

# A Synchronous Multimedia Annotation System for Secure Collaboratories

Ronald Schroeter, Jane Hunter, Jonathon Guerin, Imran Khan, Michael Henderson

*The University of Queensland*

*St Lucia, Qld, Australia 4072*

*+617 33652097*

*{ronalds,jane,guerin,imrank,mjh}@itee.uq.edu.au*

## Abstract

*In this paper, we describe the Vannotea system - an application designed to enable collaborating groups to discuss and annotate collections of high quality images, video, audio or 3D objects. The system has been designed specifically to capture and share scholarly discourse and annotations about multimedia research data by teams of trusted colleagues within a research or academic environment. As such, it provides: authenticated access to a web browser search interface for discovering and retrieving media objects; a media replay window that can incorporate a variety of embedded plug-ins to render different scientific media formats; an annotation authoring, editing, searching and browsing tool; and session logging and replay capabilities. Annotations are personal remarks, interpretations, questions or references that can be attached to whole files, segments or regions. Vannotea enables annotations to be attached either synchronously (using jabber message passing and audio/video conferencing) or asynchronously and stand-alone. The annotations are stored on an Annotea server, extended for multimedia content. Their access, retrieval and re-use is controlled via Shibboleth identity management and XACML access policies.*

## 1. Introduction

A *Collaboratory* is the term used to describe the networked environment that allows geographically dispersed teams of researchers to work together. Using information and communication technologies, researchers can access each other's remote data and information, software, tools and instruments that they require to carry out their research. Key components of collaboratories are: robust security mechanisms; multi-user video-conferencing services; shared data

repositories; instant messaging services; application sharing and shared analytical tools and services. Many research communities are generating research data in multimedia formats - images, video, audio, 3D objects and animations (dynamic data visualizations, simulations and models). As a result, research communities are increasingly seeking tools that enable them to collaboratively analyse and annotate such multimedia content, either synchronously or asynchronously. Examples of communities generating multimedia content that requires collective analysis and knowledge capture through shared annotation tools include:

- Oceanography projects such as Visions 2005 [1], which is generating hundreds of hours of underwater video and data streams that require real-time analysis by teams of scientists;
- CIMA X-ray Crystallography portal [2], which is generating video streams, data streams and x-ray diffraction patterns that when processed, produce 3D crystal structures;
- The Paradisec project [3], a web-enabled facility for collaborative digitisation, management and access to Australian researchers' ethnographic audiovisual recordings of endangered languages and music from the Asia Pacific region.
- The UK eScience Integrative Biology project [4] which is generating MPEG movies showing computer simulations of electrical propagation of the heart and the growth of cancerous tumours.

The work described in this paper focuses on the development of a secure collaborative multimedia annotation system for collaborators within eResearch environments – that can be used either asynchronously or synchronously. An essential requirement for this sector is the need to authenticate the source of the annotation and to restrict access to a particular group of trusted colleagues, for reasons of privacy, confidentiality or intellectual property. This is particularly important within eScience, where the

annotation or interpretation of the raw document or data, is often more valuable than the target of the annotation.

Our implementation involves combining and extending a number of existing open source technologies based on open standards:

- Annotea [5] – an annotation protocol developed by the W3C, as part of the Semantic Web initiative. We have extended Annotea to support annotation of fine-grained contexts within multimedia objects;
- Jabber [6] – this provides the instant messaging required for the real-time application sharing and event logging ;
- Shibboleth [7] – an Internet2 middleware initiative that enables identity management and secure access to Web resources shared amongst a federation of organizations ;
- XACML [8] (eXtensible Access Control Markup Language) – XML-based language for defining and enforcing access control policies .

## 2. Background and Previous Work

Existing annotation tools (which enable users to attach personal notes, questions, explanations, etc. to documents) can be categorized according to the media types which can be annotated (text, web pages, images, audio or video, 3D) and the extent of collaboration supported.

There are a large number of systems available [22, 23, 24] – particularly for the annotation of web pages and images. Annotations vary from simple semantic tags to rich, structured annotations (free text, hyperlinks, ranking, language, audiovisual). The systems also vary in their ability to attach annotations to fine-grained segments or regions. Some systems only allow annotations to be attached to whole files or specific types of segments, such as keyframes. The functionality of interfaces for searching, browsing and presenting annotations also vary considerably. Few of the systems surveyed provided advanced search mechanisms over different types of structured annotations or annotation metadata (e.g., creator, data, language) – requirements of many researchers.

We found the majority of existing tools to be asynchronous. Of the systems that do support synchronous annotations, they generally do it through shared whiteboards, rather than application sharing. A limitation of this approach is that the annotations are saved as an image which makes search and retrieval of the annotated content highly problematic.

None of the systems reviewed provided a flexible enough architecture to support a wide-range of high

quality media types through easy incorporation of alternative embedded plug-ins. Such media types include text, html, images, video, audio, 3D. Vannotea is capable of all of these.

Our survey also showed that some systems support private local access only whilst others permit sharing amongst groups and/or public access through storage on a web-accessible server. However none of the surveyed systems provide the level of robustness or the fine-grained role-based access control mechanisms that are required by collaborative teams of scientists engaging in eResearch.

## 3. Usage Scenario

Below is an example of a usage scenario which illustrates how all of the key features of Vannotea are typically utilized within a collaborative eResearch/eScience Marine Science project.

A Marine Biologist from the University of Queensland (UQ) browses through the latest collection of deep water video footage, which was made available by the Visions'05 project at the University of Washington (UW). UQ and UW are both part of a Shibboleth Federation. Visions'05 has set up access policies that allow staff members from the Marine Biology department at UQ to gain full access to their online repository as part of their collaboration.

The Marine Biologist opens a particular recent video within Vannotea so he can bookmark and attach his personal notes to segments of the video, keyframes or regions within frames. To secure his notes, he reuses an existing policy that grants access to all of the participants of this collaboratory. Through this policy, his notes are securely stored on an Annotation Server at UQ. Whilst the Annotation Server is shared amongst other departments at UQ and other members of the Federation, his annotations are only visible to members of the collaboratory.

The Marine Biologist notices a close up of a unrecognizable tubeworm species in a hydrothermal vent being filmed in one of the videos. He highlights it with one of the drawing tools and posts a question “unknown species of tubeworm?” to the annotation server. It is early in the morning, and his Jabber contact list indicates that his colleague, an Oceanographer at UW, is still online. He fires up a videoconferencing tool to speak to him. The Oceanographer starts Vannotea, and is invited to join a Jabber conference room by the Marine Biologist. This triggers an event that opens up the same video at the same location inside the Oceanographers Vannotea Client. All annotations by this collaboratory are also retrieved,

including the “unknown species of tubeworm” annotation the Marine Biologist posted earlier.

The Marine Biologist hits the record button. This records their conversation and the application events they fire whilst collaboratively watching and browsing the same video content. They start looking for other occurrences of the tubeworm while discussing the matter. Every time they see one, they pause the video, highlight the region and store it as an annotation reply of type bookmark to the “unknown species” annotation on the Annotation server. As a result of their discussion, they not only get a list of all locations of the unidentified tubeworm within the video, but they also narrow the classification down to three different possible species of tubeworm. They terminate the session and the Marine Biologist uploads the recorded session, including the video/audio and a time stamped log file of the application events, to a shared repository. He also creates an annotation that links to the audiovisual recording of the videoconference session.

A few hours later, another member of the collaboratory at the University of London starts his work day. He checks his RSS feeds and notices new annotations about an “unknown species of tubeworm” on the annotation servers that he has access to. A simple double-click opens the annotation and the video inside Vannotea and jumps to the frame that shows the worm. He goes through the list of bookmarks that point to the different occurrences. Still unsure, he retrieves the discussion and replays the complete session between his two colleagues earlier. Curious, he starts researching online publications relating to the three possible worm species and is able to reject them through his findings, which he links by posting further replies.

By using Vannotea, the globally distributed collaborators are able to analyse, annotate, and share their knowledge about the multimedia research data, either synchronously or asynchronously without compromising security. There is a possibility that they have discovered a new species of tube worm that only survives in high temperatures and deep sea conditions of hydrothermal vents. This finding can be kept confidential until such time that they are ready to be publicised.

## 4. System Implementation

Figure 1 shows an overview of the system architecture. Vannotea has been implemented in C# using the .NET 2.0 Framework. The color scheme illustrates how the different server side services are being accessed by different UI components within Vannotea.

The content provider(s) on the left can either be positioned inside (secure) or outside the Shibboleth Federation (publicly accessible). Vannotea can easily incorporate search interfaces to both kinds of multimedia research repositories. In the centre of Figure 1 are two researchers from different institutions that are part of a Shibboleth Federation. The researchers use video conferencing tools and the Vannotea Client to collaboratively discuss, analyze and annotate the scientific multimedia content they have retrieved from one of the content providers. Their Vannotea clients are synchronized through Jabber messages, which ensure that both researchers have exactly the same view of the multimedia content at the same time.

On the right we have the home institutions of the two researchers. Their universities provide access to services that include:

- Shibbolized Annotea Server, where they can upload and retrieve annotations;
- Jabber Server for instant messaging;
- Shibboleth Identity Provider (IdP), which authenticates users when they log on.

The IdP also releases attributes about the user to the Shibboleth Service Providers (SP) within the federation, e.g., to the content providers or the Annotea Servers, so they can make decisions regarding access rights.

### 4.1 The content provider

The content providers provide access to collections of domain-specific scientific multimedia data. They provide search, browse and retrieval interfaces to their collections through existing web portals or web applications using pre-existing metadata e.g. transcripts of linguistic data, technical specifications of microscopy images, provenance information etc.

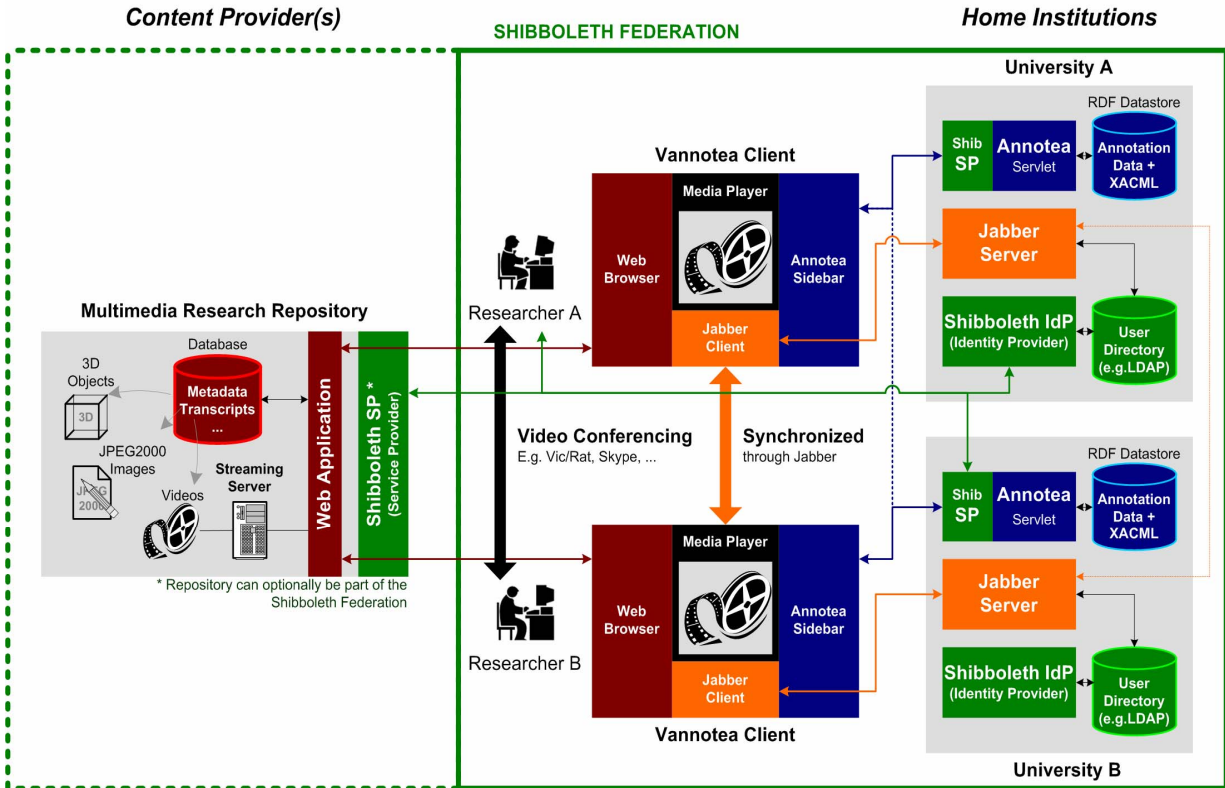


Figure 1: System Architecture

The content may be publicly accessible or protected through site-specific user accounts. An emerging mechanism for restricting access to institutional research repositories is via Shibboleth. Institutions join a Shibboleth Federation and establish a Shibboleth Service Provider (SP). This controls sharing of data and services between institutions. Users can use SingleSignOn (SSO) from their home institution to log in and gain access to data or services within the Federation. We have chosen to use Shibboleth to implement the authentication and access control over our annotation servers.

Through the integration of a Wiki, the metadata of the content could be directly edited by the convening domain experts discussing the matter, e.g. to change the actual transcript of an interview.

The multimedia content can be either hosted on a Web Server, in which case the whole file is downloaded and cached on the client side by the appropriate player/viewer, or – for video content - can be streamed without downloading through any one of the available streaming servers, such as Darwin Streaming Server, Windows Media server, Video Lan Server, etc. Generally, the larger the file size of the video, the more essential it is that the video is streamed.

We deliberately let the above decisions to be made by content providers. Vannotea’s functionalities can be applied independent of the domain, discipline, security, quality, format and network requirements.

#### 4.2. Shibbolized Annotea Server

Annotations are subjective, external comments, notes, reviews or references that can be attached to whole files, segments or regions. They can be in the form of free text, URLs or local files such as images or PDF files that are then uploaded to the server.

Annotea is a flexible Web-based annotation system that uses Resource Description Framework (RDF) to model annotations as a set of statements or assertions made by the author. Our Annotea server is implemented using a Tomcat Java Servlet. RDF is stored using the Jena API [9] and a MySQL database.

It is important to note that the annotations are stored **separately** from the content. This allows any community of users to discuss any collection but still maintain control over access to these discussions. Annotation servers also vary in the extent of their distribution. They may be:

- public and centralized for general access;

- private and centralized: installed on a single server and password protected by a specific group;
- searched in a distributed fashion: searches are performed transparently across institutional servers and the results automatically aggregated.

**4.2.1. Extended Annotea.** Each annotation stores basic metadata information including: the author, date of creation, title, type, which resource it annotates, the context within the resource (e.g. an XPointer to paragraph within a html document) and the content of the annotation. Additional information has been added to each annotation: a unique creator identifier (provided by the Shibboleth IdPs), a link to an XACML policy governing access to the annotation, and a Scalable Vector Graphics (SVG) description storing the coordinates of region outlines drawn on top of images and video frames.

**4.2.2. Shibboleth and XACML.** In addition to the operations defined by Annotea (posting, querying, downloading, updating, replying and deletion of annotations) the Annotea server and protocol have been extended to support access control policies.

Users can define XACML policies (see 5.3.3 XACML Policy Creation) and attach them to annotations. Policies are stored within the RDF repository, along with annotation bodies. They can be created either during the posting of a new annotation or independently of an annotation. Annotations are linked to particular policies through their policy property – which specified by a URL. This approach has the benefit of enabling multiple annotations to use the same policy. If a policy is modified, the changes will affect all those annotations associated with that policy.

Policies are applied when a user tries to retrieve an annotation. The user's attributes are securely retrieved through the Shibboleth SP installation which queries the Shibboleth IdP for the user's attributes. The attributes are delivered through Secure Assertion Markup Language (SAML) assertions, which are translated into an XACML Request within the Annotea Servlet. Using Sun's XACML API, an access decision

is made based on the XACML Request and the XACML Policy, eventually either denying or granting the user access to an annotation.

Policies can be defined that allow users to be aware of the existence of annotations (e.g. list all annotations of a specific user) but do not reveal the actual content. This is important, because it allows the server to return results for statistical purposes, e.g., *how many annotations does this resource have?*

**4.2.4. Notification.** A problem often identified in relation to the use of annotations in asynchronous collaboration [10, 11] is maintaining awareness of change. Cumbersome methods of dealing with this issue include manual checking or notification via email. To address this issue, we have incorporated a simple feature which outputs annotation server changes as RSS. This is achieved by transforming the Annotea RDF using XSLT. By querying the modified annotation server in the normal fashion (e.g. give me all annotations by a creator for a particular resource) within an RSS feed reader, users can keep up to date with any changes that may have occurred. This approach also ensures that the results of the RSS feed are filtered in exactly the same way as the secure annotation server, thus ensuring protection is maintained.

### 4.3. Jabber Server(s)

We use the Open Source Wildfire Server [12] for communication between Vannotea clients. It is built upon the widely adopted open protocol for instant messaging, XMPP (also known as Jabber). Jabber enables two or more entities on the Internet to exchange messages, presence, and other structured information in close to real time. Jabber messages pass through firewalls and proxies, so no additional setup is required by users or system administrators.

We also believe that Jabber represents the future for instant messaging. Wildfire integrates neatly into existing authentication systems, such as LDAP. Jabber can use the same LDAP directory as the Shibboleth architecture, hence allowing single sign-on.

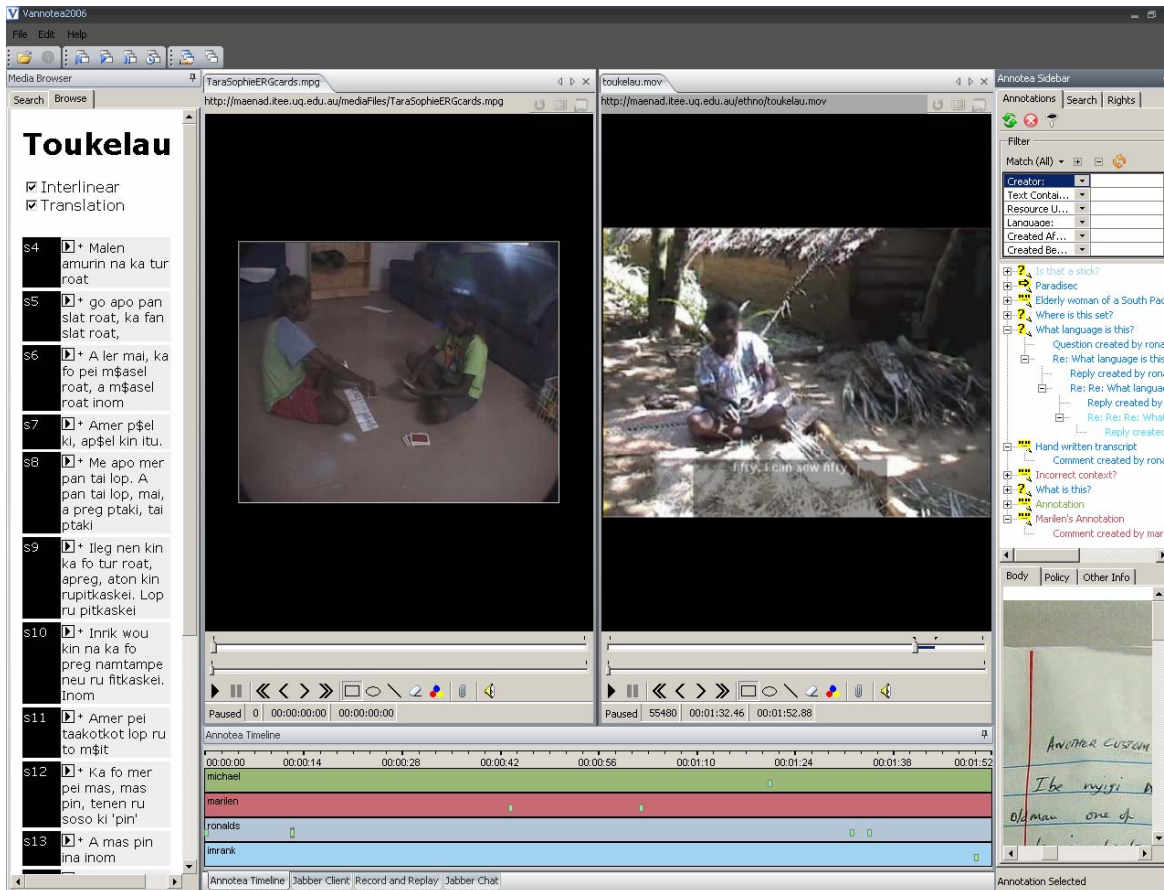


Figure 2: Vannotea User Interface

## 5. User Interface

A screenshot of Vannotea's UI can be seen in Figure 2. It shows Vannotea being used within the Paradisec Ethnographic Analysis Project. The UI consists of the following embedded components: A Web Browser (Internet Explorer); Various embedded media players (such as Quicktime); the Annotea Sidebar and the Jabber Conferencing Client.

### 5.1. Web Browser

The embedded web browser is used to access web enabled video repositories. Alternatively, users can drag and drop multimedia links from an external web browser. We embedded a Web browser to enrich user-friendliness and functionality by tailoring the web application for Vannotea. Simple JavaScript can be added to the web application that allows communication between the web site and Vannotea. By clicking a segment, the web application passes the information about the media file (URL), and the segment (id, start and end point) to Vannotea, which then jumps to the start and plays the segment. Communication also works in the other direction, e.g.

Vannotea can call scripts to retrieve and highlight current segments.

### 5.2. Media Players

Media players display the annotatable multimedia content. By using various embedded ActiveX controls, such as Quicktime and Windows Media Player, Vannotea supports the majority of video and audio formats. We also support crystallographic structure 3D formats, using JMOL [13].

Commonly required functionalities (that depend on the media type), such as play, pause and frame step, are provided. For 3D objects, users can pan around, zoom in and out and rotate the object.

Additionally a transparent drawing panel is layered on top of every player enabling the user to draw simple shapes on top of the content to point out or highlight important regions. The drawing information is stored as SVG and can be drawn collaboratively.

### 5.3. Annotea Sidebar

The Annotea Sidebar has also been implemented in C# and .NET. In addition to its use in Vannotea, it can

be installed as a sidebar plug-in for Internet Explorer, allowing users to annotate web pages asynchronously.

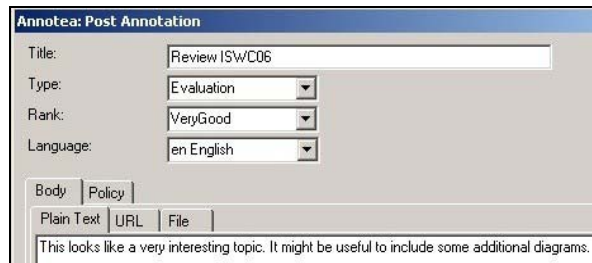
**5.3.1. Annotation Retrieval.** The Annotea Sidebar automatically displays a list (or tree for discussion threads) of all the annotations that refer to the current media file. Using the W3C Annotea Protocol over HTTP, it retrieves the annotations from a list of servers that the user has specified during system configuration. The sidebar only shows the annotations that the user has access to.

For time-continuous media files, the annotations' *context* points to specific timestamps or temporal segments. As the time information is not reflected in the list/tree view, we implemented a graphical timeline to display annotations across the duration of the media file. The timeline is constantly synchronized with the annotations listed in the sidebar. This gives the user an instant overview of the annotations on the video or animation. This provides an indication of the most important and relevant segments.

The different tracks on the timeline are by default grouped by the different authors, but they can also be grouped by annotation type or creation date.

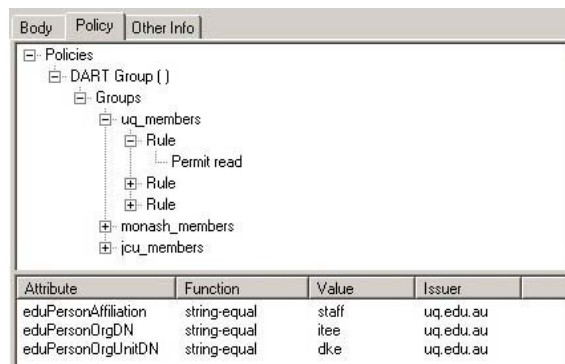
The annotation server supports searching the body of annotations as well as complex queries across the annotation metadata. For example, *show me annotations by a particular author, between given dates and containing the term "Ethnography"*. This functionality on the server side is accessible through the search interface which allows users to specify searches quickly and intuitively. Although we have chosen to implement the search user interface within the browser sidebar, we are also considering the development of a web portal style interface which users can access without having the sidebar installed. This web portal will provide a broader overview of the distribution of annotations across collections, thus indicating information which is of greater interest and relevance.

**5.3.2. Annotation Creation.** Figure 3 illustrates the user interface for creating and attaching an annotation to a digital multimedia object. We have extended Annotea to support structured annotations that contain a number of fields including hyperlinks, files, free text or controlled vocabularies through drop down boxes.



**Figure 3: Creating or Editing an Annotation**

**5.3.3. XACML Policy Creation.** Figure 4 shows the interface developed to define policies. It consists of two main parts; the definition of access control rights to a particular user group and the definition of user groups (based on particular eduPerson attributes [14]). More complete details of the access policy can be found in [15].



**Figure 4: User Interface for defining access policies**

## 5.4. Jabber Client

The embedded Jabber Client, like any other chat client, manages a list of contacts and their presence information. It provides all of the common functionalities such as adding new contacts and deleting them, allowing/disallowing users to see presence information and inviting other users into a conference room or group chat for collaboration and discussion over the same material. We used Jabber-NET, to access the Jabber Server and handle Jabber messages.

## 5.5. Event Logging and Replay

The event logging and replay is part of the Jabber Client. By clicking the record button during a collaborative session, all the incoming Jabber events

are being time stamped and stored in a simple log file. The information of the log file is displayed along a user friendly timeline. Clicking on any of the events displayed in the timeline, replays the session from that event.

## 6. Evaluation

We are currently working with researchers from the Paradisec project and the Crystallography portal, carrying out user evaluation studies of the Vannotea system within these two very different disciplines. User feedback has on the whole been very positive. The ethnographers requested the ability to display multiplexed annotations on the same segment. The crystallographers requested the ability to annotate strands or ligands, as opposed to just single atoms.

## 7. Future Work and Conclusions

In the immediate future we are planning to investigate the following aspects of and extensions to Vannotea:

- More detailed and extensive user evaluation studies;
- Semantically enriched annotations through ontologies;
- A repository of synchronized event logs and video conference recordings that can be searched and replayed.

This paper has described a secure, collaborative synchronous annotation service for multimedia, developed by combining and extending a number of existing open source technologies. The proliferation of eResearch activity has led to a demand for such a service across many disciplines. By providing researchers with the necessary support for authenticating the source and protecting the confidentiality and intellectual property of their annotations, they will be more willing to share their views and engage in inter-organizational collaborations with trusted colleagues. Moreover, the modular design and interoperable technologies that we have adopted, makes it easy to quickly adapt the server to a variety of different media types, different domains and different communities.

## 9. Acknowledgements

The work described in this paper has been funded by the GrangeNet program which is a component of the Australian Federal Government's BITS Advanced

Network Initiative (Department of Communications, Information Technology and the Arts).

## 10. References

- [1] University of Washington, "Visions 05 Project," <http://www.visions05.washington.edu/>.
- [2] Indiana University School of Informatics, "CIMA X-ray Crystallography portal," <http://iri164.informatics.indiana.edu:8080/gridsphere/gridsphere>.
- [3] Paradisec, "The Paradisec Project," <http://paradisec.org.au/>.
- [4] "The UK eScience Integrative Biology Project," <http://www.integrativebiology.ac.uk>.
- [5] J. Kahan, M.-R. Koivunen, E. Prud'Hommeaux, and R. R. Swick, "Annotea: An Open RDF Infrastructure for Shared Web Annotations," in *Proceedings of the WWW10 International Conference*, Hong Kong 2001.
- [6] JabberStudio, "Jabber-NET," <http://jabberstudio.org/projects/jabber-net/project/view.php>.
- [7] Internet2, "Shibboleth Project," 2005, <http://shibboleth.internet2.edu/>.
- [8] M. Lorch, S. Proctor, R. Lepro, D. Kafura, and S. Shah, "First experiences using XACML for access control in distributed systems," in *Proceedings of the 2003 ACM workshop on XML security* Fairfax, Virginia ACM Press, 2003.
- [9] B. McBride, "Jena: A Semantic Web Toolkit," *IEEE Internet Computing*, vol. 6, pp. 55-59, 2002.
- [10] J. J. Cadiz, A. Gupta, and J. Grudin, "Using Web annotations for asynchronous collaboration around documents," in *Proceedings of the 2000 ACM conference on Computer supported cooperative work*. Philadelphia, Pennsylvania, United States: ACM Press, 2000, pp. 309-318.
- [11] A. J. B. Brush, D. Barger, J. Grudin, and A. Gupta, "Notification for shared annotation of digital documents," in *Proceedings of the SIGCHI conference on Human factors in computing systems*. Minneapolis, Minnesota, USA: ACM Press, 2002, pp. 89-96.
- [12] Jive Software, "Wildfire Server," <http://www.jivesoftware.org/wildfire>.
- [13] "Jmol," <http://jmol.sourceforge.net/>.
- [14] Directory Working Group (MACE-Dir), Internet2 Middleware Architecture Committee for Education, "EduPerson Object Class Specification (Draft)," 2006, <http://www.nmi-edit.org/eduPerson/draft-internet2-mace-dir-eduperson-latest.html>.
- [15] I. Khan, R. Schroeter, and J. Hunter, "Implementing a Secure Annotation Service," in *IPAW06*, Chicago 2006.