

# Synote: Enhancing Multimedia E-Learning with Synchronised Annotation

Yunjia Li, Mike Wald, Shakeel Khoja, Gary Wills, David Millard, Jiri Kajaba, Priyanka Singh,  
Lester Gilbert

Learning Societies Lab

Department of Electronics and Computer Science

University of Southampton

Southampton SO17 1BJ, UK

+44 (0)23 8059 3667

{yl2, mw, sk07r, gbw, dem, jk2, ps3, lg3}@ecs.soton.ac.uk

## ABSTRACT

This paper discusses the development of a collaborative hypertext application named Synote. Multimedia has become technically easier to create, but while the WHOLE of a media resource can be easily bookmarked, searched, linked to and tagged, it is still difficult to find or associate notes or other resources with a certain PART of a resource. Synote addresses the important and pervasive user need of making multimedia Web resources easier to access, search, manage and exploit for learners teachers and other users through developing an application that supports the creation of synchronised notes, bookmarks, tags, links, images and text captions. Synote has been developed based on requirements for such a Web based e-learning environment, identified by various studies. The performance, efficiency and effectiveness of Synote was evaluated by using it with students and teachers for a number of undergraduate courses and this has shown that they like using Synote and want more recordings and lectures in this format to be made available to them.

## Categories and Subject Descriptors

K.3.1 [Computers and Education]: Computer Uses in Education – collaborative learning, distance learning; H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems – hypertext navigation and maps.

## General Terms

Management, Design, Experimentation, Human Factors, Standardization.

## Keywords

Synote, Synnotation, Synmark, multimedia annotation, tags

## 1. INTRODUCTION

With the rapid development of the Web, many types of multimedia resources, such as audio and video, have become available. The Web, which once was regarded as only holding

read-only media with note taking or modification of the actual document barred [1], is swiftly being evolved into a global hypermedia system, having users developing content through collaboration with note taking and metadata insertion services for almost all types of media. This new evolution allows users to (1) create and use external hypermedia structures imposed on stored documents, (2) annotate, link and share these documents storing links and annotations separately, (3) create mash-ups using various distributed services and (4) develop customised views for better presentation and adoption. Continuous media such as audio and video has always been considered difficult to handle on the Web due to the complex nature of their data, but their presence has seen a steady increase. With the advent of video storage sites such as Youtube<sup>1</sup>, Blinkx<sup>2</sup>, etc, it has become extremely easy for Web users to upload and share their videos. This has led to a challenging demand to find easy and efficient solutions to handle these vast data archives.

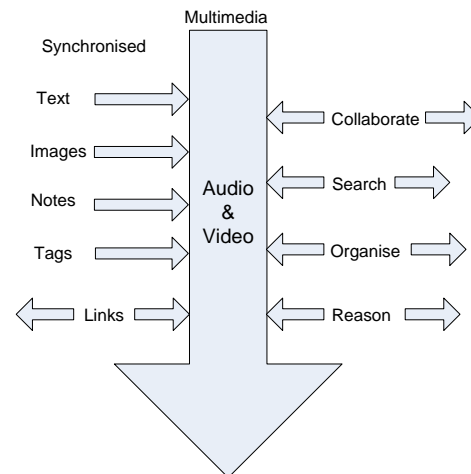


Figure 1. Features of Synote System

While the WHOLE of a media resource can be easily bookmarked, searched, linked to and tagged, it is still difficult to find or associate notes or other resources with a certain PART of a resource. As an analogy, users would clearly find a text book

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

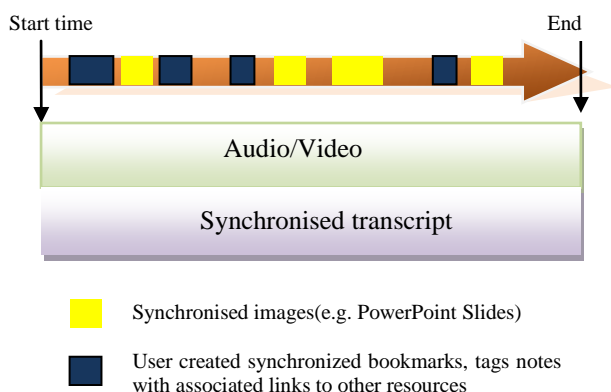
MTDL '09, October 23, 2009, Beijing, China.

Copyright 2009 ACM 978-1-60558-757-8/09/10...\$10.00.

<sup>1</sup><http://www.youtube.com>

<sup>2</sup><http://www.blinkx.com>

difficult to use if it had no contents page, index or page numbers. Therefore, the growing amount of knowledge available in multimedia format has yet to achieve the level of interconnection and manipulation achieved for text documents via the Web. So in e-learning, there are exciting improvements that can occur in 'Web 2.0' and 'social software' environments. This paper therefore addresses the important user needs of making multimedia web resources easier to access, search, manage, and exploit for learners and teachers through developing technologies which support the creation of synchronised notes, bookmarks, tags, links, images and text captions. The application discussed is called Synote. In Synote, the synchronised bookmarks, containing notes tags and links are called Synmarks. The Synchronised Annotations will be referred to by the new term 'Synnotations' to denote the link between two synchronised resources. Figure 1 and 2 schematically show the system that has been developed.



**Figure 2. Time-based View of System**

This paper further explores the idea of annotating multimedia and how Synote is designed to benefit students and teachers in e-learning environments. Section 2 of the paper reviews literature on the history of annotating multimedia. Section 3 describes the user requirements identified by various approaches of studies. Section 4 explores other annotation applications which only to some degree meet the requirements mentioned in section 3. Section 5 and 6 discuss the design, implementation and evaluation of Synote. Section 7 gives the conclusion and presents some future work to further enhance Synote.

## 2. ANNOTATING MULTIMEDIA

Annotation in multimedia is a means of marking up the objects and scenes in audio-video streams in order to facilitate the interpretation and understanding of its content. Use of annotation is divided into two main categories: metadata association and content enrichment. Metadata association methods use specific metadata models and Content Based Retrieval (CBR) data to build a semantic structure to support features such as content search, etc. MPEG-7<sup>3</sup> is an example of such a system. MPEG-7 offers a comprehensive set of audiovisual description tools to create descriptions which will form the basis for applications enabling effective and efficient access (search, filtering, browsing) to the media content. This approach requires the user to understand the underlying semantic metadata structure model in order to perform annotations that conform to the framework, which is a tedious and

time consuming task. The content enrichment annotation uses other high level elements such as text, tags, bookmarks, audio and video to enrich the multimedia objects in a media stream and generate a new composite stream. Users add annotations through collaboration, which are used for searching and aggregation. This method is also used to assist users to understand the media object more comprehensively and thoroughly.

Hypertext is defined as a collection of documents (or nodes) interconnected by links, so that a user can go easily from one place to another, as well as create links between them. Additionally hypertext can be defined as the combination of content and the structure that holds the content. So annotations are considered as a fundamental aspect of the inception of hypertext. Work on annotations in the hypertext world started with the development of Memex by Vannevar Bush [2]. Memex allowed the reader of a document to speedily transverse a vast corpus of literature, to annotate pages and to store the trails taken through this landscape of text. These annotations were shared by other trail blazers (Memex users) discovering new relations between texts and publishing them for other readers to use [3]. Englebart in the Sixties developed NLS/Augment [4] which provided its users with a collaborative, multi-window system with video conferencing, where every media could be linked to, and supported collaborative document creation (like wikis). Ted Nelson's Xanadu takes a transcursive [5] approach in which new hypertext seamlessly assimilates portions of older writings and augments a capacity for Journal systems commentary (like blogs). Despite the early emphasis on annotation, there remains substantial work to be done. Many systems were geared toward the initial design of a hypertext, not toward the annotation of existing materials [6]. Rosenberg [7] commented on these systems as writing-oriented systems, rather than reading-oriented systems. While the role of reader as navigator is widely acknowledged in the hypertext community, the reader as annotator is a far less common emphasis [6]. Web link services such as Microcosm [8], CoNoter [9], Devise Hypermedia [10], provided some features for reader annotation. In Intermedia [11] the notion of a Web as collections of links, became a useful tool for collaboration. Intermedia handled multiple users with different permissions, so that one person could create documents, available only for reading and linking by others and later publish to others.

Recently the notion of metadata has gained huge popularity through the introduction of standards such as XML and RDF. Metadata is seen as a potential means to create intelligent web applications that not only contain additional information about hypermedia but also the rules about how to use this information for intelligent and effective use. Web 2.0 has provided the ability to gather metadata from users through collaboration. Marshal [12] has argued that user generated annotations and structures are useful metadata for other users of a digital library collection who need to pursue research in the same territory as other colleagues have covered earlier. Open Hypermedia Systems research [13, 14] has been working since the early 1990s on providing general support for user controlled annotations and structuring which can be kept separate to the documents containing the information content. Mimicry system [15] is one of the efforts that allow authors and readers to link to and from temporal media (video and audio) on the Web. Web applications such as Overstream<sup>4</sup>,

<sup>3</sup><http://www.chiariglione.org/mpeg/standards/mpeg-7/mpeg-7.htm>

<sup>4</sup><http://overstream.net>

Tegrity<sup>5</sup> and Panopto<sup>6</sup> enable users to take notes that are synchronised with media on the Web. YouTube has also introduced the facility of annotating videos, along with writing comments and star rating the videos.

### 3. SYNOTE REQUIREMENT ANALYSIS

#### 3.1 Identification of User Requirements

Although annotating multimedia is not a new field of research and many applications have already made use of the annotations, it is still problematic when considering the efficiency and accuracy of the annotation. For example, on Youtube, users can easily find a whole video clip by searching the tags provided by the uploaders, but often users are not interested in the whole content, but a particular segment within the video. So it is an important user need to annotate part of the recording, so that the recording is easier to access, search and manage via these annotations. In the e-learning environment, some studies have been done to reveal how teachers and learners can benefit from making this kind of annotation to help search and manipulate recordings of lectures and link other multimedia resources on the Web to the lectures. These studies also considered the usability and accessibility issues as well as solutions.

Consultations within the University of Southampton, including the University's podcasting user group, have confirmed the user needs, interest in involvement in the concept and synergy with the University's road map to support and empower learners in their studies. Dr Wald in collaboration with the international Liberated Learning Consortium conducted a number of qualitative and quantitative user studies involving questionnaires, interviews, focus groups and observations with learners and staff that established the importance of text captions for searching multimedia transcripts [16]. User needs analyses and evaluations were conducted with learners at the University using both low and high fidelity prototype technologies. Reflection by learners on how they learn using recorded multimedia and how synchronized notes and tags could assist them were carefully analysed. These studies confirmed the importance of captions for searching the recordings, reading the transcripts, the value of being able to personally annotate the recordings (e.g. bookmarks, notes and tags) and searching these annotations. Other research also supports these identified user needs [17, 18, 19, 20]. In addition, the text and annotations need not be English, which is especially useful when English is not the first language of learners and when the materials themselves are not English. Generally speaking, annotations can enable learners:

- to search text transcripts for specific topics and then replay recordings from that point
- to read captions rather than listen to recorded speech to support learning style preference, deafness, or English as a second language
- to use the transcript of lectures to support their understanding of text books and academic articles if they find the more colloquial style of transcribed text easier to follow than an academic written style
- to insert a bookmark at a particular point in a recording so as to be able to continue later from where they left off

- to link to sections of recordings from other resources (e.g. documents, web pages etc.) or share these sections with others
- to tag and highlight sections of recordings/transcripts they don't understand fully so they can revisit them later for clarification
- to annotate recordings with notes and URLs of related resources (e.g. documents, websites etc.) at specific places in a recording to clarify issues and support revision
- to tag recordings using their own terms as a personal index
- to create a community folksonomy 'index' from everyone's tags to help provide a shared understanding and a more detailed and complete index

Annotation can enable teachers and lectures:

- to index their recordings using syllabus topic tags
- to provide synchronized slides and text captions to accompany podcasts
- to identify which topics need further clarification from the pattern of learners
- to provide feedback on learner-created recordings of presentations (e.g. 'good use of voice', 'good use of audio visual material')
- to ask learners to annotate recordings of their group meetings using unstructured or structured tags to provide evidence of their contributions as part of the module assessment
- to analyse unstructured tags learners use (folksonomy) to help create structured tags (ontology) for indexing future recordings
- to tag recordings with URLs of related resources (e.g. documents, websites, etc)
- to link to and use sections of existing multimedia on the Web without having to edit the recording

The importance of the identified user needs of making multimedia accessible and searchable through the creation of synchronised transcripts is supported by Google's encouragement for people to caption their multimedia before uploading it to Google<sup>7</sup> and their addition of a subtitling annotation facility to Youtube. For longer recordings, manual captioning is time consuming and costly but speech recognition has been demonstrated to provide a cost-effective way of automatically creating accessible text captions and transcripts synchronised with audio and video. Dr Wald has collaborated with IBM and the international Liberated Learning Consortium since 1999 to investigate how speech recognition captioning can enhance learning and to develop ViaScribe<sup>8</sup>, which uses speech recognition to create synchronized searchable transcripts, slides and captions from multimedia recordings of live or recorded speech or video. ViaScribe demonstrates the following features, which should be considered usability requirements:

- The text can be highlighted and scroll automatically in time with the speech

---

<sup>5</sup><http://www.tegrity.com>

<sup>6</sup><http://panopto.com>

---

<sup>7</sup><http://video.google.com/support/bin/answer.py?answer=26577>

<sup>8</sup> <http://www.liberatedlearning.com/technology/index.shtml>

- The controls can move backwards or forwards through the presentation or pause it
- The browser 'find' facility can search for text and play the multimedia from that position
- Selecting a slide thumbnail will move the presentation to that position
- The frames and slide thumbnails can be resized

Learners and teachers have found teaching and learning improved using this technology [21]. Portable wireless systems can facilitate personalisation of the display and manual correction can be used when required to reduce speech recognition errors, especially if there are multiple speakers [22]. But since voice recognition software is not widely used, it is still necessary to include some other mechanisms to allow users to type in and synchronise a transcript manually, e.g.:

- Allow users to type in a transcript and synchronise it.
- Allow users to upload a transcript text and then synchronise it.
- Provide support for popular captioning file formats (e.g. .srt file), so that users can create a synchronised transcript using existing captioning software and then upload this transcript.

### 3.2 Scenarios of Synote Usage

Based on the requirements identified by the investigations, two typical example scenarios are provided to depict the use of Synote:

*"Scenario 1: Mike records a narrated PowerPoint of his lecture. After the class, he uses software downloadable from the Synote website to automatically create a clickable index from the slide titles and a transcript of his presentation, which is synchronised with the recording of his voice and the slides. Then he easily uploads these resources into Synote and makes it accessible to his students (a certain group of users) so that they can add annotations. He quickly adds some questions students should think about for next week as well as the URLs of some other resources (including a section of a recording he made the previous year). These 'Synmark' resources are all automatically synchronised with the relevant parts of the lecture recording. When Sally, a first year student, sits down at her computer, all the resources Mike uploaded are available on Synote. She can search the slides, text transcript and notes for specific topics via the search engine provided by Synote. Then a single click leads her to the replay of recordings from that point and currently spoken words are automatically highlighted. She annotates the recording at specific points with notes made from the text book to aid revision for the exam and she also tags and highlights a section of the transcript she doesn't understand fully, so that Mike can clarify it for her and then prints out the information she needs. Before going to lunch she inserts a synchronised bookmark into the recording and on her return continues from where she left off."*

*"Scenario 2: When surfing on the internet, Mike finds a video clip on Youtube which is related to his lecture. As Synote player can directly play video clips from Youtube, he creates a new recording pointing to the video clip. One of Mike's students John watches the video and particularly likes the content between 2:10 to 2:20. So he adds a Synmark Synnotation with start and end time of 2:10 and 2:20. John also thinks that there are many students whose*

*first languages are not English in his class. So it might be difficult for them to understand the content of the video clip. As Synote player provides a transcript editing function, he simply adds the transcript of the video by hand in Synote player and synchronises it with the video."*

These scenarios help give an example about a few of the ways Synote might be used in real e-learning situations.

## 4. INVESTIGATION OF EXISTING ANNOTATION APPLICATIONS

Investigations of existing annotation software (including trials and discussions with developers and users) have shown the widespread recognition of the importance of the identified user need of supporting the creation of annotations.

The Memetic<sup>9</sup> (Meeting Memory Technology Informing collaboration) project's overall aim was to extend the functionality of the next generation AccessGrid<sup>10</sup> collaboration environment with advanced meeting support and information management annotation tools that were developed in CoAKTinG<sup>11</sup>(Collaborative Advanced Knowledge Technologies), which drew on the extensive multimedia retrieval, open hypermedia knowledge and experience gained through ECS Southampton's earlier research (e.g. Microcosm[8]). JISC and NSCF digital libraries programme funded the development of Project Pad<sup>12</sup> to assist learners in interacting with digital multimedia archive recordings. The NXT<sup>13</sup> set of libraries and tools was developed for the manipulation, query and analysis of multimedia language data while Vannota [23] is a research tool based on the W3C Annotea project [24]. Microsoft OneNote, Tegrity and Panopto enable learners to take notes that are synchronized with recordings. Then they can select a note and replay the recording from that point. Proprietary non-standard formats<sup>14</sup> requiring proprietary media players can be used to create enhanced podcasts by embedding markers in a multimedia file to jump to predetermined spots in a presentation. There are also some video annotation tools available online (e.g. Youtube, PLYmedia<sup>15</sup>) but they don't support the full required functionality.

As mentioned in section 3.1, a transcript is an important and necessary annotation to the multimedia. Many systems have been developed to address the problem of creating synchronised transcripts. Project Readon<sup>16</sup>, Overstream<sup>17</sup> and dotSUB<sup>18</sup> have been developed to display manually created captions for videos that are however non-searchable. Everyzing<sup>19</sup> uses speech

<sup>9</sup><http://www.memeticvire.net>

<sup>10</sup><http://www.accessgrid.org>

<sup>11</sup><http://www.aktors.org/coacting>

<sup>12</sup><http://www.jisc.ac.uk/whatwedo/programmes/dlitc/spokenword.aspx>

<sup>13</sup><http://www.ltg.ed.ac.uk/NITE>

<sup>14</sup><http://docs.info.apple.com/article.html?artnum=301880>

<sup>15</sup> <http://www.plymedia.com>

<sup>16</sup> <http://projectreadon.com>

<sup>17</sup> <http://www.overstream.net>

<sup>18</sup> <http://dotsub.com/>

<sup>19</sup> <http://www.everyzing.com/>

recognition to create a searchable transcript from podcasts or multimedia recordings. It provides further evidence for the importance of the identified user need and the value of using voice recognition software. However, Everyzing does not display fully synchronised captions and so does not meet the identified user accessibility needs.

## 5. SYNOTE DESIGN

### 5.1 Features of Synote

Considering the user requirements and the insufficiency of current annotation systems, it is therefore necessary to include the following features in Synote from a technical point of view:

- Separate the storage of multimedia resources from annotations. Make full use of the existing multimedia on the Web.
- Synchronise captions, images, tags, links, notes and bookmarks and other resources under the timeline.
- Users are allowed to add, edit, delete and search for annotations quickly and easily [25].
- Supports user authentication and permission control over all of the annotations and resources.
- Introduce usability and accessibility concepts into the system design from the very beginning, especially to overcome browser and platform compatibility problems.

Instead of storing copies of the original video and audio like Youtube, the recordings are just stored by users on their chosen web accessible space, and the Synote server only stores transcripts and other user created annotations. In this way, millions of hours of recordings can be catered for. The Synote application has the capability to handle most of the types of audio and video formats available on the web. Synote works by synchronising audio and video with accessible text captions and transcripts, allowing audio visual material to be manipulated through searching and browsing the text. Similarly, images, tags, links, notes and bookmarks are also synchronised with audio and video to provide innovative methods of annotating multimedia. Synote enables users to add and search for, annotations quickly and easily. These annotations can be further used for a number of web 2.0 type collaborative applications. Accessibility has been considered at all stages in the development of Synote.

### 5.2 Synote Conceptual Model

The work on Synote started by considering it as an Open Hypermedia application and media can be stored anywhere on the Web. In the planning phase, it was decided to introduce a common framework of hypermedia system, so its contributions could be great. First of all, it will help in keeping Web multimedia separate from embedded links and annotations. Secondly, it will provide generalization of the link (Synotations) from binary to n-ary (multi-headed) links with the realization that nodes, links and composites will be considered as separate first-class components. A simple and clear navigation structure is also defined to link resources together. In this regard, the following components and entities were initially defined to form the key elements of the Synote Conceptual model with inspirations from Dexter model [26] and FOHM [27]:

**Multimedia Resources:** All media resources will have a globally unique identifier, an anchor indexing into the component, a

presentation specification and list of attribute pairs, and are external to the model.

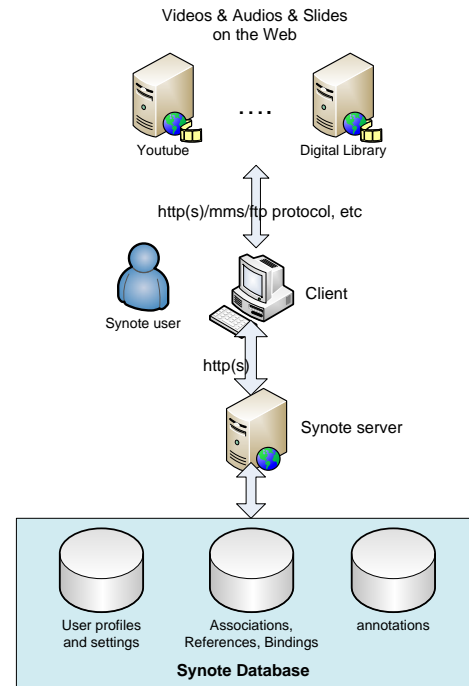


Figure 3. Synote System Overview

**Associations:** Associations are structures that represent relationships between data objects (media objects in Synote).

**References:** Reference objects are used to point at Data objects or at parts of data objects.

**Bindings:** Bindings specify the attributes of the connection between Association and Data objects through Reference objects.

**Link:** A relationship between two anchors, stored in the same or different database.

**Presentation Specification:** Determines how different media components should be presented to users.

It was believed that it would be possible to expand on the above defined model in such a way that the application will be able to integrate with the resources on the Web. Figure 3 shows the general overview and the interaction between different networked components. The Synote server manages the links for multimedia resources, stores associations, references, bindings, annotations, user profiles and communications with clients' browsers.

The key characteristics of open hypermedia systems includes multi-user access to documents and hypermedia structures, access rights to read, annotate and write, collections of hypermedia structures, span to span bidirectional links and a large object oriented application framework to facilitate extensibility. These characteristics should be implemented in the Synote system for annotating, bookmarking, transcription and synchronization into its navigation structure, interface and storage.

### 5.3 Synote Object Model

Figure 4 shows the Synote object model, which is derived from the entities and components mentioned in section 5.2. Annotation and Resource are the main entities in the object model. As the



names suggest, Annotation entity is for handling annotations, while Resource entity is used for storing media object bindings and associations.

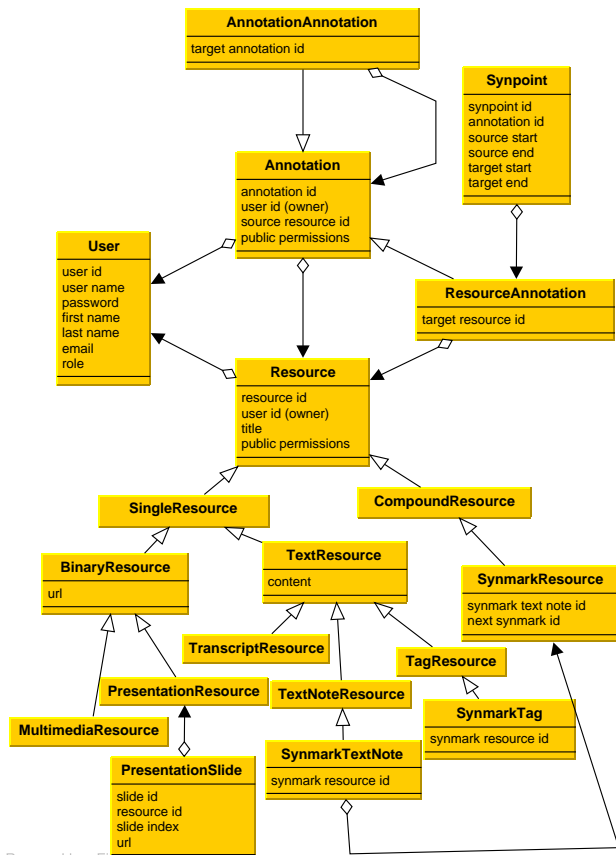


Figure 4. Synote Object Model

A Single Resource entity is used to store the URL for a media/slide or a simple note of type text. This note can be a section of the text of transcript, a tag or a Synmark description. A Compound Resource will contain multiple occurrences of text resources, such as combination of a tag and description, etc. Compound Resource is resolved and brought to the Single Resource level in order to make all the entities of Resource at the same level. Next, Annotation and Resource tables are bundled so that all their entries are brought at the same level, making all as first class elements to be used for combined querying and searching. Objects related to user authentication and permissions are not included in this figure.

Figure 5 demonstrates an example of an annotation between two resources: multimedia resource and Synmark resource. Synmarks are specially designed high level metadata storage for Synote. A Synmark has a unique ID, title, description that can contain valid HTML code, tags separated by spaces and a link to the next Synmarked part of the recording. As the annotation data is stored separately from the media URL, it can be re-used or re-linked to another media. Synmarks can be searched and queried. Currently, the downloadable format for these annotations is not available in Synote, but it is possible to convert them to a certain standard, such as MPEG-7 or SMIL, if further interoperability requirements are identified.

ResourceAnnotation is the core object of navigation structure mentioned in section 5.1. Synpoint (synchronisation point) is a data structure to store the synchronisation information, including the start and end references of the “source resource” and the “target resource”. In this example, the “source resource” is the multimedia resource and the Synmark resource is the target. The Synmark resource annotates the multimedia resource between the start time and end time. A Synpoint object can also store start and end references for target object. If the target resource is a transcript text, the start and end reference will be the start and end characters of the text. Synpoint object is kept aside from the main conceptual data model. This is due to the fact that it can contain data regarding synchronisation of media and text highlighting in transcripts. For example, in one recording of a 40 minute lecture the speaker has spoken 5454 words comprising 29386 characters, creating 10908 start and end time entries, and 2420 synchronising points. If all these entries are treated as first class elements and are used for querying and searching along with media URLs and annotations data, this will create a huge overhead and will seriously affect the performance of the system. Since these entries are used only to provide highlighting of text for transcript and synchronising media, it was decided to keep these entries out of any retrieval activity.

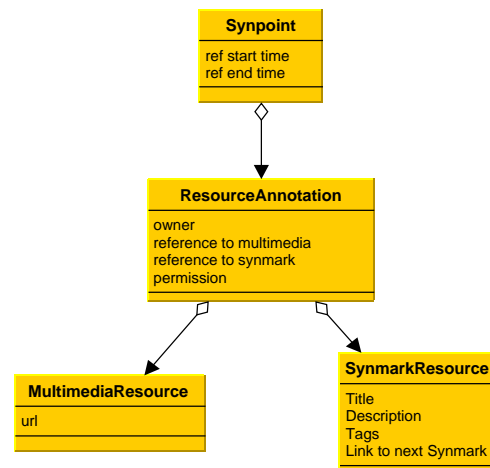


Figure 5. Synote Link Navigation Structure

## 5.4 Synote Client Interface Model

The Synote client interface, which is also referred to as Synote player, comprises of four modules: player module, transcript module, Synmark module and slide module.

The player module can play most of the audio video formats available on the Web, such as WAV, MP3, AVI, MPEG and flash videos. Play, stop, pause, rewind and resize controls are provided by the player. The transcript module maintains the synchronisation of captions with media and the spoken words are highlighted, which is achieved through the Synpoints stored in Synote server. This module’s editing facilities allow users to create Synpoints or even a completely new transcript. This facility allows users to edit the transcript while playing the media. The Synmark module allows users to create and modify synchronised bookmarks. Users can also use a “Synchain” to sequence Synmarks in a linear fashion in order to view only those marked parts of a recording. The slide module displays the related slides, all of which will have start and end times, so that the current slide will automatically be highlighted by the system.

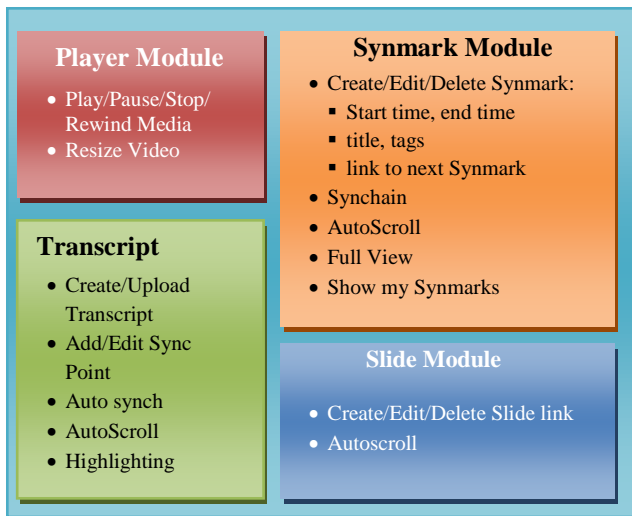


Figure 6. Synote Interface Design Graph

Figure 6 displays the basic interface model of Synote player. In order to satisfy the open hypermedia standards, it has been implemented that the user can create or modify URLs of different type of media with different associations and bindings. The interface allows users to modify and extend upon the existing collection of links and anchors.

## 6. IMPLEMENTATION AND EVALUATION

### 6.1 General Framework

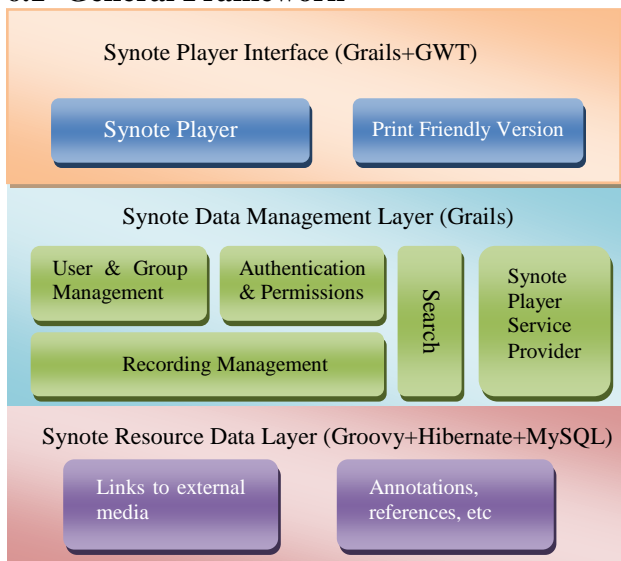


Figure 7. Synote Framework

Synote has been developed using Grails<sup>20</sup>, JAVA and Google Web Toolkit<sup>21</sup> (GWT). Based on the design, the implementation of Synote application can be divided into three layers: resource data layer, data management layer and Synote player (Figure 7).

<sup>20</sup><http://www.grails.org>

<sup>21</sup><http://code.google.com/webtoolkit/>

Resource Data Layer stores links of external media (e.g. video, audio and image slides) and Synmark, transcript resources and annotations. External media are stored in user owned or specified online spaces. When replaying this media, the media stream will be read directly from this link. Synmark, transcript resources and annotations are usually text format information, which are generated by Synote users. In this project, the objects mentioned in Figure 4 are defined as domain classes in Groovy. These domain classes are mapped into MySQL database via Hibernate.



Id	Title	Owner	Public Permission	Your Permission	Replay	Print
318	Organic Chemistry (with slides)	shakeel	PRIVATE	WRITE	Replay	Print
350	Devices and Drivers	shakeel	PRIVATE	WRITE	Replay	Print
373	Quintin Gee's lecture on Professional Issues (animated slide version)	shakeel	READ	WRITE	Replay	Print
376	Quintin Gee's lecture on Professional Issues (with video)	shakeel	PRIVATE	WRITE	Replay	Print
379	Quintin Gee's lecture on Professional Issues (audio only)	shakeel	PRIVATE	WRITE	Replay	Print
413	Smartfreight	shakeel	PRIVATE	WRITE	Replay	Print
427	Smart Freight with transcript	shakeel	PRIVATE	WRITE	Replay	Print
442	David Argle's Lecture 1	shakeel	PRIVATE	WRITE	Replay	Print
463	David Argle's lecture on Security	shakeel	PRIVATE	WRITE	Replay	Print
468	Machine Ethics Framework for Gambling	shakeel	WRITE	WRITE	Replay	Print

Figure 8. Synote Recording List Screenshot

The data management Layer provides a Web interface for Synote users to manage personal information, permissions and multimedia resources. The search engine can retrieve all resource data saved in the database. According to the framework of GWT plugin<sup>22</sup> for grails, this layer also exposes service interface for Synote player. A tool has also been developed to automatically create Synote presentations from narrated PowerPoint files and searchable Synmarks from the slide title, text, notes and timings. It can create a synchronised transcript if this is typed into the PowerPoint slide notes. Figure 8 and 9 show actual screen displays of Synote. Synote Player will be introduced in the next section.

Id:	1
Name:	abc
Shared:	true
Owner:	shakeel
Members:	mike , test
	+ Add Member
Recording Permissions:	Organic Chemistry (with slides) (READ) Quintin Gee's lecture on Professional Issues (audio only) (READ)
	+ Add Recording Permission to a Group
	Edit Delete

Figure 9. Synote Group Permissions Screenshot

### 6.2 Synote Player Implementation

Synote Player Interface comprises of two parts. One is Synote player, which is a typical Ajax application written in Google Web Toolkit (Figure 10). Considering that the heavy javascript implementation of Synote player may result in accessibility issues, a print friendly version of the media and synchronised resources is also implemented.

Synote player is being developed to add Synnotations to the synchronised multimedia captions, images, and slides

<sup>22</sup><http://www.grails.org/plugin/gwt>

demonstrated using ViaScribe [28]. The system has built on the user interface knowledge gained in developing a real-time editing system that enables an operator to select and correct errors in the speech recognition captions as they are displayed [29]. In the prototype users are also allowed to search, browse and add Synnotations (notes, tags, bookmarks, images/slides, links, and captions) to recordings.

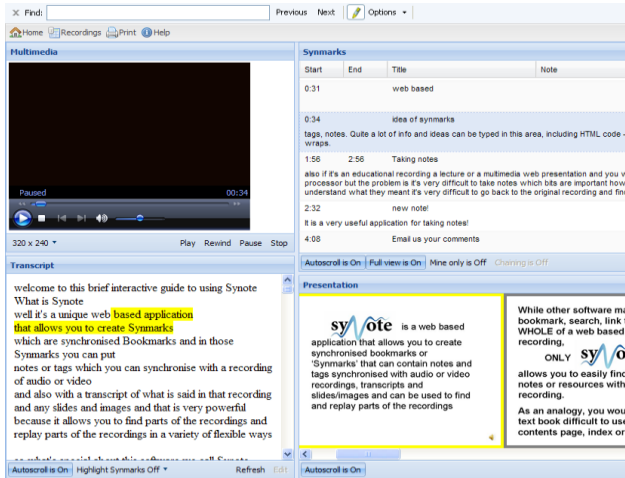


Figure 10. Synote Player Interface Screenshot

In Synote player, when the recording is played, the currently spoken words are shown highlighted in the transcript (Figure 10). Selecting a Synmark, transcript word or Slide/Image moves the recording to the corresponding time. The browser 'Find' facility can be used to search the transcript or Synmarks and a Synote search facility provides for more advanced searching of all the annotations stored on the database. Toggling Autoscroll can ensure the current selection stays visible in the panel or allows user to scroll manually. The displayed size of the slides and the video can be altered by the user, as can the amount of the screen display used to display the transcript or the Synmarks. The start time is automatically entered as the time in the recording corresponding to when the Synmark was created (Figure 11). All of the content or just one line of the individual user's Synmarks or everybody's Synmarks can be displayed and Synmark information can be sorted and hidden. Users can edit or delete their own Synmarks. Users, who are entitled to edit the transcript can edit the transcript manually or upload a .srt caption file within Synote player. Figure 12 shows the transcript editing module. The time icons are used to display the synchronisation points and moving the mouse over them will display the synchronised time. A mouse click on transcript text brings it to the edit box, where the user can edit the transcript text, whereas clicking at the spaces between words will create synchronisation points in the transcript using the current media time.

In the print friendly output, information for the whole time period of the recording or only part of it can be printed and the user can choose to print information for the whole of the recording within this time period or only the Synmarked parts of it and select whose Synmarked parts. The user can also choose whether they want to print the transcript, presentation slides/images and the Synmarks, as well as information in the Synmarks.

Since usability and accessibility issues are considered in the every stage of design and implementation, the compatibility problems are minimized in Synote player. Currently, Synote play can play

most media resources through the use of Windows Media Player, JW Player<sup>23</sup> and QuickTime Player on both Windows and MAC OS platforms. Synote will choose a proper media player according to the format of the file and the client platform. Keyboard shortcuts are supported so that users can accomplish most of the interaction operations with Synote player without touching the mouse. Users can use keyboard shortcuts to play/pause/stop/rewind media player, edit transcripts and create Synmarks. All the shortcuts are carefully designed to avoid confliction with platforms and browsers. The display of Web pages and other interaction operations have been tested in Internet Explorer, Firefox, Google Chrome and Safari browsers.

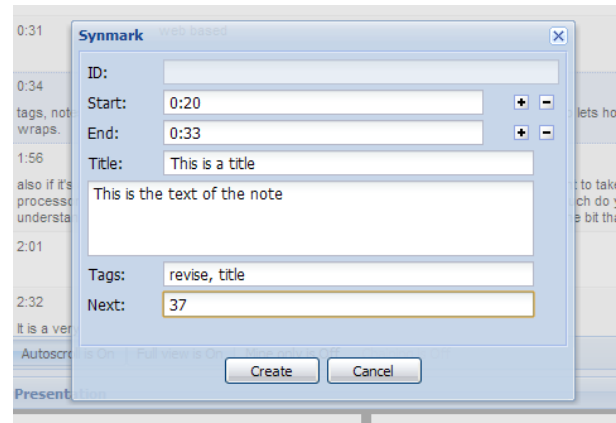


Figure 11. Create Synmark Screenshot

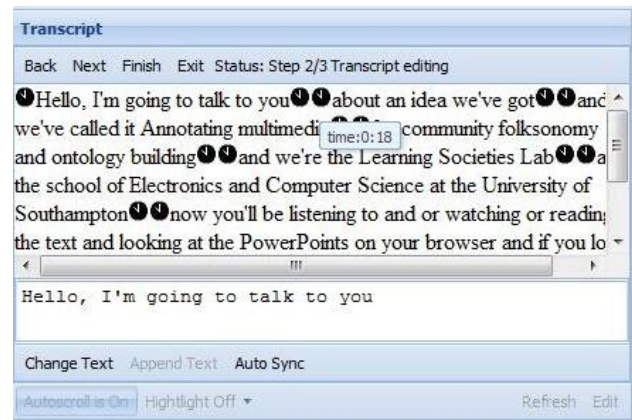


Figure 12. Transcript Editing Screenshot

### 6.3 Evaluation

Synote has been developed and evaluated iteratively with the involvement of users and with the support of JISC<sup>24</sup> and Net4Voice<sup>25</sup>. Generally, initial evaluations using questionnaires and interviews with 120 students in four classes have shown that students and teachers really liked using Synote and found the synchronised transcripts and note taking facilities very useful and wanted more recordings and lectures to be available in this way. Requirement analysis, design and implementation of demonstrators involved usability studies including iterative

<sup>23</sup><http://www.longtailvideo.com/players/jw-flv-player>

<sup>24</sup><http://www.jisc.ac.uk>

<sup>25</sup><http://spazivirtuali.unibo.it/net4voice/default.aspx>



evaluation of screen and interaction design with users. Then this process was used to finalise software design Demonstrators were evaluated using quantitative and qualitative evaluations and observations based on deployment by users.

Very enthusiastic and positive feedback has been received that supports the identified benefits to users (section 3.1) of enabling the different media's communication qualities and strengths to be available as appropriate for different contexts, content, tasks, learning styles, learning preferences and learning differences. Staff from a wide range of schools and support services at the University of Southampton participated in the system and functionality trials of the software as well as members of the JISC Emerge Community, International Liberated Learning Consortium and Net4Voice project. Staff have used the software as part of their normal teaching commitments. Teachers in nursing or health care professions especially like using Synote as it provides a platform for them to monitor their students' recorded progress through providing annotated feedback. There has been a very positive response from students and staff who have said that they would like to continue to use this facility if more material was made available. Students would also want a facility to be able to automatically incorporate notes taken in real time lectures with Synote. Students who were not native speakers especially found the synchronised transcripts of Synote particularly helpful e.g.:

*"Synote gives a very clear understanding of module and lecturer's voice is well enough to understand the main ideas. It was very useful for me especially as I am a non English native speaker."*

Table 1 shows the statistics of feedback, which has helped to improve its usability and accessibility.

**Table 1. Synote Evaluation Feedback**

Description	Result
Use Synote regularly	60%
Browser chosen	IE 63%, Firefox 30%
Think Synote very useful and easy to use	80%
The content of video and audio recordings are very useful	83%
Transcripts are useful	81%
Note taking is a very useful activity	64%

Overall, the survey suggests a very positive feedback about the usability and efficiency of the software and it is hoped that such systems can bring advancement in the learning and teaching practices of the future generation.

## 7. CONCLUSION AND FUTURE WORK

Synote is a solution to introducing multimedia annotation into the e-learning environment, so that teachers and learners can both benefit from the better use of multimedia resources available online. The system works with the operating systems and browsers most used by students and staff and links to web based multimedia while storing the annotations separately. The new terms 'Synmark' and 'Synnotation' were invented to denote the synchronised bookmarks and synchronised annotations creating using Synote. Evaluations have shown that almost everyone has found the synchronised transcripts and note taking facility useful and wanted more recorded material made available for use in Synote.

Even though the current Synote platform meets most of the user requirements defined in section 3.1, more work can be done to improve interoperability and usability. In the evaluation, some students suggested the value of the ability to create Synmarks in real-time. While lectures are making a recording of the lecture in the classroom, students could log into the system and create Synmarks via wireless or even mobile applications. Through mobile devices, it is possible to extend Synote to m-Learning (mobile learning) and integrate it with some mobile learning toolkits, such as mPLAT [30].

Large amounts of multimedia resources are available behind various authentication systems and in order to annotate these resources, Synote user management system should integrate with these authentications seamlessly. It is possible to develop a Synote application specially tailored for a certain environment (e.g. institutions, digital libraries, etc), in which Synote uses the authentication system provided by the environment. Synote can also expose a general authentication service interface. Synote player itself is a heavy javascript application and the interaction speed can be extremely slow sometimes, especially when editing a transcript with hundreds of Synpoints or displaying hundreds of Synmarks at the same page. It is therefore desirable to optimize the design and implementation of Synote player.

Synote is available at <http://www.synote.org> for any author to contribute resources and annotations.

## 8. REFERENCES

- [1] Schilit, B. N., Price, M. N., Golovchinsky, G., Tanaka, K., and Marshall, C. C. 1999. As We May Read: The Reading Appliance Revolution. *Computer*, 32(1). 65-73.
- [2] Bush, V. 1945. As We May Think, *Atlantic Monthly*, Aug 1945, 101- 108.
- [3] Kahn, P., Nyce, J. M. , Oren, T., Crane, G., Smith, L. C., Trigg, R., and Meyrowitz, N. 1991. From Memex to hypertext: understanding the influence of Vannevar Bush. In *Proceedings of the 3rd ACM Conference on Hypertext* (San Antonio, USA, Dec. 1991). HYPERTEXT '91. ACM, New York, NY, 361.
- [4] Engelbart, D. 1963. A conceptual framework for the augmentation of man's intellect. In *Vistas in Information Handling*, volume 1, 1-29, Spartan Books, Washington DC, USA, 1963.
- [5] Nelson, T. 1984. *Literary Machines*. South Bend, Indiana: The Distributors. Edition 87.1 (6th Edition), 1984.
- [6] Marshal, C.C. 1995. Toward an ecology of hypertext annotation. In *proceedings of the 9th ACM Hypertext Conference*, 40-49, 1995.
- [7] Rosenberg, J. 1996. The Structure of Hypertext Activity. In *HYPERTEXT'96: Proceedings of the seventh ACM conference* (New York, USA, 1996). ACM, New York, NY, 22-30.
- [8] Hall, W., Davis, H. C., and Hutchings, G. 1996. *Rethinking Hypermedia: The Microcosm Approach*. Kluwer Academic Publishers, Norwell, USA, 1996.
- [9] Davis, J.R., Huttenlocher, D.P. 1995. Shared Annotation for Cooperative Learning. In *Proceedings of CSCL 95* (Bloomington IA, USA, 1995). L. Erlbaum Associates, Hillsdale, NJ, 84-88.

- [10] Grønbaek, K. and Malhotra, J. 1994. Building tailorable hypermedia systems: the embedded-interpreter approach. In Proceedings of the Ninth Annual Conference on Object-Oriented Programming Systems, Language, and Applications (Portland, Oregon, United States, October 23 - 28, 1994). ACM, New York, NY, 85-101.
- [11] Costa, R. M., Moreno, M. F., and Gomes Soares, L. F. 2008. Intermedia synchronization management in DTV systems. In Proceeding of the Eighth ACM Symposium on Document Engineering (Sao Paulo, Brazil, September 16 - 19, 2008). DocEng '08. ACM, New York, NY, 289-297.
- [12] Marshal, C.C. 1998. Making metadata: a study of metadata creation for a mixed physical-digital collection. In DIGITAL LIBRARIES'98: Proceedings of the 3rd ACM International Conference on Digital Libraries (Pittsburgh, PA, 1998). ACM, New York, NY, 162-171.
- [13] Karousos, N., Pandis, I., Reich, S., and Tzagarakis, M. 2003. Offering open hypermedia services to the WWW: a step-by-step approach for developers. In Proceedings of the 12th international Conference on World Wide Web (Budapest, Hungary, May 20 - 24, 2003). ACM, New York, NY, 482-489.
- [14] Nürnberg, P. J., Leggett, J. J., and Wiil, U. K. 1998. An agenda for open hypermedia research. In Proceedings of the Ninth ACM Conference on Hypertext and Hypermedia: Links, Objects, Time and Space in Hypermedia Systems (Pittsburgh, Pennsylvania, United States, June 20 - 24, 1998). ACM, New York, NY, 198-206.
- [15] Bouvin, N.O., Schade, R. 1999. Integrating temporal media and open hypermedia on the World Wide Web. In Proceedings of the 8th International World Wide Web Conference (Toronto, Canada, May 11-14, 1999), 375-387.
- [16] Bain, K., Hines, J., Lingras, P., Yumei. Q. 2007. Using Speech Recognition and Intelligent Search Tools to Enhance Information Accessibility. In Proceedings of HCI International 2007, In Lecture Notes in Computer Science series (LNCS)
- [17] Whittaker, S., Hyland, P., Wiley, M. 1994. FILOCHAT: handwritten notes provide access to recorded conversations. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Celebrating interdependence (Boston, Massachusetts, United States, April 24 - 28, 1994). ACM, New York, NY, 271-277.
- [18] Wilcox, L., Schilit, B., Sawhney, N. 1997. Dynamite: a dynamically organized ink and audio notebook. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (Atlanta, Georgia, United States, March 22 - 27, 1997). ACM, New York, NY, 186-193.
- [19] Chiu, P., Kapuskar, A., Reitmeief, S., Wilcox, L. 1999. NoteLook: taking notes in meetings with digital video and ink. In Proceedings of the seventh ACM international conference on Multimedia (Part 1) (Orlando, Florida, United States, October 30 - November 05, 1999). ACM, New York, NY, 149-158.
- [20] Brotherton, J. A., Abowd, G. D. Lessons Learned From eClass: Assessing Automated Capture and Access in the Classroom. ACM Trans. Comput.-Hum. Interact. 11, 2 (Jun. 2004), 121-155.
- [21] Leitch, D., MacMillan, T. Liberated Learning Initiative Innovative Technology and Inclusion: Current Issues and Future Directions for Liberated Learning Research. 2003. Retrieved 15, March, 2009, from: Saint Mary's University, Nova Scotia: [http://www.liberatedlearning.com/resources/pdf/RC\\_2002\\_Year\\_III\\_Research\\_Report.pdf](http://www.liberatedlearning.com/resources/pdf/RC_2002_Year_III_Research_Report.pdf)
- [22] Wald, M. and Bain, K. Enhancing the Usability of Real-Time Speech Recognition Captioning through Personalised Displays and Real-Time Multiple Speaker Editing and Annotation. In Proceedings of 12th International Conference on Human-Computer Interaction (Beijing, China, July 22-27, 2007) Springer 2007, 446-452.
- [23] Schroeter, R., Hunter, J., Guerin, J., Khan, I., Henderson, M. 2006. A Synchronous Multimedia Annotation System for Secure Collaboratories. In Proceedings of the 2nd e-Science and Grid Computing Conference (December, 2006), 41.
- [24] Kahan, J., Koivunen, M.-R., Prud'Hommeaux, E., and Swick, R. R. 2001. Annotea: An Open RDF Infrastructure for Shared Web Annotations. Computer Networks, 39(5). 589-608.
- [25] Wald, M., Wills, G., Millard, D., Gilbert, L., Khoja, S., Kajaba, J., and Butt, P. 2008. Multimedia Annotation and Community Folksonomy Building. In Proceedings of World Conference on Educational Multimedia, Hypermedia & Telecommunications (July, 2008), 2213-2220.
- [26] Grønbaek, K., Bouvin, N. O., and Sloth, L. 1997. Designing Dexter-based hypermedia services for the World Wide Web. In Proceedings of the Eighth ACM Conference on Hypertext (Southampton, United Kingdom, April 06 - 11, 1997). ACM, New York, NY, 146-156.
- [27] Millard, D. E., Moreau, L., Davis, H. C., and Reich, S. 2000. FOHM: a fundamental open hypertext model for investigating interoperability between hypertext domains. In Proceedings of the Eleventh ACM on Hypertext and Hypermedia (San Antonio, Texas, United States, May 30 - June 03, 2000). ACM, New York, NY, 93-102.
- [28] Wald, M. 2006. Creating Accessible Educational Multimedia through Editing Automatic Speech Recognition Captioning in Real Time. 2006. International Journal of Interactive Technology and Smart Education: Smarter Use of Technology in Education, 3(2). 131-142.
- [29] Wald, M. 2007. A Research Agenda for Transforming Pedagogy and Enhancing Inclusive Learning through Synchronised Multimedia Captioned Using Speech Recognition. In Proceedings of ED-MEDIA 2007: World Conference on Educational Multimedia, Hypermedia & Telecommunications (Vancouver, Canada, July 2007), 4479-4484.
- [30] Zhang, P., Millard, D., Wills, G., Howard, Y., Faulds, S., Gilbert, L. and Sparks, D. (2008) A Mobile Toolkit for Placement Learning. In 8th IEEE International Conference on Advanced Learning Technologies 2008 (July 1- 5, 2008, Santander, Cantabria, Spain), 92-96.