Tiziana Catarci, Mohand S. Hacid, Arantza Illarramendi, Vipul Kashyap, Massimo Mecella, Eduardo Mena, Erich J. Neuhold, and Et. Emergent semantics principles and issues. In *Database Systems for Advances Applications (DASFAA 2004), Proceedings*, pages 25–38. Springer, March 2004.

[3] Adobe. Xmp specification – part 1, data and serialization models. http://partners.adobe.com/public/developer/en/xmp/sdk/XMPspecification.pdf, January 2008.

[4] Rama Akkiraju, Joel Farrell, John Miller, Meenakshi Nagarajan, Marc-Thomas Schmidt, Amit Sheth, and Kunal Verma. Web service semantics: Wsdl-s. W3C Member Submission, 07. November 2005, 2005.

[5] M Ames and M Naaman. Why we tag: motivations for annotation in mobile and online media. In *CHI 2007 - Tags, Tagging and Notetaking*, pages 971–980, San Jose, 2007.

[6] Annozilla. http://annozilla.mozdev.org/. (last accessed on 01.06.2009).

[7] Richard Arndt, Raphal Troncy, Steffen Staab, Lynda Hardman, and Miroslav Vacura. Comm: Designing a well-founded multimedia ontology for the web. In *The Semantic Web – Proceedings of ESWC 2007*, pages 30–43. Springer, 2007.

[8] Soren Auer, Sebestian Dietzold, , Thomas Riechert, and Thomas Riechert. Ontowiki - a tool for social, semantic collaboration. In *The Semantic Web - ISWC 2006, 5th International Semantic Web Conference, ISWC 2006*, pages 736–749. Springer, 2006.

[9] Melanie Aurnhammer, Peter Hanappe, and Luc Steels. Integrating collaborative tagging and emergent semantics for image retrieval. In *Proceedings WWW2006, Collaborative Web Tagging Workshop*, May 2006.

[10] Steve Battle, Abraham Bernstein, Harold Boley, Benjamin Grosof, Michael Gruninger, Richard Hull, Michael Kifer, David Martin, Sheila McIlraith, Deborah McGuinness, Jianwen Su, and Said Tabet. Semantic web services framework. W3C Member Submission 9 September 2005, 2005.

[11] Stephan Bloehdorn, Kosmas Petridis, Carsten Saathoff, Nikos Simou, Vassilis Tzouvaras, Yannis Avrithis, Siegfried Handschuh, Yiannis Kompatsiaris, Steffen Staab, and Michael Strintzis. Semantic annotation of images and videos for multimedia analysis. In *Proceedings of Semantic Annotation Workshop, co-located with EWSC 2005*, Heraklion,Greece, May 2005.

[12] Simone Braun, Andreas Schmidt, Andreas Walter, Gabor Nagypal, and Valentin Zacharias. Ontology maturing: a collaborative web 2.0 approach to ontology engineering. In *Proceedings of the Workshop on Social and Collaborative Construction of Structured Knowledge (CKC 2007) at the 16th International World Wide Web Conference (WWW2007)*, 2007.

[13] Michael Buckland. Vocabulary as a central concept in information science. In *Digital Libraries: Interdisciplinary Concepts, Challenges, and Opportunities. Proceedings of the Third International Conference on Conceptions of Library and Information Science*, Dubrovnik, Croatia, May 1999.

[14] Tobias Bürger. Towards increased reuse: Exploiting content related and social features of multimedia content on the semantic web. In *Proceedings of the Workshop on Interacting with Multimedia Content on the Web (IMC-SSW), co-located with SAMT 2008, Dez 3-5, Koblenz, Germany, 2008*, 2008.

[15] Tobias Bürger and Michael Hausenblas. Metadata standards and ontologies for multimedia content. In *Handbook of Metadata, Semantics and Ontologies (to be published)*. World Scientific Publishing Co., 2009.

[16] Tobias Bürger, Philip Hofmair, and Gert Kienast. The salero virtual character ontology. In *Proceedings of the First Workshop on Semantic 3D Media, co-located with SAMT 2008, Dec, 3.-5., 2008 (to appear)*, 2008.

[17] C. Catton, S. Sparks, and D. M. Shotton. The imagestore ontology and the bioimage database: semantic web tools for biological research images. In *EWIMT 2005. The 2nd European Workshop on the Integration of Knowledge, Semantics and Digital Media Technology, 2005.*, 2005.

[18] Erik Christensen, Francisco Curbera, Greg Meredith, and Sanjiva Weerawarana. Web services description language (wsdl) 1.1. W3C Note 15 March 2001, `http://www.w3.org/TR/wsdl`, March 2001.

[19] CiteULike. `http://www.citeulike.org/`. (last accessed on 01.06.2009).

[20] S. Dasiopoulou, V.K. Papastathis, V. Mezaris, I. Kompatsiaris, and M.G.Strintzis. An ontology framework for knowledge-assisted semantic video analysis and annotation. In *Proc. 4th International Workshop on Knowledge Markup and Semantic Annotation (SemAnnot 2004) at the 3rd International Semantic Web Conference (ISWC 2004), November 2004*, 2004.

[21] Delicious. `http://del.icio.us`. (last accessed on 01.06.2009).

[22] Diigo. `http://www.diigo.com/`. (last accessed on 01.06.2009).

[23] S. Dill, N. Eiron, D. Gibson, D. Gruhl, R. Guha, A. Jhingran, T. Kanungo, K.S. McCurley, S. Rajagopalan, A. Tomkins, J.A. Tomlin, and J.Y. Zienberer. A case for automated large scale semantic annotation. *Journal of Web Semantics*, 1(1), 2003.

[24] Ebay. `http://www.ebay.com`. (last accessed on 01.06.2009).

[25] Facebook. `http://www.facebook.com`. (last accessed on 01.06.2009).

[26] Faviki. `http://faviki.com`. (last accessed on 01.06.2009).

[27] Flickr. `http://www.flickr.org`.

[28] Flickr. Flickr blog. `http://blog.flickr.net/en`, 2009. (last accessed on 01.06.2009).

[29] General Architecture for Text Engineering (GATE). `http://www.gate.ac.uk/`. (last accessed on 01.06.2009).

[30] Freebase. `http://www.freebase.com`. (last accessed on 01.06.2009).

[31] Roberto Garcia and Oscar Celma. Semantic integration and retrieval of multimedia metadata. In *Proceedings of 4rd International Semantic Web Conference. Knowledge Markup and Semantic Annotation Workshop, Galway, Ireland*, 2005.

[32] A. Garcia-Rojas, D. Thalmann, F. Vexo, L. Moccozet, N. Magnenat-Thalmann, M. Spagnuolo, and M. Gutierrez. An Ontology of Virtual Humans: Incorporating Semantics into Human Shapes. In *The 2nd European Workshop on the Integration of Knowledge, Semantics and Digital Media Technology (EWIMT 2005)*, pages 7– 14, 2005.

[33] Lars M. Garshol. Metadata? thesauri? taxonomies? topic maps! making sense of it all. *Journal of Information Science*, 30(4):378–391, 2004.

[34] Rich Gazan. Social annotations in digital library collections. *D-Lib*, Volume 14 Number 11/12, December 2008.

[35] Geonames. `http://www.geonames.org/`. (last accessed on 01.06.2009).

[36] David K. Gifford, Pierre Jouvelot, Mark A. Sheldon, James W. O'toole, and Jr. Semantic file systems. In *In Proceedings of the 13th ACM Symposium on Operating Systems Principles*, pages 16–25, 1991.

[37] F. Giunchiglia and I. Zaihrayeu. *Encyclopedia of Database Systems*, chapter Lightweight Ontologies. Springer, July 2009. to Appear.

[38] Scott Golder and Bernardo A. Huberman. The structure of collaborative tagging systems. *Journal of Information Science*, 32(2):198208, April 2006.

[39] Web Ontology Working Group. W3c recommendation for owl. `http://www.w3.org/TR/2004/REC-owl-features-20040210/`, February 2004.

[40] Thomas Gruber. Ontology of folksonomy: A mash-up of apples and oranges. *International Journal on Semantic Web & Information Systems*, 3(2):1–11, 2007.

[41] Martin Gudgin, Marc Hadley, Noah Mendelsohn, Jean-Jacques Moreau, Henrik Frystyk Nielsen, Anish Karmarkar, and Yves Lafon. Soap version 1.2 part 1: Messaging framework (second edition). W3C Recommendation 27 April 2007, `http://www.w3.org/TR/soap12-part1/`, April 2007.

[42] Mario Gutierrez, Alejandra Garcia Rojas, Daniel Thalmann, Frederic Vexo, Laurent Moccozet, Nadia Magnenat-Thalmann, Michela Mortara, and Michela Spagnuolo. An ontology of virtual humans: Incorporating semantics into human shapes. *The Visual Computer*, 23(3):207–218, 2007.

[43] Hugo Haas and Allen Brown. W3C Working Group Note 11 February 2004. Web services glossary. `http://www.w3.org/TR/ws-gloss/`, February 2004. (last accessed on 01.06.2009).

[44] Christian Halaschek-Wiener, Jennifer Golbeck, Andrew Schain, Michael Grove, Bijan Parsia, and Jim Hendler. Annotation and provenance tracking in semantic web photo libraries. In *Provenance and Annotation of Data*. Springer, 2006.

[45] Christian Halaschek-Wiener, Andrew Schain, Jennifer Golbeck, Michael Grove, Bijan Parsia, and Jim Hendler. A flexible approach for managing digital images on the semantic web. In *Proceedings of the 5th International Workshop on Knowledge Markup and Semantic Annotation (SemAnnot 2005)*, 2005.

[46] S. Handschuh and S. Staab. Authoring and annotation of web pages in cream. In *Proceedings of the 11th International World Wide Web Conference (WWW 2002)*, 2002.

[47] M. Hausenblas, W. Bailer, T. Bürger, and R. Troncy. Ramm.x: Deploying multimedia metadata on the semantic web. In *Proceedings of SAMT 2007, Dec 4-7, Genova, Italy, 2007*, 2007.

[48] Michael Hausenblas, Susanne Boll, Tobias Brger, Oscar Celma, Christian Halaschek-Wiener, Erik Mannens, and Raphael Troncy. Multimedia vocabularies on the semantic web. W3c incubator group report 24 july 2007, W3C, July 2007.

[49] Sarah Hayman and Nick Lothian. Taxonomy directed folksonomies. In *Proceeings of the World Library and Information Congress: 73rd IFLA General Conference and Council*, 2007.

[50] Laura Hollink, Guus Schreiber, Jan Wielemaker, and Bob Wielinga. Semantic annotation of image collections. In R. Dieng S. Handschuh, M. Koivunen and S. Staab, editors, *Knowledge Capture 2003 – Proceedings Knowledge Markup and Semantic Annotation Workshop*, 2003.

[51] Jane Hunter. Adding multimedia to the semantic web - building an mpeg-7 ontology. In *Proceedings of the 1st International Semantic Web Working Symposium (SWWS 2001), Stanford, USA, 30 July - 1 August 2001*, 2001.

[52] Jane Hunter and Ronald Schroeter. Co-annotea: A system for tagging relationships between multiple mixed-media objects. *Multimedia, IEEE*, 15(3):42–53, 2008.

[53] K-Space Annotation Tool. `http://www.uni-koblenz.de/FB4/Institutes/IFI/AGStaab/Research/kat`. (last accessed on 01.06.2009).

[54] J. Kahan, M.-J. Koivunen, E. PrudHommeaux, and R. Swick. Annotea: an open rdf infrastructure for shared web annotations. In *Proceedings of the 10th International World Wide Web Conference (WWW 2001)*, 2001.

[55] Ralf Klamma, Marc Spaniol, and Dominik Renzel. Community-Aware Semantic Multimedia Tagging - From Folksonomies to Commsonomies. In K. Tochtermann, H. Maurer, F. Kappe, and A. Scharl, editors, *Proceedings of I-Media'07, International Conference on New Media Technology and Semantic Systems*, J.UCS (Journal of Universal Computer Science) Proceedings, pages 163–171. Springer-Verlag, 2007.

[56] Josef Kolbitsch. WordFlickr: a solution to the vocabulary problem in social tagging systems. In *Proceedings of I-MEDIA*, 2007.

[57] Last.fm. http://www.last.fm. (last accessed on 01.06.2009).

[58] Fred Leise, Karl Fast, and Mike Steckel. What is a controlled vocabulary? http://www.boxesandarrows.com/view/what_is_a_controlled_vocabulary_, 2002.

[59] LiveJournal. http://www.livejournal.com/. (last accessed on 01.06.2009).

[60] Magpie. http://projects.kmi.open.ac.uk/magpie/main.html. (last accessed on 01.06.2009).

[61] Bill Mann. *An Introduction to Rhetorical Structure Theory (RST)*. 1999. Published: web page.

[62] David Martin, Mark Burstein, Jerry Hobbs, Ora Lassila, Drew McDermott, Sheila McIlraith, Srini Narayanan, Massimo Paolucci, Bijan Parsia, Terry Payne Evren Sirin, and Naveen Srinivasan Katia Sycara. Owl-s: Semantic markup for web services. W3C Member Submission, 22. November 2004, 2004.

[63] Jose M. Martinez. Mpeg-7 overview. http://www.chiariglione.org/mpeg/standards/mpeg-7/mpeg-7.htm, October 2004.

[64] Adam Mathes. Folksonomies - cooperative classification and communication through shared metadata. Technical report, Graduate School of Library and Information Science. University of Illinois Urbana-Champaign, December 2004.

[65] S. MCIllraith, T. C. Son, and H. Zeng. Semantic web services. *IEEE Intelligent Systems. Special Issue on the Semantic Web*, 16(2):46–53, 2001.

[66] A. Mechouche, C. Golbreich, and B. Gibaud. Semantic description of brain mri images. In *Proceedings of the First International Workshop on Semantic Web Annotations for Multimedia (SWAMM)*, 2006.

[67] Media Info Microformat. http://wiki.digitalbazaar.com/en/Media_Info_Microformat. (last accessed on 01.06.2009).

[68] MediaRSS. http://search.yahoo.com/mrss. (last accessed on 01.06.2009).

[69] Peter Mika. Ontologies are us: A unified model of social networks and semantics. In *International Semantic Web Conference*, pages 522–536, 2005.

[70] George A. Miller. WordNet: a lexical database for english. *Communications of the ACM*, 38(11):39–41, 1995.

[71] MyOntology Consortium. Myontology word plugin. http://www.myontology.org/semantics4office/. (last accessed on 01.06.2009).

[72] Myspace. http://www.myspace.com/. (last accessed on 01.06.2009).

[73] MyStickies. http://www.mystickies.com/. (last accessed on 01.06.2009).

[74] Frank Nack and Adam T. Lindsay. Everything you wanted to know about MPEG-7 (Part I). *IEEE Multimedia*, 6(3):65–77, 1999.

[75] Frank Nack and Adam T. Lindsay. Everything you wanted to know about MPEG-7 (Part II). *IEEE Multimedia*, 6(4):64–73, 1999.

[76] Library of Congress. `http://www.loc.gov/index.html`. (last accessed on 01.06.2009).

[77] Ontoprise. `http://www.ontoprise.de`. (last accessed on 01.06.2009).

[78] Ontotext. Kim semantic annotation platform. `http://www.ontotext.com/kim/`. (last accessed on 01.06.2009).

[79] Ontotext. Owlim native semantic database engine. `http:///www.ontotext.com/owlim/`. (last accessed on 01.06.2009).

[80] Ontotext. Semantic annotation (definition). `http://www.ontotext.com/semanticannotation/`. (last accessed on 01.06.2009).

[81] Picasa. `http://picasa.google.com/`. (last accessed on 01.06.2009).

[82] V. Quint and I. Vatton. An introduction to amaya. W3C NOTE 20-February-1997, February 1997.

[83] Manuele Quintarelli, Andrea Resmini, and Luca Rosati. Facetag: Integrating bottom-up and top-down classification in a social tagging system. In *IASummit*, Las Vegas, 2007.

[84] Tye Rattenbury, Nathaniel Good, and Mor Naaman. Towards automatic extraction of event and place semantics from flickr tags. In *SIRIR '07: Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval*, pages 103–110, New York, NY, USA, 2007. ACM Press.

[85] C Reed and G Rowe. Araucaria: Software for puzzles in argument diagramming and XML. Technical report, Department of Applied Computing. University of Dundee Technical Report., 2001.

[86] Gerald Reif, Harald Gall, and Mehdi Jazayeri. Weesa – web engineering for semanitc web applications. In *Proceedings of the 14th International World Wide Web Conference*, pages 722–729, 2005.

[87] D. Roman, U. Keller, H. Lausen, J. de Bruijn, R. Lara, M. Stollberg, A. Polleres, C. Feier, C. Bussler, and D. Fensel. Web service modeling ontology. *Applied Ontology*, 1(1):77–106, 2005.

[88] Francesco Ronzano, Andrea Marchetti, and Maurizio Tesconi. Tagpedia: a semantic reference to describe and search for web resources. In *SWKM 2008: Intl. Workshop on Social Web and Knowledge Management @ WWW*, 2008.

[89] P. Schmitz. Inducing ontology from flickr tags. In *Proc. of the Collaborative Web Tagging Workshop (WWW 06)*, May 2006.

[90] R. Schroeter and J. Hunter. Annotating relationships between multiple mixed-media digital objects by extending annotea. In *Proceedings of the 4th European Semantic Web Conference, ESWC 2007*, 2007.

[91] R. Schroeter, J. Hunter, J. Guerin, I. Khan, and M. Henderson. A synchronous multimedia annotation system for secure collaboratories. In *Proceedings of the 2nd IEEE International Conference on E-Science and Grid Computing (eScience 2006)*, 2006.

[92] Seekda. `http://seekda.com/`. (last accessed on 01.06.2009).

[93] Semantic MediaWiki. `http://semantic-mediawiki.org`. (last accessed on 01.06.2009).

[94] Shoebox. `http://www.kavasoft.com/Shoebox`. (last accessed on 01.06.2009).

[95] Börkur Sigurbjörnsson and Roelof van Zwol. Flickr tag recommendation based on collective knowledge. In *WWW '08: Proceeding of the 17th international conference on World Wide Web*, pages 327–336, New York, NY, USA, 2008. ACM.

[96] Elena Simperl, Christoph Tempich, and Tobias Bürger. Methodologies for the creation of semantic data. In Miguel-Angel Sicilia, editor, *Handbook of Metadata, Semantics and Ontologies*. World Scientific Publishing Co., 2009 (forthcoming).

[97] Slideshare. `http://www.slideshare.net`. (last accessed on 01.06.2009).

[98] Barry Smith and Christopher Welty. Fois introduction: Ontology—towards a new synthesis. In *FOIS '01: Proceedings of the international conference on Formal Ontology in Information Systems*, New York, NY, USA, 2001. ACM Press.

[99] Soboleo. `http://www.soboleo.com/`. (last accessed on 01.06.2009).

[100] S Sood, K Hammond, S Owsley, and L Birnbaum. Tagassist: Automatic tag suggestion for blog posts. *International Conference on Weblogs and Social Media*, 2007.

[101] Louise Spiteri. Structure and form of folksonomy tags: The road to the public library catalogue. *Webology*, 4(2), 2007.

[102] StrikeIron. `http://www.strikeiron.com/`. (last accessed on 01.06.2009).

[103] Rudi Studer, Richard Benjamins, and Dieter Fensel. Knowledge engineering: Principles and methods. *Data & Knowledge Engineering*, 25(1-2):161–198, MAR 1998.

[104] R. Troncy. Integrating structure and semantics into audio-visual documents. In K. Sycara D. Fensel and J. Mylopoulos, editors, *The Semantic Web - Proceedings ISWC'03*, volume 2870 of *Lecture Notes in Computer Science*, 2003.

[105] C. Tsinaraki, P. Polydoros, and S. Christodoulakis. Interoperability support between mpeg-7/21 and owl in ds-mirf. *Transactions on Knowledge and Data Engineering (IEEE-TKDE)*, Special Issue on the Semantic Web Era, 2007.

[106] M. Tuffield, S. Harris, D. Duplaw, A. Chakravarthy, C. Brewster, N. Gibbins, K. O Hara, F. Ciravegna, D. Sleeman, N. Shadbolt, and Y. Wilks. Image annotation with photocopain. In *Proceedings First International Workshop on Semantic Web Annotations for Multimedia (SWAMM 2006)*, 2006.

[107] Douglas Turnbull, Ruoran Liu, Luke Barrington, and Gert Lanckriet. A game-based approach for collecting semantic annotations of music. In *Proceedings of ISMIR 2007*, 2007.

[108] TXFeed. `http://transmission.cc/xmlns`. (last accessed on 01.06.2009).

[109] University of Maryland. Smore. `http://www.mindswap.org/2005/SMORE/`. (last accessed on 01.06.2009).

[110] M. Vargas-Vera, E. Motta, J. Domingue, M. Lanzoni, A. Stutt, and F. Ciravegna. Mnm: Ontology driven semi-automatic and automatic support for semantic markup. In *Proceedings of the 13th International Conference on Knowledge Engineering and Management (EKAW 2002)*, pages 379–391, 2002.

[111] S. Vembu, M. Kiesel, M. Sintek, and S. Baumann. Towards bridging the semantic gap in multimedia annotation and retrieval. In *Proceedings of the 1st International Workshop on Semantic Web Annotations for Multimedia, SWAMM 2006 at the 15th International World Wide Web Conference, WWW 2006*, 2006.

[112] Max Völkel, Markus Krötzsch, Denny Vrandecic, and Heiko Haller. Semantic wikipedia, 2006.

[113] L. von Ahn and L. Dabbish. Labeling images with a computer game. In *ACM CHI*, 2004.

[114] W3C. Xpointer specification. `http://www.w3.org/TR/xptr/`. (last accessed on 01.06.2009).

[115] W3C. Describing and retrieving photos using rdf and http. W3C Note 19 April 2002, http://www.w3.org/, April 2002.

[116] Thomas Vander Wal. Folksonomy: Coinage and definition. http://www.vanderwal.net/folksonomy.html. (last accessed on 01.06.2009).

[117] Amy J. Warner. A taxonomy primer. http://www.ischool.utexas.edu/~i385e/readings/Warner-aTaxonomyPrimer.html. (last accessed on 01.06.2009).

[118] Webnotes. http://www.webnotes.net/. (last accessed on 01.06.2009).

[119] Wikalong. http://www.wikalong.org. (last accessed on 01.06.2009).

[120] Wikipedia. http://wikipedia.org/. (last accessed on 01.06.2009).

[121] Wikipedia. Attribute (computing). http://en.wikipedia.org/wiki/Attribute_(computing). (last accessed on 01.06.2009).

[122] Wikipedia. Spotlight (software). http://en.wikipedia.org/wiki/Spotlight_(software). (last accessed on 01.06.2009).

[123] Wikipedia. Tag (metadata). http://en.wikipedia.org/wiki/Tag_(metadata), 2009. (last accessed on 01.06.2009).

[124] Yahoo. Upcoming.org. http://upcoming.yahoo.com/. (last accessed on 01.06.2009).

[125] YouTube. http://www.youtube.org. (last accessed on 01.06.2009).

[126] Zotero. http://www.zotero.org/. (last accessed on 01.06.2009).

A Questionnaire for Determining Annotation Needs of the Use Case Partners

Motivation

The following set of questions was used in a questionnaire whose intention was to support the requirements gathering in WP2 of the INSEMTIVES project with respect to annotation models. It specifically was supposed to determine prospected needs of the use case partners and the planned scenarios to be implemented in the project. The main intention was to understand the requirements of the use case partners with respect to the types of annotations.

The questionnaire was finally used as a guideline for telephone interviews.

Questionnaire Design

The questions are divided into 5 parts:

- Information about what the partners would like to annotate and what value they expect
- Information about how the annotation should look like
- Questions about who is supposed to provide the annotations and their motivation, i.e., why do (should) people annotate, how they feel, and how they can be motivated
- Questions about the prospected use of the annotations
- Questions about experiences with annotation approaches among the partners

Information about *what* to annotate

These questions are about what the partners want to annotate, on which level of granularity, and which specific properties of the content should be annotated.

1. **What is the content type you need to be annotated ?**
   This could include either multimedia content, text, or Web services.

2. **What are the content formats used?**
   e.g., specific image or multimedia formats, specific structured/unstructured textual content formats, WS description languages used and the actual representation of the descriptions.

3. **Do you need annotations on the entire resource, or on a part of it ?**
   this is about granularity of the annotations. e.g.

   - do you need to annotate a document, a paragraph, a sentence or any region in textual content
   - do you need to annotate an entire image (like Google Image Labeler) or using a shape to define the region of the annotation (what shape makes sense for you: rectangular, polygonal, ellipsoidal, etc.) or - vague region indicators like upper-left; center; right; etc.
   - multimedia (video or games) with temporal or story line dimension: do you need annotations to have a time span characteristic; is it absolute time, or the time to reach from A to B in the story line of a game for example (say someone wants to say that the part of the game where you have to walk from the forest to the house doing nothing is boring).

4. **Do you want to annotate specific properties of the content?**
   This could include the structure of documents, but also more abstract properties such as rights, title, main color, reliability, quality, etc.

5. **How many resources need to be annotated ?**

6. **Should the user be allowed to annotate only his own resources, resources only of a group he belongs to (from a department for example) or can everyone annotate any resource?**

7. **Why is the annotation of specific content important? What advantages do you think to obtain from annotated knowledge/information?**
Interviewee should explain why annotation of specific content is important and how it will improve the current companys state of the art

Information about *how* the annotation should look like

8. **How would you like to use the annotations after they are created?**
e.g., we would like to search (what and how); navigate; visualize/display (what and how); use them as a link to existing structured knowledge.

9. **Do you intend to use predefined annotations provided to the annotator, and what are they?**
Like, being able to annotate an image with respect to two predefined aspects:

    - emotional: "fun, educational, boring, dull, etc."
    - factual: it is picturing "animal, human, nature, abstract thing"

10. **Can the set of available annotation terms be extended dynamically (or automatically), and by whom?**

11. **Can the annotators use completely new annotations that are outside of a predefined controlled vocabulary?**

12. **What is the provenance information you need for an annotation?**

13. **Do you intend to keep track of the version history of the annotations ?**

14. **What complexity do you expect the annotations to be?**

    - phrases (no links to external structured data)
    - link to existing structured knowledge
    - do you need the annotators to choose the type of the relationship between the resource and the external structured data instance/entity, or just expressing that the resource mentions it: e.g.
        - a news article refers Frodo the Lizzard; Frodo is an *entity* of type *fairy-tale hero* in our ontology; we expect the task of the annotator to only link the reference in the text to the actual entity in the knowledge base (simpler scenario)
        - annotating how entities are displayed on an image: Frodo is behind Sheila; Sheila is greener than Frodo, etc. i.e. the references already expose particular relationships, properties or attributes
    - template annotations (like: saying for a *planet* that *planet name* is Earth, and it is
        - *populated* TRUE
        - *hasWater* TRUE
        - *hasPopulation* size X Billion
        - *hasVisibleColour* Blue

Questions about *who* is supposed to provide the annotations and about their *motivation*

15. **What are the characteristics of your annotators?**
e.g., are they experts; their age group, if relevant; familiar with hierarchies/not; etc.

16. **How much time you expect your annotator to spend on a single resource?**

17. **Do you expect a single resource to be annotated by one or many annotators?**

18. **What is the expected amount of annotations per resource per annotator?**

19. **How many annotators you expect per resource?**

20. **If you want to offer pre-defined annotation terms: Who can define the set of available annotation terms?**

21. **Does the company pay the annotators for annotating content?**

22. **Does the company publicly endorse the annotators?**

23. **Does the management force them to do? In other words, what motivation do workers have to annotate content?**

Questions about the prospected *usage* of the annotations

24. **When searching for content, what does the user need to retrieve?**

    - will they have to retrieve only their own documents, documents from other users?
    - do they need to do generic search, about a topic, and find all the relevant documents?

25. **Do you intend to offer query expansion strategies?**
    e.g., when searching for resources annotated with ”big”, should the user be able to find resources annotated with ”large”?

26. **Should a search/annotation interface resolve ambiguities?**
    e.g., in annotation and search, should the user be able to differentiate between the several meanings of ambiguous terms in order to improve the precision of search? For example, when providing the term ”Java”, should the user be able to explicitly tell to the system that the term refers to the concept of the Java programming language?

27. **Should the user be able to explicitly refer to the property of the resources the user query refers to?**
    For example, the user could specify that the query ”John Doe” refers to the property ”author” of document resources and, therefore, avoid finding documents that mention ”John Doe” in their body

28. **When searching for a resource, should the user be able to find resources that have a certain relationship with another resource?**
    For example, does the user need to retrieve documents that are referred to in another document. In the Seekda use case, do the user need to search for services that require a particular service

29. **Should the scope of the search be limited to a specific group of users (e.g., within a particular department) or a single user?**

30. **Are all the resources and annotations always available to everyone? or should there be some access control features?**

31. **Are the users going to navigate the whole set of resources or just do searches?**

32. **Should the user be able to navigate from a given resource to other semantically related resources?**

33. **If they navigate, are they interested in seeing a detailed classification/taxonomy (such as by country, then by region, then by city, etc.) of the resources or a rough clustering (such as the tag cloud example in the report)?**

34. **Are the users interested to see statistics about the resources and their annotation?**

35. **Will the same people that make the annotations, search based on the annotations? Are annotators and end user different?**

Questions about experiences and existing annotation mechanisms in the applications of the partners

36. **Do you have experiences with annotation mechanism? If yes: which?**
37. **Do you already use annotation mechanisms in your application(s)? If yes: could you please explain them and the drawbacks you see in them?**

B Questionnaire and Results for PGP Use Case

Information about *what* to annotate

1. **What is the content type you need to be annotated ?**
   Images, Video, Multi-Layer Flash Content

2. **What are the content formats used?**
   Flash, arbitrary image formats

3. **Do you need annotations on the entire resource, or on a part of it ?**
   The entire resource, but also parts.

4. **Do you want to annotate specific properties of the content?**
   Yes, emotional Context in the third use case.

5. **How many resources need to be annotated ?**
   First use case: x0000 images, second use case: maximum 65 videos (perhaps limited down to 20), third use case: up to 1 million variations of landscapes.

6. **Should the user be allowed to annotate only his own resources, resources only of a group he belongs too (from a department for example) or can everyone annotate any resource?**
   Everybody can annotate any resource.

7. **Why is the annotation of specific content important? What advantages do you think to obtain from annotated knowledge/information?**
   First use case: "They do not want to park their rockets on a big rock"; geological evaluation on data sets; too large to manually evaluate. Second use case: Limit relevant videos in a production context. Third use case: to obtain an "interactive narrative".

Information about *how* the annotation should look like

8. **How would you like to use the annotations after they are created?**
   First use case: Set of criteria that people would like to search by, explore relations of data, get any info on the image; get all images which have interesting information on them Second use case: sample query: "Show me a shot where Bing is happy and has his telescope" Third use case: sample query: "Show me all the landscapes that make people feel happy, miserable, sad, frightened"

9. **Do you intend to use predefined annotations provided to the annotator, and what are they?**
   Yes. But don't know exactly what they are at the moment.

10. **Can the set of available annotation terms be extended dynamically (or automatically), and by whom?**
    Want to limit; but somebody should be able to add terms to the set (admin level)

11. **Can the annotators use completely new annotations that are outside of a predefined controlled vocabulary?**
    No.

12. **What is the provenance information you need for an annotation?**
    We need to store the username and the timestamp.

13. **Do you intend to keep track of the version history of the annotations ?**
    No.

14. **What complexity do you expect the annotations to be?**
    Link to existing structured knowledge: yes; use case will need slightly more structure knowledge.

Questions about *who* is supposed to provide the annotations and about their *motivation*

15. **What are the characteristics of your annotators?**
kids, non-experts.

16. **How much time you expect your annotator to spend on a single resource?**
1. Use case: less than a minute 2. Use case: 5-10 minutes 3. Use case: 2-3 minutes

17. **Do you expect a single resource to be annotated by one or many annotators?**
1. use case: Many 2. use case: Many 3. use case: Possible one or several

18. **What is the expected amount of annotations per resource per annotator?**
No idea.

19. **How many annotators you expect per resource?**
no idea; will vary dramatically between the use cases; on many some; on many few.

20. **If you want to offer pre-defined annotation terms: Who can define the set of available annotation terms?**
PGP.

21. **Does the company pay the annotators for annotating content?**
No.

22. **Does the company publicly endorse the annotators?**
yes, but only by user name.

23. **Does the management force them to do? In other words, what motivation do workers have to annotate content?**
No.

Questions about the prospected *usage* of the annotations

24. **When searching for content, what does the user need to retrieve?**
See question 8.

25. **Do you intend to offer query expansion strategies?**
no idea; we have to tell Carl!

26. **Should a search/annotation interface resolve ambiguities?**
No.

27. **Should the user be able to explicitly refer to the property of the resources the user query refers to?**
I do not know.

28. **When searching for a resource, should the user be able to find resources that have a certain relationship with another resource?**
Yes.

29. **Should the scope of the search be limited to a specific group of users (e.g., within a particular department) or a single user?**
No.

30. **Are all the resources and annotations always available to everyone? or should there be some access control features?**
Yes.

31. **Are the users going to navigate the whole set of resources or just do searches?**
just searches; landscapes (use case 3): probably browsing.

32. **Should the user be able to navigate from a given resource to other semantically related resources?**
Yes.

33. **If they navigate, are they interested in seeing a detailed classification/taxonomy (such as by country, then by region, then by city, etc.) of the resources or a rough clustering (such as the tag cloud example in the report)?**
Maybe.

34. **Are the users interested to see statistics about the resources and their annotation?**
Maybe. I do not know at the moment

35. **Will the same people that make the annotations, search based on the annotations? Are annotators and end user different?**
Depends on the use case.

Questions about experiences and existing annotation mechanisms in the applications of the partners

36. **Do you have experiences with annotation mechanism? If yes: which?**
Yes in SALERO.

37. **Do you already use annotation mechanisms in your application(s)? If yes: could you please explain them and the drawbacks you see in them?**
No.

C Questionnaire and Results for SEEKDA Use Case

Information about *what* to annotate

1. **What is the content type you need to be annotated ?**
   Web Services; Providers (Service + all the related content; different stakeholder)

2. **What are the content formats used?**
   WSDL; might be that we point to the functional aspects; point to terms; you could have a description in the service; other aspects are not in WSDL.

3. **Do you need annotations on the entire resource, or on a part of it ?**
   Both.

4. **Do you want to annotate specific properties of the content?**
   Yes: rights, title, etc.

5. **How many resources need to be annotated ?**
   any service that we have; couple of thousands.

6. **Should the user be allowed to annotate only his own resources, resources only of a group he belongs too (from a department for example) or can everyone annotate any resource?**
   This depends on the user; some properties can only be annotated by a provider of the service (identified by the domain) WIKI entries can be annotated by any authenticated user

7. **Why is the annotation of specific content important? What advantages do you think to obtain from annotated knowledge/information?**
   for Seekda: to get a better structured understanding from the services (can not be automatically deduced); human can provide content that the machines not can do; scalability!

Information about *how* the annotation should look like

8. **How would you like to use the annotations after they are created?**
   navigate: surely; (search by category); nice visual, graphical interface to display the annotations (readable by the human) "Is the service free/payable?" -¿ filter the services.

9. **Do you intend to use predefined annotations provided to the annotator, and what are they?**
   predefined annotations very much! GoodRelations ontology to be used as a starting point.

10. **Can the set of available annotation terms be extended dynamically (or automatically), and by whom?**
    the only stakeholder that can extend them: engineer in the company; free defined vocabularies not intended!; user can create instances, not concepts!

11. **Can the annotators use completely new annotations that are outside of a predefined controlled vocabulary?**
    No, except of tagging.

12. **What is the provenance information you need for an annotation?**
    Who and when (to keep track of the value).

13. **Do you intend to keep track of the version history of the annotations ?**
    Yes, absolutely!

14. **What complexity do you expect the annotations to be?**
    link to existing structured knowledge, but also phrases (comments, WIKI entries).

Questions about *who* is supposed to provide the annotations and about their *motivation*

15. **What are the characteristics of your annotators?**
some technical people; 4 groups in the use cases (technical advanced); IT literate.

16. **How much time you expect your annotator to spend on a single resource?**
owner of the service: quite strong involvement, because it is his resource! considerable effort ordinary users: minimal time, 1-2 mins

17. **Do you expect a single resource to be annotated by one or many annotators?**
Many.

18. **What is the expected amount of annotations per resource per annotator?**
As many as possible;

19. **How many annotators you expect per resource?**
couple.

20. **If you want to offer pre-defined annotation terms: Who can define the set of available annotation terms?**
Engineers inside of SEEKDA.

21. **Does the company pay the annotators for annotating content?**
No.

22. **Does the company publicly endorse the annotators?**
Good idea; past: no; future: yes

23. **Does the management force them to do? In other words, what motivation do workers have to annotate content?**
The management is not involved in the annotations.

Questions about the prospected *usage* of the annotations

24. **When searching for content, what does the user need to retrieve?**
All related documents; see the annotations from others.

25. **Do you intend to offer query expansion strategies?**
Never thought about it; would be nice surely; why not.

26. **Should a search/annotation interface resolve ambiguities?**
Never considered it; but nice feature; why not

27. **Should the user be able to explicitly refer to the property of the resources the user query refers to?**
Yes.

28. **When searching for a resource, should the user be able to find resources that have a certain relationship with another resource?**
Not really.

29. **Should the scope of the search be limited to a specific group of users (e.g., within a particular department) or a single user?**
Not really.

30. **Are all the resources and annotations always available to everyone? or should there be some access control features?**
Resources: yes; annotations: restricted.

31. **Are the users going to navigate the whole set of resources or just do searches?**
just searches;

32. **Should the user be able to navigate from a given resource to other semantically related resources?**
Yes.

33. **If they navigate, are they interested in seeing a detailed classification/taxonomy (such as by country, then by region, then by city, etc.) of the resources or a rough clustering (such as the tag cloud example in the report)?**
Yes.

34. **Are the users interested to see statistics about the resources and their annotation?**
No.

35. **Will the same people that make the annotations, search based on the annotations? Are annotators and end user different?**
Yes.

Questions about experiences and existing annotation mechanisms in the applications of the partners

36. **Do you have experiences with annotation mechanism? If yes: which?**
tagging in the portal; automatic: a lot of them; user annotation: textual description for every service+provider; tags with user+service; rate service+provider; comment about use; retrieve comment

37. **Do you already use annotation mechanisms in your application(s)? If yes: could you please explain them and the drawbacks you see in them?**
Drawbacks? Until you don't really have the huge traffic they do not work.

D Questionnaire and Results for Telefonica Use Case

Information about *what* to annotate

1: What is the content type you need to be annotated ?

The resources are mainly textual, in the form of web content from the intranet. In particular blog entries and forum posts.

There are also news videos that should be annotated.

2: What are the content formats used? Mainly HTML web pages from the intranet, as well as the news videos.

3: Do you need annotations on the entire resource, or on a part of it ?

The minimal requirement is on the whole resource but annotations of parts of text or videos would be interesting (e.g. sentences, paragraphs and video segments).

If segments can be annotated, it would be interesting to also annotate relations between them. For example, if a sentence talks about "Insemtives project" and a later one about "the University of Trento", it would be good to annotate this one with a "partner" link to the "Insemtives project".

4: Do you want to annotate specific properties of the content?

Useful but not the main requirement. The interest is in annotating content.

5: How many resources need to be annotated ?

A lot, it could be hundreds or thousands.

6: Should the user be allowed to annotate only his own resources?

It will depend on the resource. In most of the cases, everyone can access and annotate all the resources (in particular for news videos and blog entries). However in the case of some documents, the access should be limited to a small group of users, in general the ones contributing to the document.

7: Why is the annotation of specific content important? What advantages

To search and navigate the content.

In particular, it should be easy to navigate within the metadata for "faceted" discovery of resources.

Information about *how* the annotation should look like

8: How would you like to use the annotations after they are created?

For search and navigation. But seeing statistics of the use of particular annotations, and by whom is a must have.

9: Do you intend to use predefined annotations provided to the annotator, and what are they?

No, they would provide free text annotations. There might be a recommendation system.

added later: A bootstrapped vocabulary might be provided by "experts" but it's not limited.

10: Can the set of available annotation terms be extended dynamically (or automatically), and by whom?

the set of predefined annotation is optional and the users enter free text tags, the vocabulary is extended by everyone or automatically.

11: Can the annotators use completely new annotations that are outside of a predefined controlled vocabulary?

yes, free text annotations.

12: What is the provenance information you need for an annotation?

Who did it and When the annotation was given is a must have for auditing of the annotation.

13: Do you intend to keep track of the version history of the annotations ?

yes

14: What complexity do you expect the annotations to be?

mainly free text, but attributes and relations would be good. *there was a need to explain the difference between annotation complexity and vocabulary type.*

Questions about *who* is supposed to provide the annotations and about their *motivation*

15: What are the characteristics of your annotators?

The content creators are also annotators, they are experts in their field and in the field of the content, but not experts in annotation.

16: How much time you expect your annotator to spend on a single resource?
short time, the annotation process should be quick.

17: Do you expect a single resource to be annotated by one or many annotators?
it depends on the resources. For most, many annotators can annotate the resource. For specific documents with Access Control, only a small group (collaborators) will annotate.

18: What is the expected amount of annotations per resource per annotator?
small as it has to be quick.

19: How many annotators you expect per resource?
many, everyone might annotate every resource.

20: If you want to offer pre-defined annotation terms: Who can define the set of available annotation terms?
A small set of experts (in the content and in the annotation system) might provide a controlled vocabulary to bootstrap recommendations etc. Ultimately, the vocabulary is not controlled (but moderators might check the annotations).

21: Does the company pay the annotators for annotating content?
there is an existing award system (including money price) for creating content. The annotation creation will also provide awards.

22: Does the company publicly endorse the annotators?
NA: annotation is done by employees for internal use.

23: Does the management force them to do? In other words, what motivation do workers have to annotate content?
there is no penalties for not annotating, only awards for creating content.

Questions about the prospected *usage* of the annotations

24: When searching for content, what does the user need to retrieve
retrieve content, navigate content and annotation.

25: Do you intend to offer query expansion strategies
yes

26: Should a search/annotation interface resolve ambiguities
yes, this is even more important than the previous question.

27: Should the user be able to explicitly refer to the property of the resources the user query refers to?
yes, the user must be able to search by attribute/property and filter a search to get more specific results.

28: When searching for a resource, should the user be able to find resources that have a certain relations with another resource?
yes, also for search.

29: Should the scope of the search be limited to a specific group of users (e.g., within a particular department) or a single user?
as discussed earlier depends on the resources. Simple Access Control system by group.

30: Are all the resources and annotations always available to everyone or should there be some access control features?
same as previous answer

31: Are the users going to navigate the whole set of resources or just do searches?
navigation of metadata and annotation to find specific content.

32: Should the user be able to navigate from a given resource to other semantically related resources?
yes

33: If they navigate, are they interested in seeing a detailed classification/taxonomy (such as by country, then by region, then by city, etc.) of the resources or a rough clustering (such as the tag cloud example in the report)?
both, the user provides free text annotation, but the user might access through taxonomy (provided by experts). When annotating, the user might take time to disambiguate a free text word to a concept in the taxonomy.

34: Are the users interested to see statistics about the resources and their annotation?
yes

35: Will the same people that make the annotations, search based on the annotations? Are annotators and end user different

different

Questions about experiences and existing annotation mechanisms in the applications of the partners

36: Do you have experiences with annotation mechanism? If yes: which?

the users have no prior experience of complex annotation systems, except for a simple tagging interface.

The users will not be trained in a new tagging system and it should be straightforward so they understand it by themselves.

37: Do you already use annotation mechanisms in your application(s)? If yes: could you please explain them and the drawbacks you see in them?

We already use a basic free text tagging system.

[end of document]

