

# THE ANNODEX PLATFORM

**Shane Stephens**  
*ICT Centre, CSIRO*

Annodex is a platform for the creation and viewing of indexable annotations for media documents. In this paper we will present the Annodex Platform, comparing and contrasting the platform with other tools available to ethnographic researchers. We will examine how some of the key features of Annodex may be of use to ethnographers, and discuss interoperability between Annodex and other tools.

Apart from the generation of annotated content, some of the relevant features of Annodex include:

- the ability to view annotated content over the World Wide Web (WWW);
- the ability to create links to specific time-based features within media documents;
- the ability to create links between features in media documents;
- the ability to associate multiple independent annotation 'tracks' with a single media document.

The Annodex tools, as well as more information, are available from the Annodex project website (*Annodex.net*, n.d.). Annodex components are indicated within this article using *italics*.

## The Annodex Platform

The Annodex platform is a suite of tools designed around the idea of Continuous Media Markup. Continuous Media Markup involves the insertion of metadata (data about events within the media) and other information into arbitrary time locations of multimedia streams. We have built an extensive set of open source tools around this simple but powerful idea.

The platform allows us to host large media files, and access arbitrary time locations within these files as media streams over the World Wide Web (WWW). It also allows us to publish Uniform Resource Locators (URLs), which point to time points of interest, and to associate

information and tags with particular time locations. We have also built a CmmWiki tool that allows real-time online editing and addition of time-point tags to previously published media.

The current set of tools is built around Ogg, a free and open-source media container format. This container format wraps an arbitrary number of content streams into a file, using a loose time-ordering to help ensure that time-related elements are co-located in the file. Our annotations introduce an additional XML content stream formatted in Continuous Media Markup Language (CMML) (Pfeiffer, Parker & Pang, 2005b). CMML provides an abstraction of a 'clip', which is a bounded time period with associated annotations. Each CMML stream contains a series of non-overlapping clips describing the associated media streams. Figure 6 illustrates an example of a clip described in CMML.

```
<clip start="0:00:00.180">
  <a href="http://www.mpi.nl/tools/elan.html">
    Transcoded from ELAN EAF format</a>
  
  <desc>
    oh puzoc a nagdus-suuran / ni gasat a capaitan - oh my heart
    which has been struck by most bitter fate
  </desc>
</clip>
```

Figure 6: Example CMML.

We refer to Ogg files with CMML streams as being in Annodex format (Pfeiffer, Parker & Pang, 2005a). These files also contain some additional information organised in a skeleton stream, which is situated at the beginning of the file. The skeleton stream contains additional useful information about every other stream in the file, such as stream format, image dimensions, language, and so forth.

We mentioned above the ability to refer to Annodex media via URL. Part of this approach is the specification of time points and time segments within a URL. This enables us to craft Web-browser-understandable URLs that refer to specific time locations within media files. The Annodex approach for describing such time positions is described in a third internet draft (Pfeiffer, Parker & Pang, 2005c).

The tools in the Annodex platform can be divided into three groups:

- creation tools, which are responsible for generating Annodex files from media and CMML documents. These tools include *anx-tools* (*anxenc*,

*anxrip*, *anxinfo*, *anxed*, *anxcheck*, *anxgrep*) and oggz-tools (*oggzdiff*, *oggzdump*, *oggzinfo*, *oggzmerge*, *oggzrip*, *oggzscan*, *oggzvalidate*);

- server tools, which are responsible for understanding temporal URLs and generating media streams. These tools include *CmmlWiki* (which also provides online CMML editing facilities) and *mod-annodex*;
- browsing tools, which enable users to view Annodex media and annotations. These tools include the proof-of-concept Firefox plugin, and a third-party open-source Java browser called Cortado.<sup>4</sup>

## Key Features of the Annodex Platform

In this section we discuss key features of the Annodex Platform and their specific relevance to ethnographic researchers.

### View annotated content over the WWW

Annodex integrates with existing browser and server technology, which allows annotated content to be readily deployed and accessed over the WWW. This capability allows widely distributed groups to share content libraries, and also makes possible the establishment of a persistent high-quality reference site containing content and annotations.

An Annodex document is a single object, referenced by a single URL, which contains both annotations and media. This natural and permanent association helps prevent mismatch between annotations and media, and also allows annotations to remain correctly anchored even when streaming from part of the way through a file.

### Create links to specific features within media documents

By using temporal URL resources, researchers in a field that uses audiovisual material as a primary resource can provide consistent and guaranteed access to specific time points within hosted reference material. These URLs can be used in published papers to reference primary materials, and inline<sup>1</sup> browsers such as Cortado can be embedded directly into online versions of these papers, providing the audiovisual equivalent of an image or diagram.

### Create links between media documents

Using temporal URLs, links can be created between specific clips of published media documents. This provides a mechanism for researchers to

specify related content, secondary examples and other auxiliary information.

### Create annotations in already-published media

With the *CmmWiki* server installed, it is possible for annotations to be added to, modified and improved over time, even after media items have been published to the WWW. This approach allows researchers to re-use the same media over the course of several studies, marking up important features as required. It also allows several researchers to make use of the same media documents for several different purposes, without prior knowledge that this collaboration could occur.

### Associate multiple independent annotation 'tracks' with a single media document

We have observed that some existing annotation tools (ELAN, ITE and SignStream)<sup>2</sup> provide the ability to associate multiple annotations tracks (also called tiers) with a single media item. This feature is often used to separate different annotation tasks, and to provide a structure in which phrases can be broken down into words, words into phonemes, and so forth.

The CMML standard and Annodex file format are also capable of supporting multiple annotation tracks. At this stage, we have not implemented multiple track support in the server or browser components of the tool-chain, but these are planned future features.

## Typical Configurations

### Annodex file-hosting

Figure 7 illustrates the components required to convert files into Annodex format, store these files on a server, and make them available for browsing through an Annodex browser plugin. Media files in Ogg format are combined with annotations in CMML format by *anxenc*. If desired, media can first be converted into Ogg format from a range of other formats using the *ffmpeg2theora* tool (*Ffmpeg2theora*, n.d.), or the GStreamer framework (*Gstreamer*, n.d.), and annotations can be converted into CMML from other annotation formats as described below.

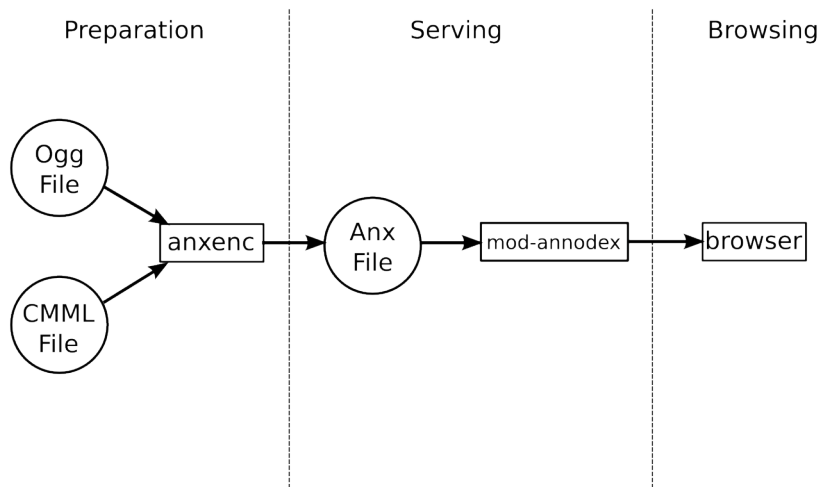


Figure 7: Annodex file-hosting.

Annodex-format files are then uploaded to an Apache<sup>3</sup> Web-server with the *mod-annodex* plugin installed. This plugin enables Apache to understand temporal URI requests, so that media can be streamed from the appropriate time point.

An Annodex Firefox plugin can be used to view and navigate the Annodex files on the server. Figure 8 shows a typical screenshot of the Firefox plugin.

### Request-time Annodexing

Figure 9 illustrates an alternative approach to Annodex file hosting: request-time Annodexing. In this approach, media files are stored on the server in Ogg format, and CMML annotations are stored in a database. When annotated media is requested by a browser plugin, a Web application retrieves the relevant CMML annotation and Ogg file, and both merges and streams these files in real-time, using the *libannodex* library. *CmmlWiki* is an example Web application with this capability.

Although request-time annodexing requires more server infrastructure, it also provides more flexibility. The annotation database allows modification and addition to CMML annotations even after the Ogg file has been uploaded, and search engines are more easily incorporated into the system. These optional features are illustrated in Figure 9 as dotted lines. Figure 10 shows a screen-shot of *CmmlWiki*, using the Cortado player for playback.

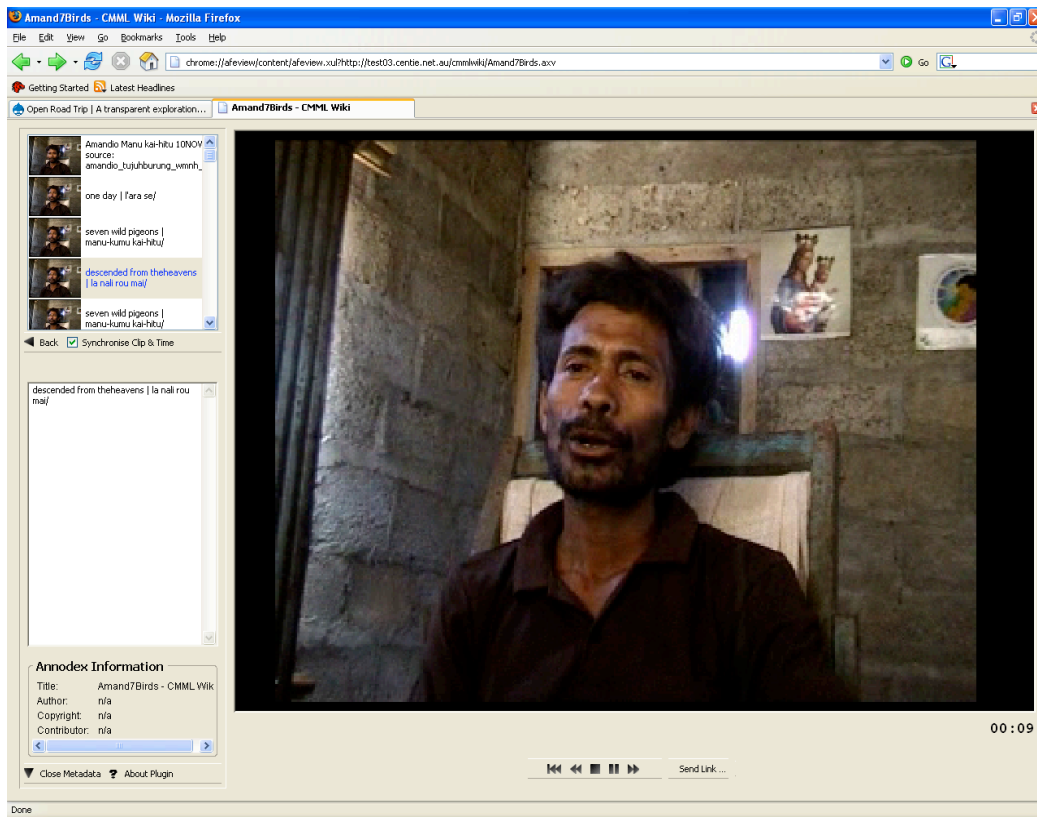


Figure 8: Screenshot of Firefox plugin.

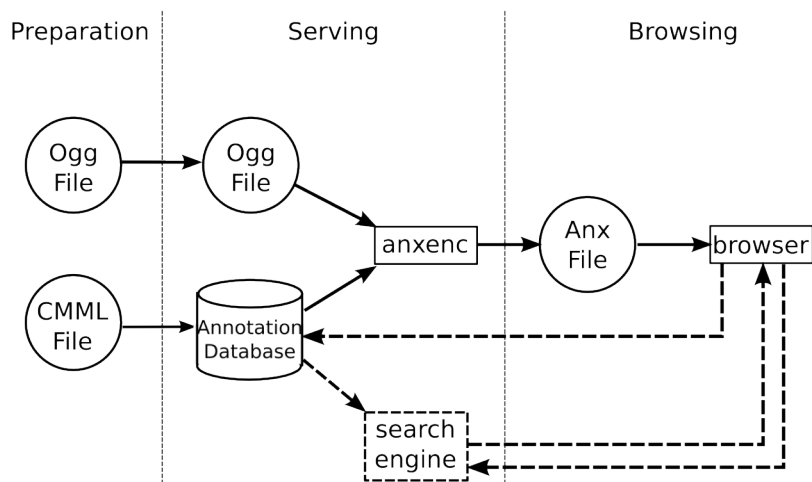


Figure 9: Request-time Annodexing.

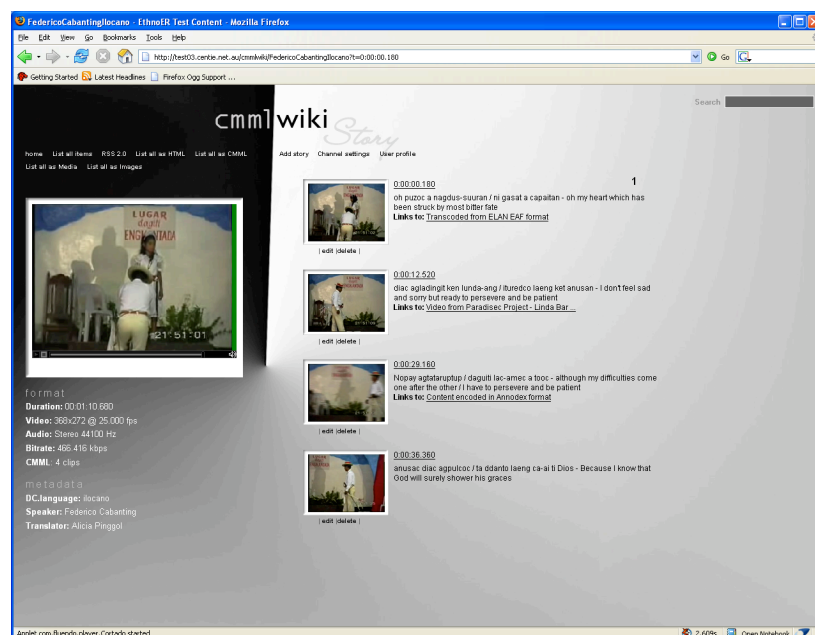


Figure 10: CmmWiki screenshot.

## Comparison with Other Ethnographic Tools

### Annotation tools

Several different tools for generating time-based annotations for a media file exist, many specifically for the purpose of ethnographic annotation and research. These tools include Transcriber, ELAN, ITE, CLAN and SignStream.<sup>4</sup> A disparate range of features is provided by the tools, as summarised in Figure 11.

<i>Tool</i>	<i>Features</i>
Transcriber	Multilanguage support, ease of use, remote access to media, <b>audio only</b>
ELAN	Multilanguage support, linking between annotations, annotation tiers, format converters, searching tools
ITE	Multilanguage support, automatic glossing lexicon construction, annotation tiers (predefined only)
CLAN	Multilanguage support, data analysis and validation tools
SignStream	Multiple tracks, comparison of utterances, searching tools, global data repository, <b>sign-language only</b>

Figure 11: Features List

These tools provide a capability that is not yet available within the Annodex platform—they facilitate the rapid generation of time-coded annotations of media files by domain experts. However, as the tools produce text files with fixed delimiter or XML-based formats, simple scripts can be generated to convert between these formats. Indeed, ELAN already provides scripts that convert between its native EAF format and the Transcriber format. As CMML is itself an XML-based format, conversion to CMML format is a reasonably trivial task. Hence, Annodex is able to integrate with existing annotation tools using simple conversion scripts, which enables researchers to use their favourite (and most productive) tool and still take advantage of the publishing, referencing, and post-publication editing features of the Annodex Platform. Furthermore, existing annotated media can be converted to Annodex format if desired.

The SignStream tool also provides a centralised repository of sign-language data. This collection and distributed availability of annotated media is viewed as an important component of the project, as it allows sharing of examples and more vivid demonstrations of sign-language phenomena that have until now been described by English-language approximations in papers. Transcriber takes a slightly different approach, with a client-server architecture that is capable of supporting file servers local to an organisation or group. This facilitates information sharing strictly between group members. The Annodex platform is capable of providing both types of repository—large public repositories and smaller organisational repositories. Furthermore, as references to media are encoded as URL strings, it is possible to merge several private and public data sources together for the purposes of searching.

### Vannotea

The Vannotea Collaborative Indexing and Annotation platform (Schroeter, Hunter & Kosovic, 2003) is a Windows-based collection of client components that allows video to be displayed, segmented, annotated and discussed simultaneously at several remote locations, and a server component that stores media and annotations, and maintains synchronisation between the clients. Vannotea has several innovative features, including an integrated Jabber client for discussion, the ability to annotate on a graphical overlay that sits on top of the video screen, and a multi-user hierarchical annotation approach that supports persistent annotation dialogues.



Annodex does not currently provide a number of these features, including graphical overlays, integrated chatting and collaborative annotation. On the other hand, the Annodex platform is designed to operate on a range of operating systems and environments (including Linux, Windows and Macintosh), and has been demonstrated to operate on mobile platforms such as 3G phones as well. This capability paves the way for field tools with access to global ethnographic repositories.

Furthermore, the Annodex Platform tightly integrates into Web browsers, and inline browsing elements such as the Cortado plugin permit annotated video to interact with other page components. In this manner, an entirely Web-based annotation approach can be constructed, avoiding the need for local software installation. The existing Annodex Platform components are also open source and hence freely available for use and modification.

Comparing the two systems directly, Vannotea provides remote client synchronisation that is not readily achievable with Annodex, while Annodex provides a more portable and cross-platform approach.

## Conclusion

The Annodex platform has several features that make it potentially useful in the field of ethnographic research. Annodex can be integrated with existing ethnographic transcription tools, provides Web-based publication and browsing features for multimedia content, allows post-publication modification of annotations, and supports searching across annotation text. Annodex also provides URL references to specific time locations or segments, which allows consistent referencing of video or audio within both online and offline documents, including journal publications.

## Endnotes

<sup>1</sup> That is, embedded on an HTML page, along with other contents of the page.

<sup>2</sup> *MPI-PL Tools Elan*, n.d.; *Interlinear Text Editor*, n.d.; Neidle, Sclaroff & Athitsos, 2001.

<sup>3</sup> *Apache*, n.d.

<sup>4</sup> Barras, Geoffrois, Wu & Liberman, 1998; *MPI-PL Tools Elan*, n.d.; *Interlinear Text Editor*, n.d.; *Using CLAN*, n.d.; Neidle, Sclaroff & Athitsos, 2001.

## References

*Annodex.net: Open standards for annotating and indexing networked media.* (n.d.). Retrieved October 23, 2006, from <http://www.annodex.net>

- Apache Software Foundation.* (n.d.). Retrieved October 23, 2006, from <http://www.apache.org>
- Barras, C., Geoffrois, E., Wu, Z., & Liberman, M. (1998). Transcriber: a Free Tool for Segmenting, Labelling and Transcribing Speech. *Proceedings of the First International Conference on Language Resources and Evaluation (LREC)* (pp. 1373-76), Grenada (Spain), May 1998. Retrieved October 23, 2006, from [xml.coverpages.org](http://xml.coverpages.org) website: <http://xml.coverpages.org/Transcriber-LREC1998.pdf>
- Cortado Java Applet.* (n.d.). Retrieved October 23, 2006, from <http://www.flumotion.net/cortado>
- Ffmpeg2theora.* (n.d.). Retrieved October 23, 2006, from <http://www.v2v.cc/~j/ffmpeg2theora/index.html>
- GStreamer Open Source Multimedia Framework.* (n.d.). Retrieved October 23, 2006, from <http://gstreamer.freedesktop.org>
- Interlinear Text Editor.* (n.d.). Retrieved October 23, 2006, from [http://michel.jacobson.free.fr/ITE/index\\_en.html](http://michel.jacobson.free.fr/ITE/index_en.html)
- MPI-PL Tools ELAN.* (n.d.). Retrieved October 23, 2006, from <http://www.mpi.nl/tools/elan.html>
- Neidle, C., Sclaroff, S., & Athitsos, V. (2001). SignStream(TM): A Tool for Linguistic and Computer Vision Research on Visual-Gestural Language Data. *Behaviour Research Methods, Instruments and Computers*, 33(3), 311-320.
- Pfeiffer, S., Parker, C. & Pang, A. (2005a, March). *The Annodex exchange format for time-continuous bitstreams, Version 3.0 (work in progress)*. Retrieved October 23, 2006, from <http://www.annodex.net/TR/draft-pfeiffer-annodex-02.txt>
- Pfeiffer, S., Parker, C. & Pang, A. (2005b, March). *The Continuous Media Markup Language (CMML), Version 2.0 (work in progress)*. Retrieved October 23, 2006, from <http://www.annodex.net/TR/draft-pfeiffer-cmml-02.txt>
- Pfeiffer, S., Parker, C. & Pang, A. (2005c, March). *Specifying time intervals in URL queries and fragments of time-based Web resources (work in progress)*. Retrieved October 23, 2006, from <http://www.annodex.net/TR/draft-pfeiffer-temporal-fragments-03.txt>
- Schroeter R., Hunter J., & Kosovic D. (2003). Vannotea—A Collaborative Video Indexing, Annotation and Discussion System for Broadband Networks. In *Knowledge Markup and Semantic Annotation Workshop, K-CAP 2003*, October, 2003, Sanibel, Florida. Retrieved 21 October 2006 from University of Queensland e-prints archive: <http://eprint.uq.edu.au/archive/00004535/01/schroeter-kcap03.pdf>
- Using CLAN.* (n.d.). Retrieved October 23, 2006, from <http://childes.psy.cmu.edu/clan>