

ANNODEX-ing Broadcast TV News for Semantic Browsing and Retrieval

Alan F. Smeaton, Colum Foley and Kieran McDonald

Adaptive Information Cluster, Centre for Digital Video Processing,
Dublin City University, Glasnevin, Dublin 9, Ireland.
{Alan.Smeaton, Colum.Foley and Kieran.McDonald}@computing.dcu.ie

Abstract

The development of a semantic web of text, images, structured information, and continuous media information including video and audio, depends on being able to annotate content with semantic meaning and then use that content description directly in applications. ANNODEX is a recently announced initiative to allow continuous media files to be integrated directly with their own content description, into one unified source. In this short paper we take a state of the art video indexing, browsing and retrieval system, FÍSCHLÁR-NEWS, whose architecture and design is based on a separation of content from content description, and we re-engineer it using ANNODEX. In doing this we demonstrate the improvements this new approach makes over our system and the increased opportunities it offers in terms of functionality.

1 Introduction

To date much of the work on the semantic web has focussed on text resources – web pages mostly – on automatic content description of those resources and on then building the semantic web on top of those content descriptions. In our work we are interested in content description for multimedia information, specifically video. The conventional approach taken to annotate this media is to separate the actual continuous media content (such as MPEG-1 or MPEG-2 files), from the content description (such as MPEG-7 files). The downside to this approach is that because they are separated, it is impossible to do anything further with the content itself once it is indexed and we rely on creating an exhaustive and accurate content description, which of course is not the case.

ANNODEX [Pfeiffer *et al.*, 2003] (Annotating and Indexing) is an open source family of technologies developed by the Commonwealth Scientific and Industrial Research Organisation for embedding annotations and hyperlinks directly within digital audio and video files. By embedding the semantic description of media within the media itself the combined resource becomes just like any web document which has content and content description bound into one.

Given the potential of ANNODEX we decided to develop a simple ANNODEX-ED version of FÍSCHLÁR-NEWS which

demonstrates the idea of content and its semantic description woven together, and which illustrates some of the possibilities that this offers, including semantic-based linking from within the combined content/description to outside resources.

2 The FÍSCHLÁR-NEWS System

FÍSCHLÁR-NEWS [Smeaton *et al.*, 2005] is a video retrieval system specifically designed for access to video news. It forms a test-bed for evaluating research ideas concerning video news retrieval and interaction. The current system provides search, browsing and video playback for a collection of daily news stories covering over two years.

The video and closed captions (subtitles) of the local terrestrial evening news channel (RTÉ) are automatically captured and processed each evening. Each broadcast is segmented into stories using a Support Vector Machine (SVM). The enriched video description consisting of stories, closed captions, shots and keyframes is encoded as an MPEG-7 description and is imported into our video retrieval system. After capture of a TV news broadcast, the MPEG-1 video file is transferred to a streaming server to provide efficient and direct playback from any point within the video stream.

To view a particular news story a user can click on a keyframe to begin playback from a given point within the broadcast or choose to playback the entire broadcast. In both cases a pop-up video player window displays the video playback and is thereby a dead-end in terms of user interaction with the semantic descriptions of the content. After playback the user must retreat back to the browsing system to select another story to view.

3 ANNODEX-enabled FÍSCHLÁR-NEWS

ANNODEX is developing open standards to enable the building of webs of distributed media known as the Continuous Media Web (CMWeb). In the CMWeb, continuous media such as audio and video files can be searched and browsed in a similar fashion to text. This is achieved through embedding anchor points within the media file and marking up these anchor points with semantic annotations and hyperlinks.

Although the MPEG-7 [2002] standard contained much of the description we needed to annotate our media files, it was not built specifically for web applications and thus does not facilitate embedding links to other resources. SMIL [W3C,

2001], the Synchronised Multimedia Integration Language, another XML based markup language, allows annotation of multimedia data, however it does not provide markup at the level of granularity of a semantic concept, instead markup can only be included for a complete file. For news data this presented a significant problem, as browsing and retrieval could only be achieved at the level of a whole news broadcast, and there was no support for browsing or retrieval at the story level.

We decided to evaluate the applicability of ANNODEx to the development of a system like FÍSCLÁR-NEWS, but with increased semantic linking capabilities to see whether it could be used as a basis for semantic web applications. Using a news story as the obvious semantic unit, for each broadcast we provided story-level segmentation and annotation. We took approximately one month's data from FÍSCLÁR-NEWS broadcasts from February and March 2005. This represents about 500 individual news stories.

Content descriptions for media files are encoded into XML-based Continuous Media Files (CMML). Much of our work was concerned with generating these CMML files, which required a parser for the FÍSCLÁR-NEWS MPEG7 database to extract data related to each news story or <clip>, such as the time offset of this story within the broadcast and a textual description of the story. Within FÍSCLÁR, related stories are generated by performing text-based similarities between the story's subtitle text and storing the results in XML pages. We extracted the most relevant related story's *id* and provided it as a hyperlink for this story. To demonstrate the semantic linking possibilities of an ANNODEx-based FÍSCLÁR-NEWS system, we provide hyperlinks to HTML pages from external news sources and this allowed us to exploit the ability of ANNODEx files to link to files of other media types. We used the ANNODEx plugin for the popular browser 'Firefox'. Opening an ANNODExed news story within Firefox the user is presented with the interface shown on the right of Figure 1. Rather than using a separate pop-up window to view the video content, the ANNODExed FÍSCLÁR system enables users to playback video content within the same browser window. Each of the stories from the broadcast are displayed on the left of the interface complete with their keyframe and clip description. While playing back a news story, a context-sensitive hyperlink is displayed underneath. Therefore unlike the current FÍSCLÁR-NEWS system, the user can continue their browsing by following these links out of the news story without having to retreat back to the browsing system. In effect the video content can now contain both the anchors and the targets of hyperlinks.

4 Analysis and Future Plans

The ability to combine continuous media material with its own content description into the one source brings the idea of a true multimedia semantic web one step closer. When consuming an ANNODEx file, the browser becomes a media player, displaying the file's clips, descriptions and hyperlinks but still retaining the traditional properties of a web browser such as a browsing history, and in essence this enables *Media Browsing*. An advantage of this approach over our cur-



Figure 1: The ANNODEx-based Firefox Browser interface

rent FÍSCLÁR system is that ANNODEx file fits neatly into the browsing and surfing paradigm on which the web was founded unlike FÍSCLÁR-NEWS in which the browsing and viewing aspects are separated

In our future work we will investigate providing links at a higher level of granularity than at present, and rather than per story we would like to provide context-sensitive links at specific times in a story and for particular objects within a story. We would also like to investigate applying the Multitrack feature of ANNODEx to our system. This feature allows for multiple links to be associated with a particular clip. We also plan to investigate the notion of users providing their own annotations and links using the Multitrack feature.

5 Conclusions

To date, almost all of the work done on the semantic web is done on artifacts which are either highly structured or text/HTML in nature. While this has allowed work on the development of semantic web infrastructures to proceed, it has not enabled a truly *multimedia* semantic web to be considered. The ANNODEx markup language, allows continuous media content to be combined with content description into one file and this allows direct linking to/from such combined files. This provides the tools needed to build a true multimedia semantic web and in this paper we have shown this to be possible by constructing the ANNODEx-based FÍSCLÁR-NEWS System and extending the functionality of our previous FÍSCLÁR system.

References

- [MPEG7 Committee, 2002] MPEG7 Committee. *MPEG-7: Multimedia Content Description Interface*. ISO, 2002.
- [Pfeiffer *et al.*, 2003] S. Pfeiffer, C. Parker, and A. Pang. *The Annodex annotation and indexing format for time-continuous bitstreams, Version 2.0 (work in progress)*. <http://www.ietf.org/internet-drafts/draft-pfeiffer-annodex-01.txt>, December 2003.
- [W3C, 2001] W3C. *World Wide Web Consortium (W3C). Synchronised Multimedia Integration Language (SMIL 2.0)*. <http://www.w3.org/TR/smil20>, August 2001.
- [Smeaton *et al.*, 2005] A. F. Smeaton, N. O'Connor and H. Lee. FÍSCLÁR-News : Multimedia Access to Broadcast TV News. In *ERCIM News*, No. 62, July 2005.

ANNODEX FÍSCHLÁR-NEWS System Architecture

Our system diagram is shown in Figure 2. The first stage in this work was to transcode the MPEG-1 news data into the Ogg format, an encapsulation format used for ANNODEX bitstreams, which was done using the `ffmpeg2theora` MPEG-Ogg application. Before we could ANNODEX an Ogg file we needed to generate a valid CMML description. The information required for the CMML files was gathered from both local and remote sources. We developed a Java parser to gather this data and create the CMML files.

The parser began by crawling through our MPEG7 database for FÍSCHLÁR-NEWS and extracting the following data related to each story or `<clip>`:

- *Clip id*, The anchor point associated with each clip element. Each story in the FÍSCHLÁR-NEWS system is referenced by a unique `storyId` for example 18237.1 where 18237 refers to the id of the captured news programme and 1 indicates that it is the first story from that day.
- *Clip Start Time*, A time offset indicating the position within the media file where this story occurs.
- *Clip Description*, The description of the news story. This is the subtitle text provided with each broadcast as extracted by the FÍSCHLÁR-NEWS system and stored in the MPEG-7 description.
- *Image*, An image representative of the story. In the FÍSCHLÁR-NEWS MPEG-7 descriptions each story has an associated *keyframe* which is automatically taken from the first anchor-person shot within the story.

The remaining element of the clip, the hyperlink, was generated by locating a related news story within the ANNODEX files we were working with. In order to collect these related stories our Java parser called these XML pages directly and extracted the most relevant related story's *id* and its associated text. We also decided to provide hyperlinks to HTML pages from external news sources. We chose the RTE News site (<http://www.rte.ie/news>) as the destination target for outgoing links as it provides back-dated news stories and is Irish in its news coverage, hence likely to be related to much of the news stories in our system. In the event of not finding a related story within our subsection of the FÍSCHLÁR system the parser will issue a query to the RTE news website using the text of the first line from the subtitles of the TV news story. The parser will then extract the top link from the returned results and provide it as the related story link. This neatly ensures that every story in our ANNODEX-based FÍSCHLÁR-NEWS system has a link to a related story, either another TV news story or a news story in the RTE News website.

Having generated the CMML file with outgoing links from each news story, the `anxenc` encoder then merged this file with its associated Ogg video producing an ANNODEXed news broadcast. This file was finally transferred to an Apache 2 Webserver with an ANNODEX extension module installed. This module allows the Apache server to supply the Annodexed media with timed offsets, preprocesses Annodexed

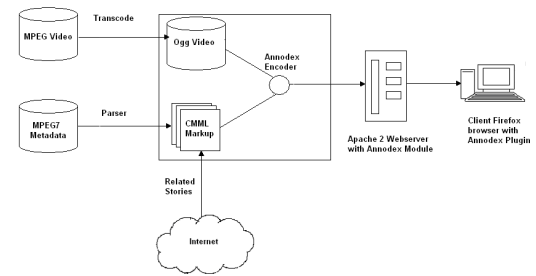


Figure 2: Architectural Diagram of the ANNODEX-based FÍSCHLÁR-NEWS system

media from CMML and enables content descriptions to be extracted from the CMML.

ANNODEX FÍSCHLÁR-NEWS System Demonstration

We will demonstrate the ANNODEX FÍSCHLÁR-NEWS system on a laptop, as illustrated in this short paper, showing browsing and linking between news stories and to outside, non-FÍSCHLÁR resources.

This will be done using a Mozilla Firefox browser with the ANNODEX plugin installed as in Figure 1. Each of the stories from the broadcast are displayed on the left of the interface complete with their keyframe, clip description and date of broadcast. For ease of viewing only the first couple of lines of the description are displayed here, clicking on a particular story will enable a preview of the clip by displaying the entire subtitle text in the description box and displaying its related hyperlink in the Hyperlink box. Double-clicking on any clip will begin video playback at the relevant offset in the file, and while the video of a story is playing, the active clip is highlighted on the lefthandside and the story's content sensitive hyperlink is displayed underneath the video window. In essence the clips can be viewed as state changes in the system where a clip remains active until a new clip is encountered.

Thus we will illustrate the end-user aspects of the Firefox enabled ANNODEX FÍSCHLÁR-NEWS system, integrated video browsing and playback and cases of (hyper-)linking from a TV news story (video) to another TV News story (also video), and from that story to a story on a news website, in text. What this contributes to a conference on the *Semantic Web* is a working illustration of how continuous media, including video and audio, can have their content description (MPEG-7 in our case) embedded within the actual content rather than stored separately. With this we can open up the development of *semantic web* technologies to media other than structured text and this must be of interest to attendees at ISWC.