# WLAP: THE WEB LECTURE ARCHIVE PROJECT

## The Development of a Web-Based Archive of Lectures, Tutorials, Meetings and Events at CERN[1] and at the University of Michigan

NORA BOUSDIRA

*CERN – Ecole Nouvelle d'Ingénieurs en Communication, Lille*
*E-mail: bousdira@elv.enic.fr*

STEVEN GOLDFARB

*University of Michigan, Ann Arbor*
*E-mail:Steven.Goldfarb@cern.ch*

ERIC MYERS

*University of Michigan, Ann Arbor*
*E-mail : myers@umich.edu*

HOMER A. NEAL

*University of Michigan, Ann Arbor*
*E-mail: haneal@umich.edu*

CHARLES SEVERANCE

*University of Michigan, Ann Arbor*
*E-mail: csev@umich.edu*

MICK STORR

*CERN*
*E-mail: Mick.Storr@cern.ch*

GIOSUE VITAGLIONE

*CERN – University of Naples, Italy - University of Michigan, Ann Arbor*
*E-mail: gio@qos.net*

## Abstract

This paper summarizes the results of a project to develop an electronic repository of "content-rich" lectures, talks, and training activities on the World-Wide Web. The work was carried out from July 1999 to July 2001 by a collaboration consisting of the University of Michigan ATLAS Collaboratory Project, the University of Michigan Media Union, the CERN HR Division, supported by the CERN IT and ETT Divisions and the CERN Academic and Summer Student Programs. In this document, we describe the software application chosen to synchronize the slide presentations to the video recordings, provide technical solutions to the various recording and archival challenges encountered during the project, and propose a set of research and development issues we feel merit further investigation. We also present the concept of a "Lecture Object" and suggest the adoption of standards so that lectures at multiple institutes can be seamlessly shared and incorporated into federated databases world-wide.

---

[1] CERN - the European Organization for Nuclear Research, commonly known as the European Laboratory for Particle Physics.

# Contents

**Trademark Notice**

All trademarks appearing in this document are acknowledged as such.

# 1   Introduction

The primary motivation for the creation of the World-Wide Web was the facilitation of collaboration between scientists [1] . There was a need for a better way for scientists to rapidly exchange large amounts of information, ranging from experimental data and results of analyses to organizational and strategic details related to ongoing experiments. The rapid proliferation of the web and web-related applications, as well as the ever-increasing size and international scope of scientific collaborations, has by now clearly demonstrated the value of the web as a common and necessary tool for research. In addition, it has enhanced the dissemination of scientific knowledge to the general public through the publication of online documents and other web-based media.

This document reports on an effort to explore the usage of the web for archiving of "content-rich" material, which we define to be lectures, seminars, or other events which include audio, video and visual support materials. The work targets a segment of the tasks that must be completed for scientists and others to optimally draw upon the web for transmitting information for training and archival purposes, as well as for keeping colleagues informed of strategic, technical and administrative decisions.

CERN was chosen as a focal point for this research because of its historical participation in web development, its continuing role as a center for scientific research and information exchange, its rich education and training programs, and the new challenges it faces during the current construction and future running of the next generation of experiments for the Large Hadron Collider (LHC). These experiments will be run by teams of sizes heretofore unseen in most sectors of the scientific community, with thousands of members literally spread around the globe.

The involvement of the University of Michigan in the CERN ATLAS experiment, as one of that experiment's largest groups, is one of the reasons for its interest in this project. Augmenting this reason are the roles played by the University of Michigan in the inauguration of U.S. participation in the CERN Summer Student program, along with its affiliation with Internet2, and its work in bringing CERN into Internet2. These reasons, together with the presence of pioneers at the University in the development of multi-media educational tools, all provided a shared rationale and stimulus for the University of Michigan and CERN to examine the possible future role of web-based archiving in the general area of highly collaborative large-scale research involving universities and international laboratories.

With this background, a major focus of our efforts has been to investigate how to best facilitate the work of large, globally dispersed scientific collaborations. Another has been to study how to best reinforce CERN's education and training programs and make them accessible to as wide a community as possible.

This paper seeks to examine the relevance of web-based archiving to this set of challenges. We approach the topic by examining how we have used web-based archiving to record a series of content-rich presentations at CERN over the past two years.  The issues covered range from the technical details of how such recordings are made, to questions of how the technology can be improved, and how such material could be confederated to address certain larger goals.

## 2   Project Motivation

In this section, we describe the principal motivations for our present study. Though we cite specific applications, the results presented herein have a clear relevance for a variety of scientific fields and educational venues.

### 2.1   Communication in modern high-energy physics experiments

A prime motivation of the WLAP project was the hope that web-based archiving technology could address some of the key challenges that face the high-energy physics community. To understand these challenges, it is instructive to consider the anatomy of a modern high-energy physics experiment. Once a set of physics goals has been established for an experiment, the achievement of these goals requires the massive generation and refinement of novel ideas for how to solve the myriad of attendant technical problems. A talented set of individuals must be assembled to assimilate these ideas, to design and build the detector components, and to integrate the components into the overall detector, a process that can span a decade and involve extensive communication among thousands of experts in numerous countries. The running detector must be maintained and monitored and the resulting data must be analyzed, an activity that may involve yet another decade. The required funding must be repeatedly applied for, dispersed, expended, accounted for and reported on. At every stage extensive communication is required among dispersed participants. In such scientific enterprises the communications aspect can rapidly become as daunting as the scientific and technical challenges themselves.

LHC experimental collaborations will have hundreds of participating university groups. Depending upon the precise responsibilities accepted by a group, it may have the need to interact daily with colleagues at a half-dozen other universities. Many practical questions quickly arise. For example, how does one convene frequent meetings with colleagues in real-time on modest budgets, when time differences of as much as 12 hours are involved for the participants, many of whom have major responsibilities in addition to those that form the focus of the meetings? How does one offer a tutorial to two thousand colleagues on some paradigm he or she has developed, when initially there are only one or two true "expert peers" on a given topic in the entire collaboration and it is your job to train all of the others?

Since many of the large experiments may run for as long as twenty years, involving numerous generations of Ph.D. students, how is information recorded and passed on to subsequent generations? When major talks are given on results from a running experiment by the author of a particular analysis, how does that information get captured and made available to members of the collaboration who may only be able to access the talk hours (or years) later? How do findings and major strides emerging from these experiments get captured and interwoven into classroom materials for later presentation?

These are but a few of the questions that arise in the conduct of large high-energy physics experiments. Given the nature of the World-Wide Web and the original function seen for it, one would naturally be led to inquire if the Web itself might provide possible solutions to facilitate the communication requirements of the very large experiments it had helped make scientifically possible.

### 2.2   Enhancing learning capability and dissemination of education and training

There is another motivation for our work in the area of web lecture archiving: its potential use in education and training. Traditional lectures and seminars follow a

sequential pattern in which the lecturer prepares a presentation and delivers it, often accompanied by visual support material. The delivery mechanism can vary in style, with the lecturer using different techniques for displaying the visual support material, for example, an overhead projector, a computer slide projection or a blackboard. Questions may be taken during the presentation, at the end, or not at all. In each case, students must rely on their notes and/or a copy of the support material to recall the key points of the lecture at a later date.

People who are unable to attend or miss a session have to make do with a copy of the visual support material when it is available and even in the best of cases find it very difficult to reconstruct in detail what has been presented verbally.

Having access to some form of audio/video reproduction of the original lecture, however, can greatly facilitate the learning process and allow many more people, in addition to those who physically attend, to benefit. Such a reproduction can exist in a variety of media, including audio or video recordings. Unfortunately, the dissemination of the material on audio/video tapes is cumbersome, thereby limiting access.

Again, recent technological developments based on the accessibility of the Internet and the widespread utilization of the World-Wide Web lead us to conclude that these difficulties can now be overcome.


## 3    Project Implementation

### 3.1    The pilot project

The Web Lecture Archiving Project (WLAP) activity started in 1999 as a pilot-project [2] funded by the U.S. National Science Foundation and the University of Michigan (UM). The primary aim was to examine the feasibility of using a software tool, called Sync-O-Matic [3] , to record and archive slide-based lectures in a variety of situations. Following the success of the pilot project [4] , a collaboration was formed between the CERN HR Division Training and Development group [5] , the UM ATLAS Group [6] and the UM Media Union [7] , supported by CERN IT Division. The objective was to demonstrate the feasibility and usefulness of archiving lectures, seminars, tutorials, training sessions and plenary sessions of ATLAS experiment meetings by focusing first on the archiving of the prestigious CERN Summer Student Lectures.

### 3.2    The archive application

The Sync-O-Matic application, successfully tested at CERN during the pilot project described above, was adopted for the implementation of the joint project. Sync-O-Matic was written by Charles Severance, then Associate Director of the University of Michigan Media Union. Documentation describing the software can be found on its web site [3] . It is freely available and there is a mailing list of users and developers who can be contacted for help and support.

Sync-O-Matic produces slide-based web-lectures for viewing with a standard web browser and the freely available RealPlayer plug-in. Its output is a multi-media lecture that combines the audio and video playback of the lecturer with digital images of the visual support material, synchronized to the video, and displayed in a browser window. **Figure 1** illustrates an archived lecture. Note that the video and slide indexes can be used

to rapidly locate sections of the lecture or to review the slides in order to select specific topics.
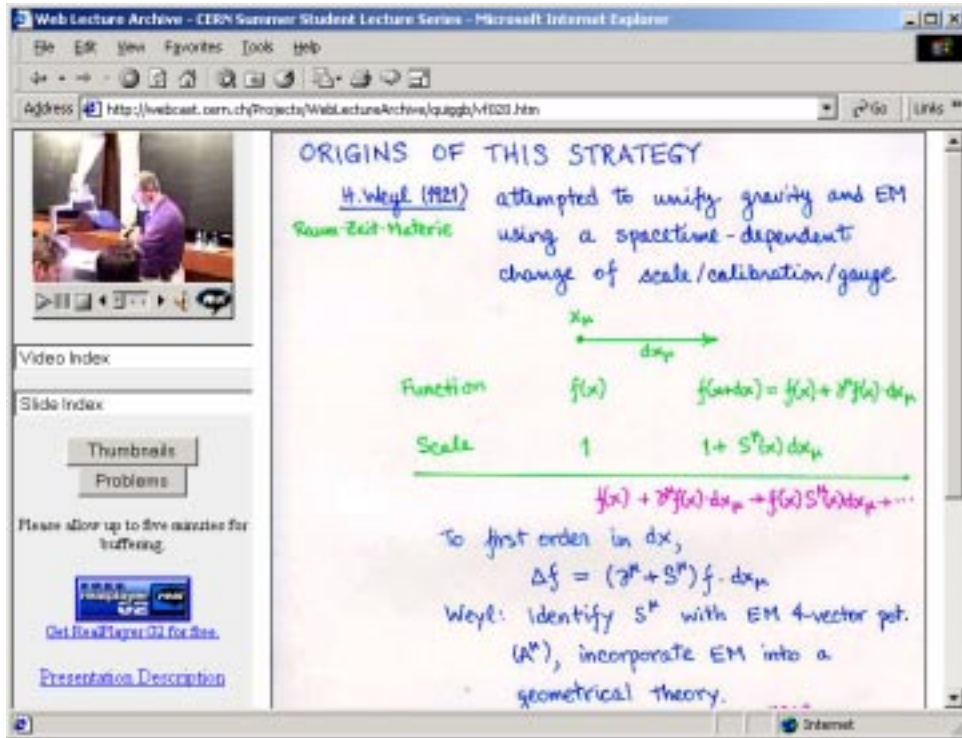


**Figure 1**: A typical archived lecture as viewed from a web browser. The video image of the speaker appears in the upper left-hand corner of the page. The visual support material (in this case, scanned transparencies) appears in the large window on the right. The changing of the transparencies is synchronized to the timing of the video.

These important features distinguish Sync-O-Matic from the historically common approach to lecture archiving based on a video recording of the event which combines views of the speaker and visual support material. Such a video is necessarily a compromise, as a choice has to be made between focussing on the lecturer or support material, or more precisely part of this material to facilitate readability. The camera operator therefore makes the choice for the viewer as to what is of primary interest at any time. This is in marked contrast to a Sync-O-Matic lecture in which the speaker and support material are always in view.

Another drawback of a standard video is that although good video resolution can indeed reproduce the slides in a readable format, it does so only at a significant cost of network bandwidth and archive size. For example, we find that using MPEG-1, a bandwidth of 1500 Kb/s is necessary to ensure readability using historical approaches, to be compared to a typical Sync-O-Matic archive that can be readable at 50 Kb/s.

Regardless of advances to the technology, there will always be some inefficiency introduced by the transmission of a video stream rather than a fixed image. In addition, the video stream lacks the slide preview and rapid search/location functionality provided by the indexed Sync-O-Matic archive. As we will discuss below, such indexing could also be exploited for the development of web-based lecture databases and search engines.

### 3.3    The archive process

Sync-O-Matic was originally designed for use by an individual teacher, operating alone or in a staffed distance-learning studio, using Microsoft PowerPoint materials [12] . In this mode, Sync-O-Matic imports a PowerPoint file and converts the slides to GIF/JPEG images as shown in

**Figure 2**.

While the teacher gives the lecture, Sync-O-Matic records the audio and video from a microphone and camera using the RealProducer [8] ActiveX control. As the lecturer changes slides, Sync-O-Matic records those actions and the timing of each action in internal text files.
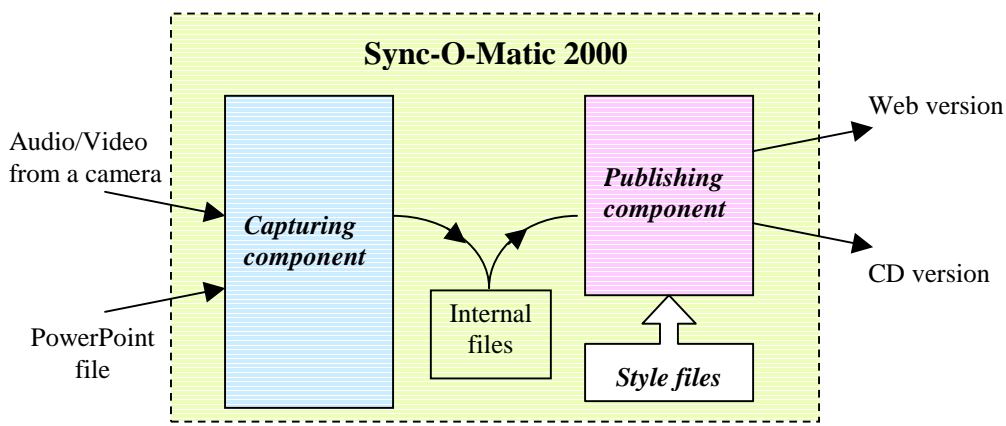


**Figure 2**: Sync-O-Matic standard operational procedure. The speaker uses Syncomatic on a PC connected to a camera and a microphone. The speaker loads the PowerPoint file into Syncomat, and then starts recording the audio and video, and changes the slides as he/she talks. The capturing component of Syncomatic creates internal files with the audio, video, slides, and timing information. Once the recording is over, the speaker can publish the web lecture in a format suitable for a CD-ROM or for the Web. The "Style files" can be edited to change the structure and the "look and feel" of the web lectures.

At the end of the presentation, two archived lectures are produced, one suitable for viewing with a standard web browser from CD-ROM, and another suitable for viewing directly from a web server. The look and feel of the resulting lectures is controlled using Sync-O-Matic specific "style files". These style files are written in HTML with some Sync-O-Matic specific markup included. The result of the publishing process is a directory with several HTML files, text files used to allow random navigation of the lecture, and the media files.

The media files make up the bulk of the disk usage. An average quality video and high quality audio (160x120 pixels, 24kps video + 16kbps audio) lecture requires about 36 Mbytes per hour of lecture. This relatively small amount of disk storage allows 15-20 hours of lectures to be stored on a single CD, and a large number on a web server without resorting to exotic storage technology. To watch a web lecture the user can use any popular web browser.  The first time the user views a lecture, it may be necessary to download and install the RealPlayer software (the free version is sufficient), although this software is now commonly bundled with most web browsers.  JavaScript is not required but some advanced navigation features are available if it is enabled.

Although adequate to meet the requirements of a lecturer in a teaching environment as described above, one of the major challenges of the WLAP project was to extend the functionality of Sync-O-Matic beyond its original design in order to handle the demands of live lectures and to cope with a number of challenging recording scenarios, such as hand-written transparencies and blackboard material. The various scenarios encountered and the solutions developed are detailed in section 4 below.

### 3.4    The WLAP archive

During the pilot project, a significant portion of the 1999 CERN Summer Student Lecture Series was recorded and made available to the participants and to the general public. As a result, significant demand was generated at CERN and throughout the physics community to continue the recordings on a regular basis.

The CERN HR Division Technical Training group, supported by the CERN Amphitheatre technical support team, took up the challenge of demonstrating the feasibility of recording a wide variety of lectures, while simultaneously investigating and applying technical advancements to simplify the process and to reduce the manpower needs. Over the next 15 months many important CERN colloquia, Technical Training seminars, Academic and Summer Student Program lectures, and software training tutorials were recorded, either by request, or for the purposes of testing the technology under a variety of recording conditions.

So successful was this effort that since January 2001 this activity has been taken over by a team from the CERN ETT Division who have not only recorded all Academic and Summer Student Program lectures since that date, but have also developed and improved the operational procedure. Several steps of the file manipulation process have been automated and the lecture 'events' are now integrated into the ETT developed CERN event calendar system.

The ATLAS Collaboration served as a test-bed for much of the material recorded, placing the focus on the challenges of a large-scale, globally dispersed scientific collaboration. The events archived for ATLAS include collaboration meetings, plenary sessions, subsystem workshops and tutorials. In addition, the collaboration profited from its members accessing and providing feedback on all of the archived lectures. The feedback was, in general, very positive, and often resulted in requests for a greatly extended service. Nearly all suggestions regarding technical improvements were eventually implemented to the recording and archiving procedures.

#### 3.4.1    The CERN WLAP archive

One of the reasons mentioned for choosing CERN as the target of the project is the richness of its physics program. Indeed, the project team quickly found the number and frequency of interesting recording opportunities to exceed its ability to record, regardless of the facility of the process. Given the modest resources, however, a significant number of lectures were archived and published on the WLAP site, covering a large spectrum of laboratory activities.

The current archive, which is growing literally every day, comprises more than 400 lectures. Among the more notable archives are colloquia by Prof. Martinus Veltman on the history of the Standard Model [9] , presented in the CERN auditorium shortly following his reception of the 1999 Nobel Prize in Physics, and by Dr. Paul Kunz on the

birth of the World-Wide Web, its early stages of development at CERN, and his involvement setting up the first web server in America [10] .

In **Table 1**, we briefly summarize the current content of the WLAP archive at CERN[2]. The full catalogue can be viewed at the WLAP archive web site - http://www.cern.ch/WLAP/.

**Table 1**: Summary of the current contents of the WLAP lecture archive

| Category | Number of lectures |
|---|---|
| Academic Training Program Lectures 99/00, 00/01 | 109 |
| ATLAS Plenary Sessions Meetings | 25 |
| ATLAS & LHC Software seminars | 65 |
| General Colloquia & Seminars | 5 |
| Summer Student Lectures 1999, 2000 & 2001 | 36+95+66 |
| Technical Training & Safety Seminars | 7 |

3.4.2   The ATLAS GEANT4 workshop

In addition to the events recorded at CERN, the University of Michigan members of the team archived a series of lectures given by Andrea dell'Acqua at an ATLAS-sponsored workshop [11] on GEANT4 held on the Ann Arbor campus.

GEANT4, a recently completed re-write in C++ of the well-tested GEANT3 application, is a software package for simulating the passage of particles through material. It is currently being tested for use by the LHC experiments.

Usage of the software requires a significant effort to bring the physics community, typically well versed in the usage of FORTRAN, up to speed in both the syntax of C++ and the concepts of object-oriented analysis and design.

Andrea dell'Acqua is one of the key developers of the ATLAS simulation software. It was clearly advantageous at this time for him to make the effort to train new users and developers in its usage. In the future, new collaboration members, or those just beginning to contribute to the software after having completed contributions to other aspects of the detector construction, will be able to access the archived lectures at a web-based training site [11] which will include documentation, problems and solutions, and a listing of frequently asked questions with answers.


## 4   Details of the Implementation

Contrary to the original design concept of Sync-O-Matic, the bulk of the lectures recorded for the WLAP archive were not prepared in advance using Microsoft PowerPoint, nor were they presented in a controlled environment with the main focus being the production of a quality lecture archive. Rather, the lectures were presented to a live audience using a variety of visual media and the lecturer was insulated as much as possible from the recording process. Given these constraints, it was necessary for the archive team to develop a number of new operational procedures for the production of a quality archive, using a reasonable amount of resources and manpower. In this section, we present a summary of these procedures.

---

[2] For completeness the table includes lectures from the Academic Training Program 2000/01 and Summer Student Program 2001 recorded by ETT Division since January 2001 which are not strictly part of the WLAP archive.

### 4.1    Audio and video capture

Because high quality audio and video are essential for producing good streaming media, we separately taped the lectures with a high quality wireless microphone and allowed the camera operator to concentrate on the video. Initially the video was encoded from tape but it was soon realized that live encoding, with tape as backup, was not only feasible but saved a significant amount of archive production time.

Our target audience was identified from the outset as being world wide, with access to the Internet provided via research network backbones, as well as using the most currently standard analog modems from home. Based on this, we chose to deliver a total bandwidth of 40 Kb/s, divided into 24 Kb/s for video and 16 Kb/s for audio streaming, in order to ensure high-quality audio. In order to optimize usage of the client bandwidth, we encoded using the RealProducer Surestream technology, which delivers reduced quality video if necessary to maintain sound quality.

One slightly negative impact of using Surestream for encoding is that the media files becomes somewhat larger (approximately 40 MB per lecture hour), decreasing the average number of lectures that can fit on a CD-ROM to about 15. This is not deemed to be a crucial problem, however, as the vast majority of access to the lectures is direct, via the web-server. Rather than handing out a CD-ROM to each student at the end of the summer, as was the practice for the pilot project, CD-ROM's of specific lectures were provided to any user, upon request.

### 4.2    Scenarios for handling the visual support material

Presenters differ in the way that they choose to display their visual support material and the techniques encountered ranged from electronic presentations with PowerPoint, through transparencies and overhead projector, to chalk and a blackboard. For each scenario the challenge was to record the order in which the information was presented and the time of display in order to synchronize with the audio and video. Furthermore, we chose not to burden the speaker with the technical aspects of starting Sync-O-Matic at the right moment and learning to use its interface in order to give their lecture.

For the pilot project, the camera operator recorded the timing of the transparency changes during the lecture in a notebook. During the production of the Sync-O-Matic archive following the event, this information was entered into the formatted text files internal to the software. For presentations given on other media than PowerPoint, the slide images were converted to GIF format (scanning overhead transparencies, if necessary) and then placed in the appropriate subdirectories for production of the archive. This process was relatively time-consuming, requiring between two and five hours of post-production work for each hour of presentation.

During the Summer Student Lecture Series of 2000, efforts were focused on reducing the time and manpower required to produce the archived lecture. The simple text format of the Sync-O-Matic internal files made it possible for us to create several software macros, loaded as background processes on the presentation PC, to capture the timing of the slide changes, during the presentation.

In the simplest case, shown in **Figure 3**, when the speaker used Microsoft PowerPoint for the presentation, we developed a VBA macro called *CarpePpt [13]*  to capture the timing when the speaker changed the slide in the PowerPoint viewer. The macro is started well before the presentation and is completely transparent to the speaker who uses PowerPoint as usual to show the slides. At the end of the presentation, a file is generated

containing the timing information in the Sync-O-Matic format.  This file is then used by Sync-O-Matic to publish the resulting lecture.
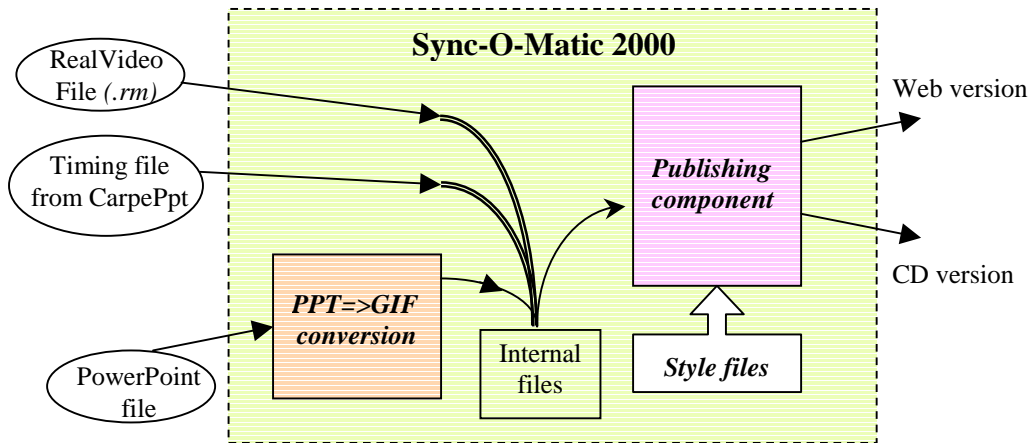


**Figure 3**: Operational procedure for the PowerPoint scenario. The speaker is not involved in the recording operations. The audio/video signal coming from a camera is encoded by a computer using RealProducer, that generates a RealVideo file. CarpePpt captures the timing transparently for the speaker who uses PowerPoint normally to show the slides, and generates a timing file in the Syncomatic internal format. In post-production, the video file and the timing file are copied and named as if they were produced by Syncomatic. The publishing component generates, in the same way as in Figure 1, the Web lecture in the formats suitable for the Web and for the CD.

When speakers used Postscript [14] or Adobe Acrobat [15]  for their prepared slides we developed another tool called *CarpePdf* [16] to capture the timing information during the lecture. This method was also used when the speaker was able to make the transparencies available in advance so that they could be scanned and converted into PDF for the presentation. In this case, the scanned images were also used to generate GIF files for the Sync-O-Matic archive.

The real challenge is when the transparencies are not made available in advance or when the presenter writes on a blackboard. In these cases, we used a second video camera to record images from the display screen or the blackboard and we entered the timing values by hand during the post-talk production using a Sync-O-Matic feature called TimeIT. The images were eventually replaced with higher quality scanned images of the transparencies or still pictures of the blackboard, when they were available after the presentation.

In 2000, Charles Severance developed a new product called ClipBoard-2000 [17] . ClipBoard produces files that are compatible with Sync-O-Matic's publishing process and introduces many new features, including two-camera support. For the challenging case discussed above, one camera could now be used to capture the video of the speaker with the second camera dedicated to the overhead screen or blackboard. A technician then pressed a button on the PC to capture a high quality image each time the speaker changed the slide, showed an object, or drew on the blackboard.

The introduction of ClipBoard not only aided in automating the recording process, but also in the post-talk production, when the high quality scanned images were substituted for the preliminary screen snapshots. Because many speakers give their presentation in a non-sequential fashion, skipping from slide to slide or moving from the

slide to the blackboard and back, the image substitution can be slightly complicated. To aid in this process, Giosue Vitaglione developed a tool called Snaps-O-Matic [18] . Snaps-O-Matic reads in the Sync-O-Matic internal files, the scanned transparencies, and the snapshot images and presents them to the publisher in a graphical interface. From the interface, the publisher can drag and drop the scanned transparencies to the snapshots. Upon completion, Snaps-O-Matic writes a Sync-O-Matic intermediate file that describes the new lecture. The whole process for this scenario is shown in detail in Figure 4 and Table 2 summarizes the complete set of scenarios.

While the tools described above do significantly aid in the recording and production processes, it remains the case that lectures presented on non-digital media require some amount of post-talk production work. Typically, this can be estimated as 1 FTE-hour for each hour of lecture. One could imagine using new devices, such as automatic scanning projectors and timing mechanisms, but they do not (yet) exist in the CERN Auditorium.

For important events, such as the presentation of a Nobel Laureate at CERN, we believe the value of the final product certainly justifies the effort. For lower profile events, conference conveners may consider requesting either digital presentations or the submission of the non-digital media before the event, with adequate time for scanning and preparation. This will guarantee immediate publication of the material, following the presentations. This is a significant asset for audiences who would like to participate in or follow an event, but who are either unable to attend locally or via videoconference, or who are separated from the event location by several time-zones.

| Scenario | Display Technique | Timing Capture |
|---|---|---|
| PowerPoint | PowerPoint | CarpePpt |
| Postscript or Adobe Acrobat | CarpePdf | CarpePdf |
| Transparencies received before lecture | Scan and display using CarpePdf | CarpePdf |
| Transparencies not available beforehand | Overhead Projector | Clipboard & Snaps-O-Matic |
| Blackboard | Blackboard | Clipboard & Snaps-O-Matic |

**Table 2**. Summary of the various scenarios encountered for visual support material and the techniques employed to record the data
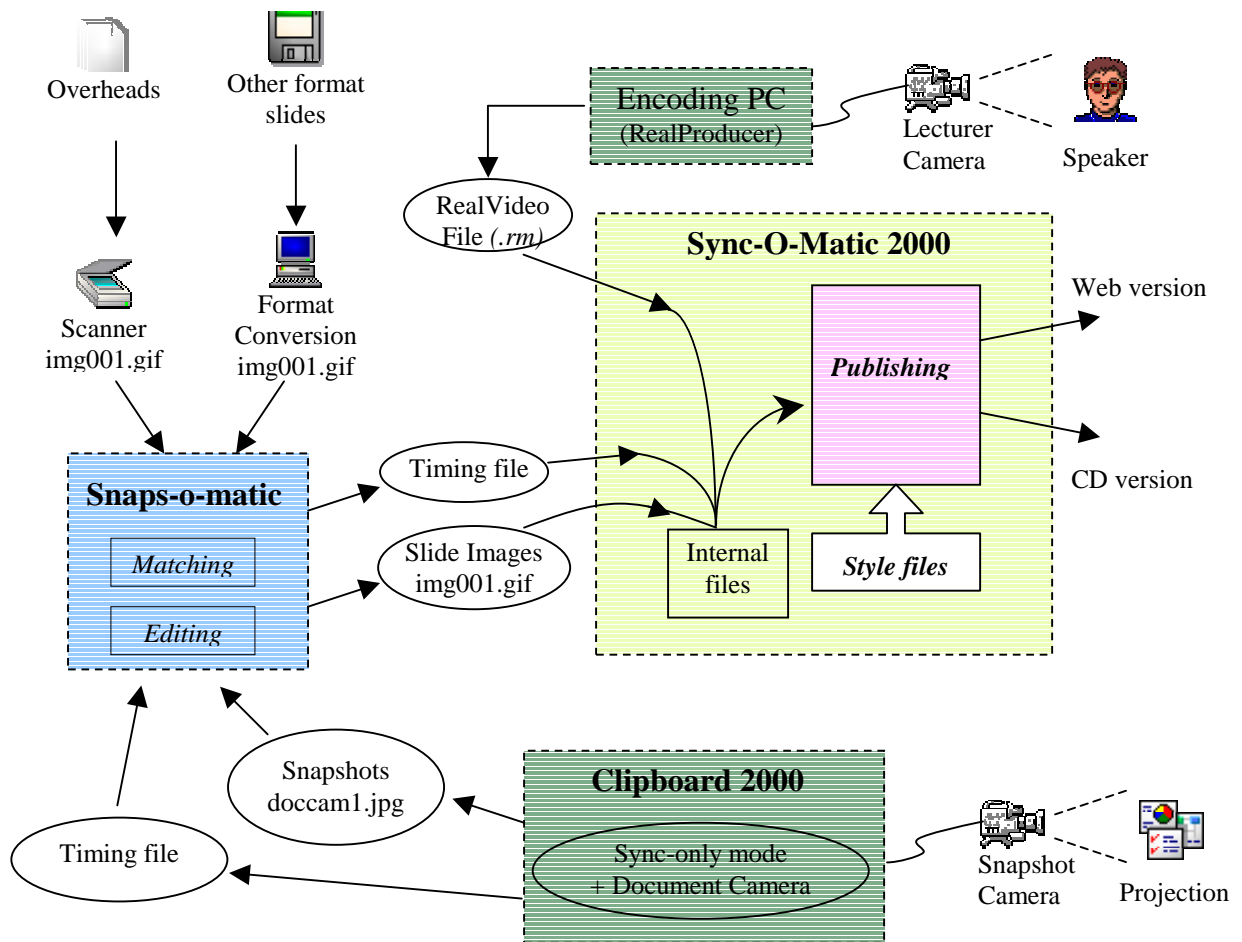
**Figure 4**: A description of the procedure when slides not available in advance: The lecturer camera (top, right, in the figure) records the face of the speaker and the audio, and a RealVideo file is created. The snapshot camera (bottom, right) takes pictures of the projected transparencies, and the cameraman, using Clipboard2000, takes a snapshot each time there is a new transparency. Clipboard creates a timing file and a set of JPEG images that go to Snaps-o-matic. In post-production either the transparencies are scanned or an electronic version of the slides is converted. In either approach, a set of files called, img001.gif, ..., is created. A technician, using Snaps-o-matic, matches the high quality GIF files with the snapshots and substitutes them editing the timing if needed. Finally the video file, the slides and the timing are processed by Syncomatic to produce the Web lectures for CD and the Web.

## 5    The Lecture Object

As we built up a significant number of lectures in the WLAP archive and people started to access it on a regular basis, we realized that constraints imposed by the technology we were using would unnecessarily limit its lifetime.

Ideally we would like the archive to be persistent over a number of cycles in technological advances such that it would remain viewable over many years. This led us to the concept of the Lecture Object which we believe not only solves the problem of persistency but is a useful new paradigm for archiving and sharing web lectures on a world-wide basis.

In this section, we briefly describe what the Lecture Object is, and our principal motivations for proposing it as a standard for archiving web lectures (section 5.5). More information can be found in [19] [20] .

Section 5.4 includes a brief description of some prototype software developed compliant to the Lecture Object architecture.

### 5.1    Lecture Object Architecture

The Lecture Object architecture is based on the distinction of three tasks: "Capturing, Archiving and Delivering", as shown in Figure 5.

The Lecture Object is a *standard* format for *archiving* web lectures and for *exchanging* them between different archives.

The typical life-cycle for a Lecture Object is: (1) A lecture is recorded, and a capturing tool creates a Lecture Object, (2) It is archived in a Lecture Object repository, (3) A publishing tool converts it in a presentation format for delivery to the client (e.g. a Web browser) for visualization.
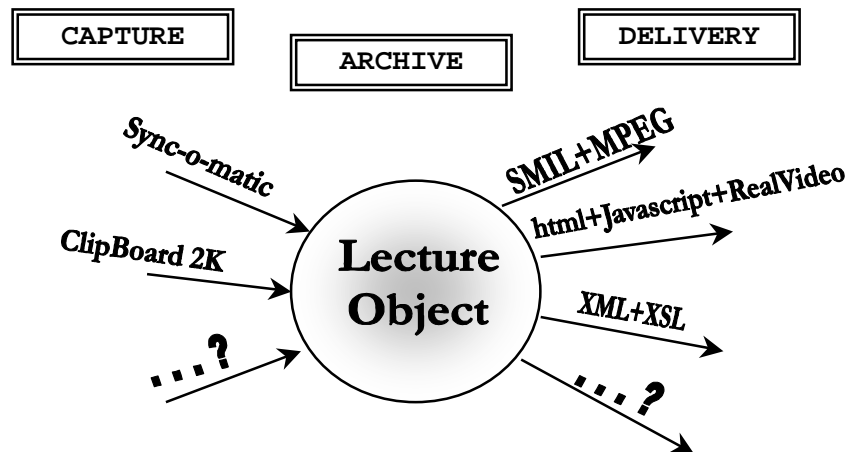


**Figure 5**: The Lecture Object architecture is based on three steps: Capturing, Archiving and Delivering. Different capturing tools, optimized to work in a certain scenario, produce standard Lecture Objects. The lecture objects are archived in a repository and are available to be used by software agents, for example a transformation to a presentation format, indexing software, etc…. The adoption of a standard format allows interoperability.

Different capturing tools, used in different scenarios, produce standard Lecture Objects; different publishing tools can transform them into several presentation formats.

The distinction between the archiving format and the presentation format is crucial. The Lecture Object is the archiving format and it is seen as a long-life entity that can be easily converted into other formats, even formats which do not yet exist.

The archiving format adopted has to be independent from proprietary technologies. This is because multimedia companies can change their strategy, limiting the wide use of a technology (e.g. limited platform support, or player software no longer freely available), or imposing constraints on the clients (e.g. advertisements) or compromising the long-term availability of the archived material (e.g. preventing conversion to other formats).

We think that the possibility to change the presentation format, independently from the archiving format, puts users in charge of their information, without tying in the archived content to a specific multimedia company.

Furthermore, starting from a Lecture Object, different presentation formats can be created to fully satisfy the expectations of the clients, always using the best available technologies. When a newer presentation format become available, a new transformation program can be developed to process all the Lecture Objects in the archive and publish them in the new presentation format.

A further advantage of the three-step architecture is that monolithic software applications are not flexible enough to be both efficient in all capturing scenarios and to allow publishing in all the desired presentation formats. The new three-step approach solves this limitation, allowing the use of interchangeable tools.

A Lecture Object must contain all the data (audio, video, slides, timing, order, annotation, etc…) and meta-data (title, author, keywords, description, prerequisites, etc…) describing the lecture. This information is related to the real lecture and not to the web lecture that the clients will watch. It is also independent of how it has been created, and is independent of how it will be used.

A Lecture Object stored in a repository can be transformed by software that runs on the server or on the client. These transformations can be performed at request-time (dynamic transformations) or in batch in anticipation of client requests.

### 5.2    Draft Specification

Starting from the idea discussed above [20] , we have developed a draft of the Lecture Object specifications [21] . This is a first attempt at defining the basis for a standard; more work is needed to refine the design and to review the technical decisions taken.

Figure 6 shows the structure of a Lecture Object with the data on the left side, and the meta-data on the right. The entity in the middle, called "Lecture", is the description of what happened during the lecture. It contains information such as: "The speaker started talking, then showed a slide, after a minute another slide, …".

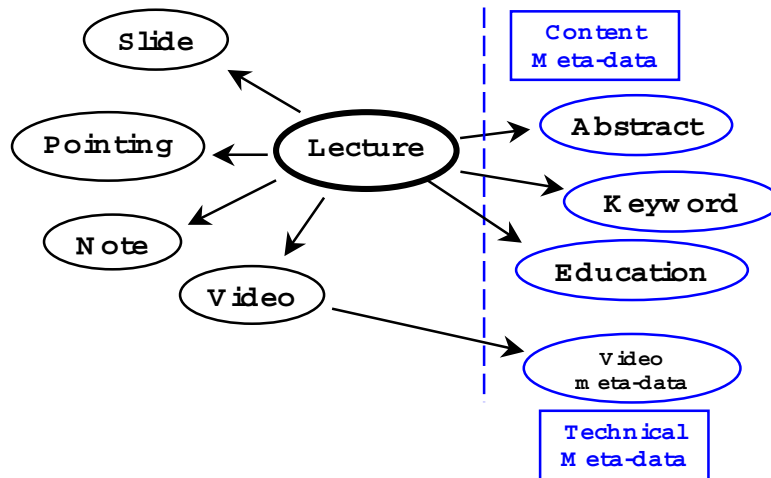We propose to use XML (eXtensible Markup Language) [22]  for this component. Here is a simple example:

**Figure 6**: Components of a Lecture Object: data (on the left) and meta-data (on the right). The content meta-data contain information about the lecture. Other technical meta-data can give information relating to a component of the Lecture Object.

```xml
<?xml version="1.0" encoding="UTF-8"?>
<LECTURE>
<PAR>
  <VIDEO TITLE="Welcome to WLAP" REGION="speaker-face"
         SRC="rtsp://webcast.cern.ch:5540/giosue.rm" />
  <SEQ title="Sequence of slides" region="slide">
   <SLIDE TITLE="First" TYPE="image/gif" BEGIN="00:00:00"
      REGION="slide" SRC="http://webcast.cern.ch/img001.gif" />
   <SLIDE TITLE="Second" TYPE="image/jpeg" BEGIN="00:00:12"
      REGION="slide" SRC="http://webcast.cern.ch/img002.jpg" />
  </SEQ>
</PAR>
</LECTURE>
```

This is a description about what happened during the lecture. All the elements between <PAR> and </PAR> happened at the same time. The elements between <SEQ> and </SEQ> happened in sequence. Thus in this case, there was a speaker talking and at the same time a sequence of slides. This is very similar to SMIL [23] (Synchronization Markup Integration Language), a W3C recommendation. Anyone already familiar with SMIL will easily understand this description of the lecture structure.

The other data components have to be in a widely used format that can be easily converted into others. Good candidates for the video are Mpeg2 or AVI (using a freely available, multi-platform codec, like Indeo3.2 or MJpeg). For slides, we typically used: GIF, JPEG, PNG, HTML or SVG. More formats can be kept in the archive in order to facilitate an easy lossless conversion to other formats.

For the content meta-data, our focus has been on IEEE LOM (Learning Object Metadata) [24] taking into account the activity of other initiatives and working groups, like: Dublin Core [25] , IMS Global Learning Consortium [26] , Ariadne [27] , Gestalt [28] and CEN ISSS LT [29] . We promote the usage of a subset of the IEEE LOM metadata set, with a complete mapping to Dublin Core.

In addition, starting from an idea of the Gestalt work, we propose to archive meta-data about the network resources needed to use a certain component of a Lecture Object. Our proposal is based on expressing the network resources in term of QoS [30] (Quality of Service) parameters using a model of sequences of Token Buckets to characterize a data flow. This allows the mapping of these meta-data on different networking architectures for using guaranteed services and advanced reservation.

## 5.3 Distributed Architecture

Multiple archives can share meta-data about the weblectures they contain, as shown in Figure 7. Software agents called "Brokers" expose all or only a part of their lecture meta-data to the other Brokers. The model proposed is similar to the approach used in the Open Archives Initiative [31] , where "Service Providers" present the services available to the clients, and "Data Providers" physically host the data.

Some development is needed to define the exchange protocols and all the data structures involved. In particular, the meta-data harvesting protocol and the data structures for describing the status of the servers. This will allow the brokers to assign dynamically a certain resource (e.g. video server) to a specific client, adopting special policy of server/network load balancing, based on performance, costs and availability.

The organizations (Universities, Research lab, Companies, etc…) taking part in this "global archive" of web lectures could define agreements for content exchange and mirroring services.
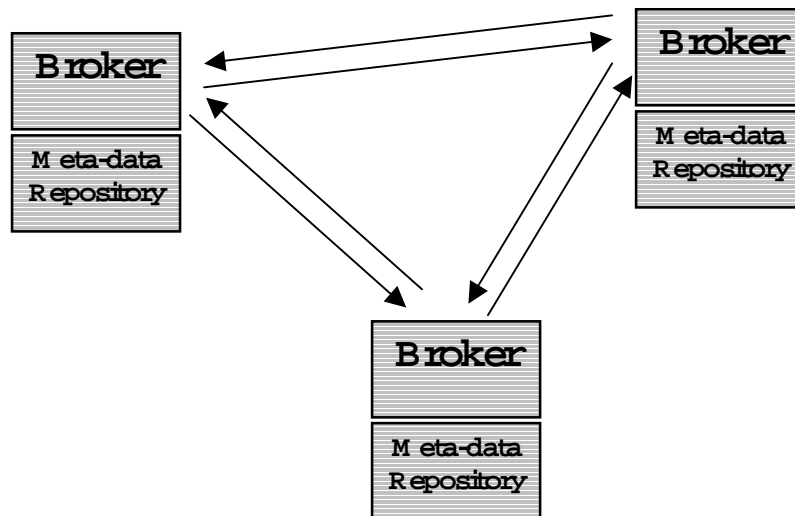


**Figure 7**: Multiple archives exchange meta-data of the Weblectures they host.

## 5.4 Prototype developed

As proof-of-concept for the Lecture Object architecture, we developed some software that we successfully integrated into our production process.

We developed capturing tools for recording the timing when the speaker uses a PowerPoint presentation (CarpePpt v.2.0 [13] ) or Postscript and Acrobat files (CarpePDF v.2.0 [16] ). Both of these create and upload a Lecture Object on the repository.

An updated version of Snaps-o-matic [18] is compliant to the Lecture Object architecture. It is used in post-production in case the speaker uses plastic transparencies. This program can also be used as a simple Lecture Object editor.

A simple server-side dynamic transformation, written in PHP [32] , has been written to transform a Lecture Object in a standard SMIL presentation. And an enhanced version of the same software produces SMIL plus RealPix and RealText, technologies developed by RealNetworks [33] .

We also started using Websentation [34]  (Web-Presentation), an automatic Java-based publishing program, to transform Lecture Objects in presentation formats on a Linux server. Although this software is still in development, and needs to be enriched with more functionality, it has proved the validity of the Lecture Object architecture for reducing the workload in producing web lectures in different presentation formats.

## 5.5   *Advantages of standardization*

We think that the broad usage of a common format for archiving and exchanging web lectures would have a very positive impact on their diffusion as a tool for distance learning and collaborative working.

In addition to the arguments mentioned in the previous sections, we would like to emphasize the following:

- Persistent Archives.  The adoption of a widely used platform-independent standard with no proprietary technology, would allow the lectures to survive external technological cycles.
- Uniform indexing and retrieving.  A common meta-data set would greatly facilitate global indexing. Search interfaces would be more uniform, and the users would be able to perform more advanced searches.
- Interoperable software tools.  Different tools for capturing, publishing, indexing, editing, etc., would be able to work together. Archivists working in particular environments could use the best tools, in their scenarios, for each task in the weblecture production. This would reduce the production workload.
- Sharing content between archives.  Having a common standard would facilitate different organizations setting up agreements for exchanging contents.
- Convergence tools development.  A common format would focus the efforts in software development. This would improve the software quality and facilitate the availability of applications for specific tasks.

We will be encouraging various relevant organizations to provide support for the adoption of standards embodying the general features described above.

## 6  Other Web Lecture Archives and Technologies

### 6.1  Other archives

There is a gradually increasing usage of the various web lecture technologies in the educational field. The individuals publishing these lectures are generally motivated by the exploration of new technologies. They typically produce a single lecture or perhaps as much as an entire course.

There are several organizations, however, which have moved beyond using web lectures on an exploratory basis and are beginning to create large general-purpose lecture repositories with diverse content similar to the web lecture project at CERN. These efforts include:

- **Microsoft Seminars Online** – This effort is featured prominently on the Microsoft home page. The quality of the seminars is excellent and the topics range from detailed technical briefings to sales and marketing presentations. The content is captured from some of the Microsoft presentations at various professional venues. The lectures consist of both PowerPoint and audio, interspersed with animated screen shots of product demos. The overall experience is very good. There is little detail as to how the lectures are produced and the effort involved. The Microsoft search engine indexes the seminars and a direct link which launches a seminar may appear as a result of a text search.

- **Fermilab Media Services Streaming Lecture Archive** – This archive consists of about 300 lectures ranging from technical physics presentations to more general presentations for a general public audience. Fermilab uses our Sync-O-Matic software to produce its lectures.

- **Berkeley Multi Media Research Center – Lecture Browser** – This repository contains several hundred lectures ranging from courses to distinguished lectures. It is produced using software specifically developed at Berkeley which provides a novel lecture navigation interface showing a timeline with slide duration that allows instant navigation to any slide. The lecture browser also provides a keyword search of the lecture repository content.

- **Institute for Theoretical Physics** - This theoretical physics institute at the University of California at Santa Barbara has for many years recorded lectures and seminars and made the audio and slides available on the Internet. The presentation technology is not as advanced as Sync-O-Matic but the site has nevertheless had a strong impact on the physics community.

These lecture repositories, along with the CERN Web Lecture Archive Project, are beginning to demonstrate that this type of multimedia production can be scaled to the point that it becomes a normal part of operations for an organization's media production group.

### 6.2  Other technologies

The commercial applications most similar to Sync-O-Matic and ClipBoard are Microsoft's PowerPoint-2000 lecture presenting capability and RealNetworks RealPresenter. The fundamental problem with these commercial products is their lack of extensibility. Because there is no standardized interchange format, when these

commercial products are used, the content cannot be reused outside of the proprietary platforms for storage and serving. There are a number of smaller firms with web lecture production solutions [35] but these typically involve a proprietary Java-based player or other constraints on the reuse of the material.

While it may seem like the best interest of each commercial venture is to build closed proprietary technologies for web lectures, we feel that the content of web lectures should belong to the people who create the material. Depending on any single commercial technology for streaming video is dangerous when creating materials with a potentially long shelf life. Company priorities change over time and multimedia companies have shown a propensity for "breaking" their old software solutions in order to force users to upgrade to the newest (often incompatible) version of their product.

For users and organizations to invest heavily in web lecture technology, we feel that the only solution is an open standard which allows interoperability of many products and puts users in charge of their information.


## 7    Planned Future Applications and R/D

An endless variety of potential applications exist for the archive methods developed and tested by the project team. In this section, we discuss those we have found to be of primary interest to the high-energy physics community, focusing specifically on our experience recording lectures for the ATLAS experiment and for CERN. We note those particular services, for which we have identified a clear need for future support and activity.

### 7.1    ATLAS

Feedback received thus far indicates that a large, globally distributed collaboration such as ATLAS, has a strong interest in the archival of collaboration meetings and subsystem workshops, either to allow remote participation or as an accurate means of recording important events and milestones. Indeed, the initiative for recording the ATLAS plenary sessions came from the collaboration management and feedback from viewers indicates a clear demand for the archival of the plenary session, as well as smaller, subsystem meetings, to become a regular service.

In addition to the recording of meetings, the project examined the feasibility and utility of lecture archiving for the purpose of collaboration training. During the life of an experiment, it is common for new tools and/or procedures to be developed, which require the training of a significant fraction of the collaboration. In this case, the developer or expert may provide one or more training lectures or tutorials and present them to her or his colleagues at the laboratory and/or at several of the home institutes. Following the initial training session, however, new collaborators (or collaborators who could not attend the original session) also require training, albeit in much smaller groups and at times scattered through the duration of the experiment. For this reason, the team recorded lectures both during the scheduled training sessions and in sessions dedicated toward the production of a training archive [36] . Feedback from the collaboration again greatly supports this service and there is strong demand for its continuation.

As the experiments ramp up and the collaboration members are required to spend significant time at the laboratory, the need for collaborative tools to maintain student-teacher communication will increase. It is foreseeable that the lecture archiving techniques developed for the training sessions described above could be applied to help maintain and

re-enforce this important relationship. While, the existence of an archived lecture does not replace the presence of a teacher in the classroom, it can indeed be used to supplement the live lectures and to provide reference material. In addition, presentations made by a visiting student to his colleagues at the laboratory could be recorded for the benefit of the supervisor who is unable to attend.

With these applications in mind, we identify two primary services as being essential to the needs of a large collaboration such as ATLAS. The first of these is a service for the recording and archiving of lectures, meetings and events at the laboratory. This service would require the necessary audio and video recording equipment, software servers and presentation laptops, as well as a team of dedicated personnel to operate and maintain the equipment and to publish the lectures. The second service would be the provision of a personal lecture archiving facility, to be reserved by individuals wishing to record and publish lectures without the presence of an audience. This facility would be optimized for quality audio and video recording and would be maintained by the archive staff, which would be available to aid in the recording and publication of the lectures.

These services could be provided by the collaboration itself or, perhaps more efficiently, by the host laboratory. In this case, the laboratory could charge a standard facility fee to the collaborations on an event-by-event basis, to support the team and equipment. Such a payment method, similar to existing fees for video or phone-conference services, would avoid abuse of the system and could be integrated into a web-based agenda scheduling system, such as the CDS system at CERN.

## *7.2  CERN Particle Physics Distance Education Program*

As one of the main objectives of this project was to demonstrate that samples of CERN's rich education and training program could be made available for viewing on the web, it is natural to investigate if we can further capitalize on the success of this effort. To this end we are exploring the possibility of reviewing and repackaging the existing set of lectures to build a prototype CERN Particle Physics Distance Education Program.

The material will be classified under well-defined headings and eventually developed into curricula, e.g. accelerator technology, theoretical particle physics. Missing topics will be identified and proposed to the seminar organizing bodies as priority candidates for inclusion in future series of lectures. In this way it should be possible to build up a unique, catalogued set of lectures to support the learning and understanding of particle physics.

The target user would typically be students and working physicists who wish to complement and update their knowledge of the subject. It could also be used by academic institutions to support their existing particle physics education programs or provide basic material for those who are unable due to lack of resources to teach the subject at the present time.

The technological developments described previously will greatly facilitate the success of this venture by improving the archive process in terms of efficiency and quality of the result, by ensuring stability of the material and by giving access to an even wider audience with fewer technical constraints.

The CERN Academic Training Committee, the body that oversees, plans and organizes many of the seminars, recognizes the significant potential of this proposal and has endorsed the creation of a pilot project to develop these ideas.

### 7.3 Web accessible Basic Safety Training

Following up on one of the CERN WLAP test activities, it is intended to use the WLAP experience and tools to design and archive a suite of CERN Basic Safety Training courses. This will enable visitors to familiarize themselves with CERN safety procedures and to satisfy mandatory requirements prior to arrival. In this way they will save time and be able to complete the requirements in advance to get site access and access to the experimental areas rather than have to wait to attend standard courses.

Use of web lectures for this type of purpose implies the introduction of a validation process and probably will imply extended use of multi-media techniques.

### 7.4 Planned Future Technology Development

There are a number of basic areas of work that will be part of the ongoing project. We outline a few of them here.

The next step in the project is to place a mirror site at the University of Michigan Physics Department. Because this entailed republishing every single lecture, we decided to invest some time in further automating the publication of the lectures. So Giosuè Vitaglione developed a complete Linux based publishing capability. While this tool is still under development and not yet ready for release, its basic features include: (1) a Perl script (called CopyCat) to convert a published Syncomatic web lecture in the Lecture Object format, (2) a Java based publishing program [34] able to parse the Sync-O-Matic style file syntax (affectionately known as ChuckScript), (3) producing the HTML+Javascript lecture using of the Linux RealProducer capabilities.

The server at University of Michigan will use this new software. When it will be complete, the archive will be available at http://wlap.physics.lsa.umich.edu. This undertaking will help us perfect the process of creating new mirror sites to assist others desiring to do the same, provide a testbed for testing and extending various applications, and start a US repository for archives likely to be heavily used there.

In addition, we intend to address the following set of specific tasks in the months ahead -

- While using RealMedia has provided excellent audio and video quality for the amount of network bandwidth and disk space it consumes, we are concerned that the format is not suitable for long-term archival purposes. We are leaning towards archiving a higher quality format such as Mpeg2 or AVI that can be converted into any number of web-capable formats including RealMedia, QuickTime Streaming, and Microsoft Media. While this takes more disk space, it will insure the long-term viability and reusability of the media. As part of the publishing process, we will dynamically encode the media into the appropriate web-ready format. ClipBoard-2000 already captures higher quality media and then down samples it for the web.

- As we add dynamic encoding capabilities, we can present the user with different quality levels based on their connection and viewer capabilities. An experimental version of ClipBoard-2000 Server Edition takes a single uploaded lecture and publishes it on the server in QuickTime, Sync-O-Matic, and Mp3 formats.

- We propose to convert ClipBoard-2000 and Sync-O-Matic-2000 to use the Lecture Object as their native intermediate formats while maintaining backward compatibility.

- We intend to finish development on the Linux publishing capability and release that product.

- We will continue to work on improving the capture capabilities for the single end-user using PowerPoint. Simple prototypes of plug-ins to PowerPoint have been developed which make lecture capture extremely simple.

The ultimate goal is to have a strong suite of tools which all operate interchangeably using the Lecture Object as their interchange format. As discussed above, we have begun to explore the pursuit of a standard for a Lecture Object to allow commercial tools to produce and/or publish Lecture Objects in a standardized way.

## 8    Conclusions

The need is burgeoning for cost effective, ubiquitous techniques for allowing users to access content-rich materials via the World Wide Web. Access needs to be reliable, available upon demand, and of high quality. Key elements of live presentations should be retained in terms of audio, video and lecture materials.

In the WLAP Project we have attempted to demonstrate that the technology exists for achieving these goals, and that the promise for future improvements is strong. Even with modem network access at 56kb, a free browser, and inexpensive speakers a user anywhere can experience and conceptually understand a Nobel lecture given at any other place in the world. In the WLAP Project we have focused on how to continuously improve the quality of the end-user experience. We also were aware of the importance of placing minimal requirements on the speaker and on the lecture setting. We evolved and developed software and hardware solutions to make the best use of whatever materials the speaker made available. These materials ranged from complex PowerPoint presentations, to FrameMaker documents, to PDF files, all the way to a blank blackboard and a piece of chalk.

Because Sync-O-Matic was open and extensible, innovation could be incorporated quickly on a number of fronts by different participants in the project. We have begun an effort to develop a usable standard for interchange between multiple lecture publishing environments. With the growing number of sites developing the capability for producing high quality archives, one can imagine the day when a rich variety of materials will be available globally, in searchable form. This will, of course, require adherence to certain standards, and we are engaging in activities that will encourage the development of these standards.

Web-based archiving of content rich material is a very important undertaking. We have only glimpsed what will be possible in this arena. Higher bandwidth networks, bandwidth reservation technologies, high definition displays, as well as continued research in reaching and learning modalities, will, in our belief, make web-based archiving the centerpiece of many future educational, training and scientific collaboration activities.

## 9    Acknowledgements

## 10 References

[1] The proposal of the WWW, Tim Berners-Lee, CERN March 1989, May 1990

[2] Proposal for a Web-Based Lecture Archive System for CERN, S. Goldfarb, National Science Foundation Project Proposal, (1999).
http://webcast.cern.ch/Projects/WebLectureArchive/Project/Proposal99.pdf

[3] The Sync-O-Matic software synchronizes lecture slide presentation
with video playback to make on-line lectures.
http://www.syncomat.com

[4] Project Summary: A Web-Based Lecture Archive System for CERN,
S. Goldfarb, E. Falaise, National Science Foundation Project Report, (1999).
http://webcast.cern.ch/Projects/WebLectureArchive/Project/Summary99.pdf

[5] CERN HR division, Training and Development group.  http://www.cern.ch/training/

[6] The ATLAS experiment is being constructed by 1850 collaborators in 150 institutes around the world.  http://atlasinfo.cern.ch/Atlas/Welcome.html

[7] University of Michigan, Media Union.  http://www.ummu.umich.edu/

[8] Real Producer, RealNetworks.  http://www.realnetworks.com/products/producer/

[9] Hiding the Infinities, Prof.M.Veltman, 1999 Nobel Prize in Physics.
http://webcast.cern.ch/Projects/WebLectureArchive/nobel/veltman/

[10] Bringing the Web to America, P.Kunz.
http://wlap.physics.lsa.umich.edu/cern/lectures/colloq/1999/kunz/

[11] "Hands-on Introduction to GEANT4", Andrea dell'Acqua. From 26/2 to 3/3/, University of Michigan, Ann Arbor.  http://wlap.physics.lsa.umich.edu/atlas/umich/geant4/2001/

[12] Microsoft, PowerPoint, http://www.microsoft.com/office/powerpoint/

[13] CarpePpt, timing capturing.  http://www.vitaglione.it/giosue/wl/carpeppt/

[14] Postscript, Adobe.  http://www.adobe.com/products/postscript/

[15] Acrobat, Adobe. http://www.adobe.com/products/acrobat/readstep.html

[16] Carpe PDF, timing capturing.  http://www.vitaglione.it/giosue/wl/carpepdf/

[17] ClipBoard 2000.  http://ct.si.umich.edu/cb2k/

[18] Snaps-o-matic.  http://www.vitaglione.it/giosue/wl/snapsomatic/

[19] "Lecture object : an architecture for archiving lectures on the Web", G.Vitaglione, N.Bousdira, S.Goldfarb, H.A.Neal, C.Severance, M.Storr

http://weblib.cern.ch/abstract?CERN-OPEN-2001-020
[20] Lecture Object Architecture Web Page.  http://home.cern.ch/giosue/lectureobject/
[21] Lecture Object specifications draft. (obsolete)
  http://home.cern.ch/giosue/wlap/docs/wlap-20000217.doc
[22] XML, eXtensible Markup Language, http://www.w3.org/XML/
[23] Synchronized Multimedia Integration Language,  http://www.w3.org/AudioVideo/
[24] IEEE P1484.12 Learning Object Metadata Working Group, http://ltsc.ieee.org/wg12/
[25] Dublin Core Metadata Initiative,  http://purl.org/dc/
[26] IMS Global Learning Consortium. http://www.imsproject.com/
[27] ARIADNE Foundation for the European Knowledge Pool, http://www.ariadne-eu.org
[28] GESTALT Project, Getting Educational Systems Talking Across Leading-Edge Technologies, http://www.fdgroup.com/gestalt/about.html
[29] CEN ISSS LT, http://www.cenorm.be/isss/
[30] Quality of Service networking, QoS.Net Project, http://www.qos.net/
[31] The Open Archives Initiative. http://www.openarchives.org/
[32] PHP Hypertext Preprocessor,  http://www.php.net/
[33] RealNetworks. http://www.realnetworks.com
[34] Websentation, Web-Presentation. http://www.websentation.com
[35] Tegrity, http://www.tegrity.com/
[36] WLAP, ATLAS Training,  http://wlap.physics.lsa.umich.edu/atlas/

### Bibliography

1. Andrew S.Tanembaum, Computer Networks, 1996, Prentice Hall, ISBN:0133499456
2. CALIBER-NET: Quality in European Open and Distance Learning
   http://www.eadtu.nl/activities/ projects/CALIBERNET/default.htm
3. D.Carnevale, The Chronicle of Higher Education, Recent trends in distance education, Feb.18, 2000, Vol.59, http://chronicle.com/indepth/distance/
4. DOM, Document Object Model,  http://www.w3.org/DOM/
5. Robert Cailliau, James Gillies, How the Web Was Born: The Story of the World Wide Web, Oxford Univ Pr., ISBN: 0192862073
6. Distance Education Clearinghouse, University of Wisconsin,
   http://www.uwex.edu/disted/
7. DUNE: Distance Education Network of Europe
   http://www.eadtu.nl/activities/projects/DUNE/default.htm
8. Edd Dumbill, The role played by XML in the next-generation Web,
   http://www.xml.com/pub/2000/09/06/distributed.html
9. MPEG, Moving Picture Expert Group, http://www.cselt.it/mpeg/
10. Reuven Aviv, Educational Performance of ALN via Content Analysis Open University of Israel,   Journal of Asynchronous Learning Networks Volume 4, Issues 2 - September 2000 ISSN 1092-8235
    http://www.aln.org/alnweb/journal/ Vol4_issue2/le/reuven/LE-reuven.htm
11. Resource Description Framework (RDF), W3C http://www.w3.org/RDF/
12. Robin Mason, Models of Online Courses, Institute of Educational Technology, The Open University, ALN Magazine Volume 2, Issue 2 - October 1998 http://www.aln.org/alnweb/magazine/vol2_issue2/Masonfinal.htm
13. Simple API for XML, http://www.megginson.com/SAX/

14. Synchronized Multimedia Integration Language, http://www.w3.org/AudioVideo/
15. Tim Berners Lee, Semantic Web Road map, September 1998, http://www.w3.org/DesignIssues/Semantic.html
16. THETA Project: Telematics in Higher Education   http://projects.europace.be/theta/
17. Greg Kearsley, Explorations in Learning & Instruction: TIP http://www.gwu.edu/~tip/
18. Video Codec WareHouse, http://www.geocities.com/codecwarehouse
19. World Wide Web Consortium, "Leading the Web to its Full Potential", http://www.w3.org/
20. Scott B. Wgner, et Al., The Effects of Internet-based Instruction on Student Learning, JALN Volume 3, Issue 2 - November 1999, http://www.aln.org/alnweb/journal/Vol3_issue2/Wegner.htm
21. Tim Berners-Lee, Mark Fischetti, Weaving the Web, Harper San Francisco; Hardback: ISBN:0062515861,