XML saves the day: Porting a Rich-Media Collection to a Mobile Platform in Three Weeks Flat

Rodrigo Arias. CEO, nTropic, Inc., Guatemala. *rodrigo@glifos.com*
Matthias Reichenbach. Lead XML Developer, nTropic, Inc., Guatemala. *matthias@glifos.com*
Grete Pasch. Director, Academic Information Resources, Universidad Francisco Marroquín, Guatemala. *gpasch@ufm.edu.gt*

Abstract

Since 2001, we develop specifications and tools for creating and deploying rich-media content. Our GLIFOS Markup Language (GML) is an XML-based specification that separates the content (videos, transcriptions, slides) from the rendering of the rich-media player. Using our tools, the University Francisco Marroquín (UFM, Guatemala) has created and hosts an online library of over 300 hours of enriched streaming video (www.newmedia.ufm.edu) to support web-based and classroom teaching.

GML is technology, platform, and format independent. We claim that these qualities guarantee content portability to diverse platforms in use today, as well as to those that will arise in the future. We had backed this claim by proving how simple it is to port GML to the leading platforms (e.g., Real Player and Windows Media Player), however, we had yet to port it to a truly new platform—that is, one that arose after the design of our XML specification and our content creation tools.

This paper describes our XML specification and content creation tools, and gives a week by week description of our experience porting the UFM's collection to the new PDA models. Such mobile devices are able to stream video through their Wi-Fi connection, without clogging their CPUs or draining their batteries. Finally, we discuss how this experience has validated the importance of XML for digital preservation and our tools' ability (so far) for porting content to new platforms.

Keywords

Rich-media, digital preservation, XML/XSLT, mobile learning

1 Introduction

When working with rich-media, we always keep two basic principles in mind. First, rich-media content should integrate video, audio, text, and images in such a way that the richness of each individual format is fully realized. For instance, the more convenient way to present a list of tools and materials may be plain text, while video is a richer medium for reproducing the process of making a particular knot, and a high-res image can show in great detail the *broca* worm trap you need for protecting your coffee trees (Fig 1). This mix of media also helps support various learning styles and reinforces the topic being taught in complementary ways.

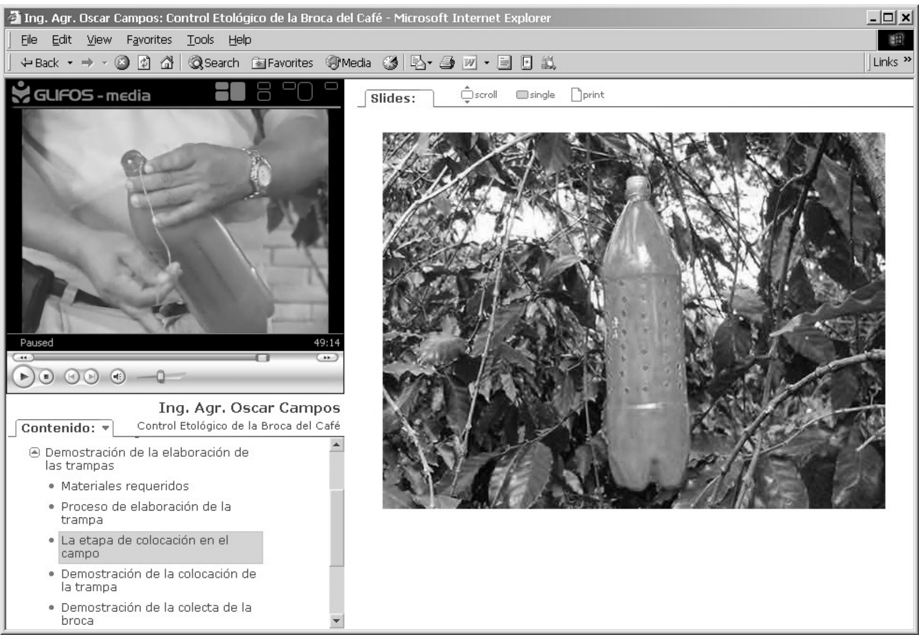

Figure 1. Building a *broca* worm trap: rich-media presentation of a manual process

The second principle we refer to is inspired in the time-tested paratextual devices employed in bookmaking—that is, we strive to give video and audio the functionality that a *text* enjoys once it is built into a *book* (Genette 1997; Pasch 2004, p.33-40.) Thus, among other actions, users should be able to easily navigate the media content, to find specific sections or terms, to make annotations, to bookmark specific parts, and to extract references. Such features allow users to interact with the media, which has been shown to support learning (Shephard, p. 296)

One perceived limitation of rich-media is its lack of interaction, both between student and instructor and among students. This "limitation" is shared by books, where the interaction between reader and author is non-existent or, at best, asynchronous. Still, for centuries the written word has successfully been (and will continue to be) an important part of most learning environments, just like "on-line learning resources that embed streamed video can increase opportunities for independent learning" (Green et al., 2003). Although neither books nor rich-media by themselves may satisfy the learning needs of those students who prefer face to face learning, recent studies show that students enrolled in distance learning programs prefer by far the convenience and control of asynchronous media content over more interactive and expensive synchronous set-ups and even face to face classes (Baessa & Fernández, 2002; Beyth-Marom et al, 2005; LeBlanc, 2004.)

The value of rich-media resides in the aggregation of content into a *collection* large enough to support academic reference and research—that is, a rich-media library. We have been building a rich-media library since 2001, and by mid-2005 it contains over 500 hours of content. Obviously, the value of such a library depends on its usage, both in the short and in the long term. In the short term, our software tools should facilitate access to its content, enabling users to locate individual segments within rich-media contents via the full-text search capabilities of the library catalog. And to ensure its long term usage, our digital preservation strategy should guarantee content portability to future platforms.

Within this context, we decided to port our rich-media collection to the new wireless PDA platform. We did so for several reasons:

1. **Proof of concept**. We have argued elsewhere that our XML-based digital preservation strategy guarantees portability to future platforms (Pasch & Arias, 2003; 2004), but until now, we had yet to port content to a truly new platform-- that is, one that arose after the design of our XML specification and the deployment of our content creation tools.

2. **Lower access cost**. Currently, a PDA with WiFi and color display costs half as much as a PC, and only one third of a laptop. Thanks to its high mobility, PDAs are even more suitable than laptops for some learning environments.

3. **A research tool**. Some trends indicate PDA functionality is being subsumed within mobile telephone devices, while such devices continue improving its processing capacity, display resolution and data bandwidth. Thus, a wireless PDA within university campus boundaries gives us today a test bed for what will be available out in the "real world" in the coming years.

2 Rich-media Content Creation

Rich-media production at the Universidad Francisco Marroquín (UFM) in Guatemala began in mid-2001 at the newly formed New Media department (www.newmedia.ufm.edu.) The purpose of this initiative was preserving unique, locally produced content and distributing it widely via the Internet (Pasch & Stewart, 2002; Pasch, 2004a.) By mid-2005, over 500 hours of streaming video have been produced in-house and deployed as rich-media content. About 90% of these videos present one-hour lectures enriched with tables of contents and/or PowerPoint slides. There are two main uses for such content: as primary material for web-based courses and as reference sources in both distance and face to face classes. Courses include topics like business ethics, forensic science, and everyday logic. Speakers include important and once in a lifetime visitors to the UFM campus, such as Vernon Smith (Nobel in economics), Jeff Taylor (CEO, monster.com) and Carol Brey (President, ALA). Often, such visitors may show up unexpectedly, but the New Media team is able to move from preproduction to completed, web-posted product within the same day. Successful and fast development depends on a well thought-out workflow, which briefly consists of three steps.

1. Preproduction: discuss desired uses for presentation and plan the production schedule

2. Media production:
    a. Tape the presentation, or convert existing video or audio to digital format
    b. Gather and prepare additional materials, such as accompanying images (projected PowerPoint slides, overhead slides, photo slides, blackboard scribbles and drawings), audio/video (VHS, CD, DVD, mp3, etc.), printed materials (handouts, photocopies, books) and other digital objects (websites, simulations, demo programs and data files).
    c. Digital editing of the video or audio. When required, some additional materials may be edited into the final video or audio file
    d. Media rendering into desired format/s (streaming Windows media, QT, etc)

3. Rich-media production:
    a. Index and synchronize the media file with the additional materials
    b. Render all elements for viewing on desired platforms
    c. Deploy via Internet or CD
    d. Archive physical and digital products

A project manager experienced in rich-media production is in charge of preproduction and also checks all intermediate and final results. Media production is under the direct supervision of video production professionals who are aware of the needs for streaming media and rich media products. The rich-media production stage requires the application of specific rich-media tools, such as the GLIFOS-media suite developed by us (Figure 2) and applied by subject specialists for the analysis and publishing of content.

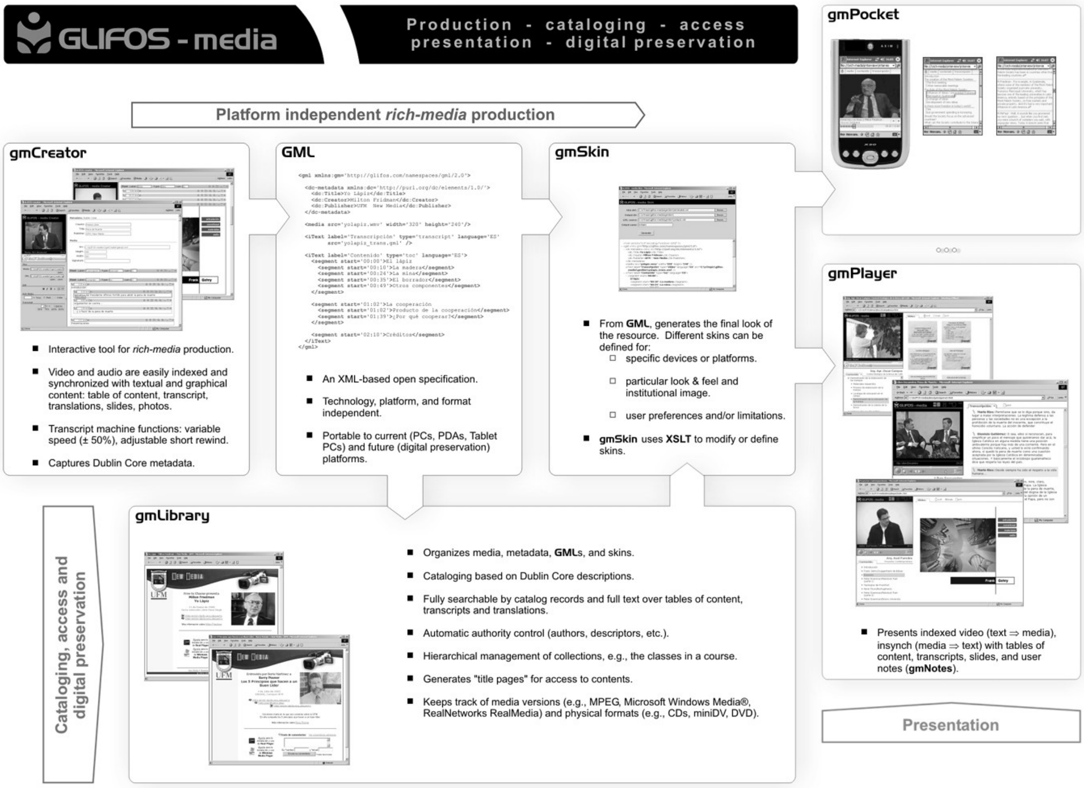

Figure 2. The GLIFOS-media toolset

First, gmCreator provides an easy to use, graphical interface for media indexing, transcription, synchronization, and metadata markup (Figure 3). For applications that require licensing, content can be signed with one mouse click. This module outputs a GML (GLIFOS Markup Language) file that follows the specification described in section 3.

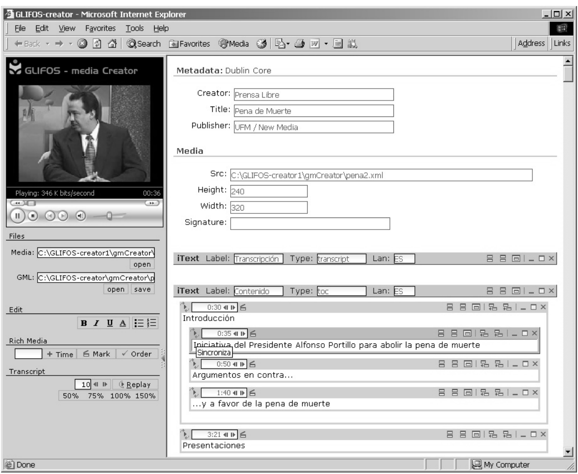

Figure 3. The gmCreator interface

Next, gmCreator is used to publish the final rich-media presentation by applying the desired skin or skins for rendering on various platforms. The skins are expressed as XSLT files that are tailored to each client's requirements. Figure 4 shows three views of the same presentation. The presentation can be published to a website or can be distributed on CDs or DVD-RAMs.

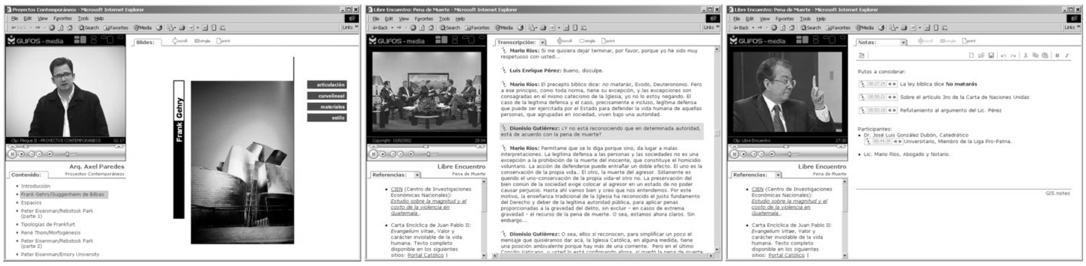

Figure 4. Final rich-media presentations created with the GLIFOS-media toolset. Left: a "normal" view with streaming video, clickable table of contents, and PowerPoint slides. Middle: video, links to related web resources, and full-text transcript. Right: video, references, and annotation pane that allows users to synchronize the video to their own notes.

Finally, gmLibrary supports the creation of a rich-media library by keeping track of the various elements in the production of rich-media, that is: the media files themselves, the skins used, the corresponding GML files, final renderings, and auxiliary files.

3. A case for content oriented XML

An open and widely adopted XML-based standard does not suffice to ensure the digital preservation of content. By the late 1990s, SMIL 1.0 (Synchronized Multimedia Integration Language) was a widely accepted XML standard, and RealNetworks' RealPlayer was the platform of choice for both streaming and locally stored media. It is easy to understand the basic component and features of SMIL by studying a simplified example. Figure 5 shows a video enriched with a table of content and synchronized transcript.

Figure 5. SMIL-based rich-media content shown with the RealPlayer

The SMIL files needed to generate this video are shown in figure 6. First, **`f.smil`** shows how `<layout>` distributes the three content regions. Note the functional dependencies in the `<region>` attributes `top`, `width` and `height` (Glushko & McGrath, 2002). Next, **`f_toc.rt`** shows the tricks that are needed to properly indent the table of contents (e.g., `<ul><ul>`). However, none of these layout and formatting codes has any meaning on other platforms. On the small display screens of mobile devices it would make more sense to show just one "region" at a time (video, table of contents or transcript) and let the user switch between them. For such an interface the `<layout>` section has no meaning. Similarly, the `<ul><ul>` antics work only for RealPlayer.

SMIL File: **f.smil**

```
<smil xmlns="http://www.w3.org/2001/SMIL20/Language">

<head>
  <meta name="author" content="Hoover Institution" />
  <meta name="title" content="Entrevista con Rose y Milton Friedman" />
  <meta name="copyright" content="(c)2004 InfoLib, S.A." />

  <layout>
    <root-layout backgroundColor="#000000" width="540" height="410" />
    <region id="wmedia" top="0" right="0" width="320" height="240" />
    <region id="transcript" top="250" right="10" width="300" height="100%" />
    <region id="toc" top="0" left="0" width="220" height="410" />
  </layout>
</head>
<body>
  <par endsync="wmedia">
    <video region="wmedia" src="f.rm" />
    <textstream region="toc" src="f_toc.rt" />
    <textstream region="transcript"
                src="f_transcript.rt" />
  </par>
</body>
</smil>
```

RealText File: **f_toc.rt:**

```
<window version="1.5" type="generic"  width="220" height="420"
        bgcolor="#282828" link="#B0B0C0" underline_hyperlinks="false" >
<pos x="15" y="20"/>
<font face="Arial" color="#D0D0D0" size="4">
<b>Contenido</b></font>
<font face="Arial" color="#404040" size="3">
  <ul>
    <a target="_player" href="command:seek(0:00)">
      <font color="#5555CC"><b>I</b></font>ntroduction
    </a>
  </ul>
  <ul>
    <a target="_player" href="command:seek(0:15)">
      <font color="#5555CC"><b>T</b></font>he creation
       of the Mont Pelerin Societies
    </a>
  </ul>
  <ul><ul>
    <a target="_player" href="command:seek(0:25)">
       <font color="#5555CC"><b>T</b></font>he first meeting
    </a>
  </ul></ul>
  <ul><ul>
    <a target="_player" href="command:seek(3:02)">
      <font color="#5555CC"><b>O</b></font>ther memorable meetings
    </a>
  </ul></ul>
   ...
</font>
</window>
```

Figure 6. Simplified SMIL and RealText files

Moreover, if the SMIL code is not generated but hand-coded, different content will surely end up with different tagging techniques that produce similar or different results. Even worse, as RealPlayer evolved through its versions G2, 7, 8, One, etc., the SMIL code also evolved to take advantage of the new features and to adapt to the new specifications (SMIL 2.0). As a result, porting a hand-coded rich-media collection based on SMIL and RealPlayer that started back in 1998 to newer platforms might pose a hard problem. Rule based systems may work if just a manageable number of different patterns are needed, but developing and fine-tuning those systems takes time, and their results are rarely 100% accurate.

Since digital content collections routinely outlive their comparatively short-lived platforms and standards, and migration costs rise proportionally as collections grow, the long term perspective of digital preservation becomes more apparent and urgent. This perspective guides our implementation priorities in several ways: 1) content becomes much more important than the current system, 2) each system should facilitate content migration to the next standard, and 3) content should be stored in a format that is as independent from current technologies as possible.

```
<gml xmlns:gm='http://glifos.com/namespaces/gml/3.0'>
  <dc-metadata xmlns:dc='http://purl.org/dc/elements/1.0/'>
    <dc:Creator>Hoover Institution</dc:Creator>
    <dc:Title>Entrevista con Rose y Milton Friedman</dc:Title>
    <dc:Publisher>UFM  New Media - Guatemala</dc:Publisher>
  </dc-metadata>

  <media src="f.rm" width="320" height="240"/>

  <iText label="Contenido" language="en" type="toc">
    &lt;segment start="0:00"&gt;Introduction&lt;/segment&gt;
    &lt;segment start="0:15"&gt;The creation of the Mont Pelerin Societies
      &lt;segment start="0:25"&gt;The first meeting&lt;/segment&gt;
      &lt;segment start="3:02"&gt;Other memorable meetings&lt;/segment&gt;
     ...
    &lt;/segment&gt;
  </iText>
</gml>
```

Figure 7. GML 3.0 example

The GML version of the previous example is given in Figure 7. This presents exactly the same content as the SMIL example, while avoiding all previous inconveniences:

- It avoids any ephemeral layout and format information.
- It avoids functional dependencies (Glushko & McGrath, 2002).
- It eliminates ambiguity, e.g., table of contents hierarchy is formally specified.
- Changes in platform versions and format do not affect the content.

We have verified the advantage of GML in practice. Every time we have ported GML content to a different platform, or to upgraded versions, we have effortlessly and consistently generated the code for those platforms. For instance, it is simpler to generate SMIL from GML than vice versa (in fact, the example shown in Figure 6 was generated from GML). More than a curiosity, this is evidence of the importance of semantic information -- for translating SMIL to GML we need to add knowledge, but for mapping GML to SMIL we just need to derive code.

After showing how simple it is to port GML content to the leading platforms (e.g., RealPlayer + SMIL and Windows Media Player + DHTML) and noticing that nothing in GML specifically considers those platforms, we argued that GML's technology, format and platform independency ensures our content portability to future platforms. However, we had yet to port it to a truly new platform—that is, one that had appeared after the design of our XML specification and the implementation of our content creation tools. When the current PDA models were able to stream video through their Wi-Fi connections without clogging their CPUs or draining their batteries, we were eager and ready to validate our hypothesis with a real life case.

4. Portability in real life: UFM case study

4.1. Week 1: mobile platform evaluation and development of a functional prototype

As soon as the Dell Axim X50 520MHz Pocket PC arrived, we ran some basic tests to measure its prowess. In contrast to older Pocket PC models, the new Axim was very capable to down stream via its WiFi a full hour of near VHS quality video full-screen (512kbps, 320x240) without draining its batteries. After such promising results, we proceeded to review the available software platforms for presenting rich-media on Pocket PCs. We decided not to use Macromedia Flash Player for Pocket PC as this point, mainly because it isn't (yet) a strong player in the streaming media segment. A second possibility was the Microsoft .NET

Compact Framework, however, this platform is specific for Microsoft's Windows Mobile operating system, which would hinder our porting our implementation to other mobile operating systems. Finally, we decided to double check Microsoft's Pocket Internet Explorer. Although the name may imply that this is sort of a Mini Me (Wikipedia, 2005) version of Internet Explorer 6.0 for PCs, it is actually a very different, limited version of IE 6.0 that does not support most of the DHTML functionality.

We made a list of the essential functionality that Pocket IE must provide to be a feasible platform, and we organized the list into two categories:

- Access and control of the Microsoft Windows Media Player mobile (WMPm) object model: Pocket IE should be able to stop, pause and play the WMPm at specific times, WMPm should provide connection status and progress time information, and there should be a way to generate a time based event.

- Layout redesign: since the pocket PC resolution (320x240pixels) is less than 1/10th of the XGA resolution (1024x768pixels), a complete layout redesign was needed. The new layout should be able to present just one rich-media format at a time (video, table of contents, transcript, etc) and switch between them easily.

After a brief testing and validation period, we weighed the inherent platform limitations. For instance, since the media progress slider built into WMPm wasted too much display real estate, we decided to implement our own slider. Also, for the time being we decided to present just the first two levels of the table of contents (more levels looked confusing in such a narrow space), and we did not include slide presentations (although we added them later).

At this point, we had a working prototype (Figure 8) but we weren't concerned with look and feel, because one of the main advantages of XML/XSLT is the ability of separate content and format not just at the conceptual level, but also at the developer's role level.

Figure 8: First-run results (Week 1)

4.2. Week 2: XSLT based skin development for rendering content on mobile platform

After producing a functional proof of concept, the rest of the work became quite mechanical. Our graphic designer proposed a look for our new player, and our skin developer started the generalization of the working prototype to create its corresponding skin. This particular skin is mostly XSLT code that automatically generates the HMTL and Javascript code needed to present the rich-media on the new Pocket PCs. Both at the design and skin creation level, special care was given to several factors:

- User interface should be intuitive and natural
- A clean and clear look and feel was desired
- Any wasted space should be eliminated

Figure 9 shows the final result of the new skin design.

Figure 9: Final results (Week 3)

4.3. Week 3: streaming media capabilities, testing and refinement

The remaining tasks, before the beta release for end user testing and feedback, were:

- Skin generalization and validation using a wider sample of GML contents
- Real life testing (e.g., campus-area testing, checking Internet bottlenecks)
- System testing, bug fixing and fine-tuning

Even though we have performed a robust unit testing during weeks 1 and 2, the system as a whole exhibited some problems during its real life tests. For example, after long media buffering delays through Internet bottlenecks, we decided to add a buffering progress message. Some noise artifacts were solved when our skin developer eliminated certain inefficiencies in the Javascript code.

At the end of the third week, the skin was beta released for batch and on-the-fly rendering of rich-media content for pocket PCs. During the whole process we did not have to change anything in our GML specification, nor did we feel the urge to do so. The advantages of separating content and format were evident all the way down to the developers' roles. This clean cut division of specialized labor allowed us to build a robust port to a brand new platform in three weeks flat from the time the Pocket PC arrived to our office.

5. Conclusions

The experience described in this paper has led us to appreciate the advantages of our approach for digital preservation, portability to different platforms, and ease in the division of labor.

First, two well-known threats to digital information are deteriorating media and obsolete hardware. Preventive measures include periodic re-copying, LOCKSS ("lots of copies keep stuff safe"), and maintaining required hardware in working order. However, even if one can successfully read the data off a 20-year old floppy disk, there's still the question of how much one can do with that data. If one is limited to running old applications (say, Visicalc) one is also limiting the usefulness of the data (for example, the ability to include newer data, to create graphs or interact with other applications) unless it is migrated to a modern format (such as Excel.) The GLIFOS-media system offers clear advantages related to preservation. First, GMplayer does not rely on "marks" embedded into the video or audio files. Thus, if the media needs to be re-rendered (e.g., for a new codec) it is not necessary to re-mark the new file with indexing points. Second, one must assume that a new set of tools may emerge at anytime, and if this happens, the most valuable asset one needs to hold on to is data. In the case of rich-media, this includes first and foremost all indexing, transcription, and synchronization information. The GLIFOS-media system uses XML to store all data and metadata. XML files are simple text files, and their structure is human-readable, therefore no applications are absolutely required in order to exploit the data, and the data can be easily migrated to new, desired formats. In addition, the GML files contain its own metadata, down to the level of detail desired by the producer, thus no separate metadata system is needed.

Regarding portability, we believe this experience is a step in the right direction in porting our content to mobile platforms. First, since the UFM is a WiFi campus, a reliable, fast, and inexpensive platform is in place to test our functionality. Second, the functionality of PDAs may end up being subsumed within mobile telephony, but at the same time mobile telephony is improving its bandwidth, functionality and screen resolution.

All of these factors help lower the cost of managing the content and extending its lifespan. Although XML offers a very useful and novel platform, it does not by itself guarantee good specification design, in the same way that using a relational database management system does not guarantee a good database design. Methodologies specific for the particularities of XML are needed.

6. References

Ayars, Jeff et al., editors. 2005. *Synchronized Multimedia Integration Language (SMIL 2.0) - [Second Edition]*. Available at: `www.w3.org/TR/smil20/`

Baessa, Yetilú and Javier Fernández. 2002. *Comparación de dos modalidades de enseñanza del curso de lógica y retórica en la Universidad Francisco Marroquín, Guatemala*. Unpublished research report. Available at: `www.newmedia.ufm.edu.gt/logica/baessafernandez.pdf`

Beyth-Marom, Ruth, Kelly Saporta, and Avner Caspi. "Synchronous vs. asynchronous tutorials: factors affecting students' preferences and choices." *Journal of Research on Technology in Education*, 37 (3, Spring 2005): 245-262.

Genette, Gerard. 1997. *Paratexts: thresholds of interpretation*. Cambridge: Cambridge University Press.

Glushko, Robert J. and Tim McGrath. 2002. "Document engineering for e-business". Proceedings of ACM DocEng'02, November 8-9, 2002. Pp. 42-48.

Green, Sue M. et al. "Evaluating the use of streaming video to support student learning in a first-year life sciences course for student nurses." *Nurse Education Today*, 23 (4, May 2003): 255-261.

LeBlanc, Glenn. 2004. "Student and faculty survey reveals attitudes to streaming video." *Distance Education Report,* 8 (12, June 1, 2004): 8.

Pasch, Grete. 2004. *Identifying, selecting, and organizing the attributes of web resources*. Unpublished dissertation, University of Texas at Austin. Available at: `www.ischool.utexas.edu/~gpasch/dissertation/gpasch_dissertation.pdf`

Pasch, Grete. 2004a. "Cautious but Decisive: Ten Years of Information Services Implementation at the Universidad Francisco Marroquín Guatemala," *Science and Technology Libraries,* 23 (2/3, 2004): 171-179.

Pasch, Grete and Rodrigo Arias. 2004. *eLearning Libraries: Managing rich media content in an academic institution*. Proceedings of the 70th IFLA World Library and Information Congress, August 2004, Buenos Aires, Argentina. Available at: www.ifla.org/IV/ifla70/papers/140e_trans_Pasch_Arias.pdf

Pasch, Grete and Rodrigo Arias. 2003. *Fast, cheap, and in control: managing metadata for streaming media*. Poster presentation at IA2003, Portland, OR. Available at: www.newmedia.ufm.edu.gt/gpasch/ia2003/pasch_arias.html

Shephard, Kerry. 2003. "Questioning, promoting and evaluating the use of streaming video to support student learning". *British Journal of Educational Technology*, 34 (3): 295-308.

Wikipedia. 2005. "Mini Me" article. Available at en.wikipedia.org/wiki/Mini-Me