Interactive publications: creation and usage

George R. Thoma*, Glenn Ford, Michael Chung, Kirankumar Vasudevan, Sameer Antani
National Library of Medicine, 8600 Rockville Pike, Bethesda, MD, USA 20894

ABSTRACT

As envisioned here, an "interactive publication" has similarities to multimedia documents that have been in existence for a decade or more, but possesses specific differentiating characteristics. In common usage, the latter refers to online entities that, in addition to text, consist of files of images and video clips residing separately in databases, rarely providing immediate context to the document text. While an interactive publication has many media objects as does the "traditional" multimedia document, it is a self-contained document, either as a single file with media files embedded within it, or as a "folder" containing tightly linked media files. The main characteristic that differentiates an interactive publication from a traditional multimedia document is that the reader would be able to reuse the media content for analysis and presentation, and to check the underlying data and possibly derive alternative conclusions leading, for example, to more in-depth peer reviews. We have created prototype publications containing paginated text and several media types encountered in the biomedical literature: 3D animations of anatomic structures; graphs, charts and tabular data; cell development images (video sequences); and clinical images such as CT, MRI and ultrasound in the DICOM format. This paper presents developments to date including: a tool to convert static tables or graphs into interactive entities, authoring procedures followed to create prototypes, and advantages and drawbacks of each of these platforms. It also outlines future work including meeting the challenge of network distribution for these large files.

Keywords: interactive publication, dynamic table, multimedia document, biomedical media

1. INTRODUCTION

In a remarkably prescient narrative describing the multimedia digital libraries of tomorrow, Lindberg and Humphreys[1] envision a scenario featuring "rich interconnections among genetics research data, aggregated clinical and public health data, published literature, and high-quality health information in many languages." Central to that vision is the existence of published literature, readily accessible online, that not only contains rich multimedia but also allows the reader to manipulate, use and analyze the information contained therein.

Multimedia documents have been in existence for a decade or more, and in common usage they refer to entities that consist of text that links to images and video clips. More often than not, the latter media types reside in databases apart from the text, and are accessed independently of the text, often leading to a loss of context.

The challenge is to create a comprehensive, self-contained and platform-independent multimedia-rich "interactive publication." By self-contained we mean the document is either one large file embedded with all media objects, or a 'folder' in which the files are tightly linked. This document, of either 'embedded' or 'folder' type, could contain many media objects: text, video, audio, bitmapped images, spreadsheets, presentation graphics, or animation sequences.

While using such a document, the reader should be able to: (a) view any of these objects on the screen; (b) hyperlink from one object to another; (c) interact with the objects in the sense of exercising control over them (e.g., start and stop video); and *(d) reuse the media content for analysis and presentation.*

It is the last objective that differentiates "traditional" multimedia documents from an "interactive publication." Such a publication (none appears to exist today) would have the characteristics of a 'document' in the sense of a completed work of an author presenting hypotheses, findings and conclusions, but also gives a reader the facility to check the underlying data and possibly derive alternative conclusions leading, for example, to more in-depth peer reviews. In

* Contact: George Thoma (thoma@nlm.nih.gov); phone +1-301-496-4496; fax +1-301-402-0341

Appears in Proceedings of IS&T/SPIE Electronic Imaging 2006: Digital Publishing. San Jose, CA. Jan 15-19, 2006, SPIE Vol. 6076.

other words, the “document” serves as a research tool. In addition to this “research tool” aspect, from a library’s preservation point of view, a self-contained document is clearly preferable to avoid dependence on the content providers (publishers) to maintain their databases of multimedia items in perpetuity, a point forcefully made by Lindberg[2].

There are major challenges in the quest for an interactive publication. One is the need for standards that support the above requirements. A step in addressing this has been taken by researchers at the University of Ulm who have compared SMIL[3, 4] and HyTime[5, 6] among other multimedia document models. They conclude that these have at least some of the specifications needed for content reuse, though not adaptability. To rectify these shortcomings, they propose a new standard (ZyX) though this does not seem to have been adopted elsewhere[7, 8]. They have used ZyX to create a training package for cardiology (Cardio-Op), but this is not a self-contained document, but rather relies on a database repository of multimedia content[9, 10].

An equally serious problem is the lack of suitable open source tools for authoring an interactive publication and for reading or using it. There are authoring and reading tools for “traditional” multimedia documents from Microsoft as well as freeware listed in sourceforge.net (for SMIL). Examples of freeware are: *ambulant* and *X-Smiles* (for reading SMIL documents) and *SMILgen* for authoring. Tools for HyTime were not found. It should be noted that these free software packages are neither easy to use nor well documented.

For authors to conveniently create their work as interactive publications, or for publishers to do so on the authors’ behalf, there must be authoring tools that are easy to use. The creator of such publications must be able to develop text in the conventional manner (as done with MS Word or Corel WordPerfect, for example), import all media types of interest, place them in the desired locations in the document, and intuitively provide links for navigation from one item to another, and allow access to tools for analysis and viewing. The marked lack in the literature for information on reliable open-source authoring and reading tools suggests an opportunity for developing and freely distributing such tools that are both open source and platform-independent. In the work reported here we use common document authoring tools in existing software environments, viz., Adobe Acrobat, Microsoft Word, Macromedia Flash and the HTML format.

The rest of this paper is organized as follows: Section 2 lists attributes of interactive publications that we consider important; Section 3 gives procedures to make tables and graphs dynamic, and to create interactive publications that include these as well as other multimedia objects; Section 4 provides a brief comparison of the four prototype documents created; and Section 5 describes the next steps in this ongoing research project.

2. DESIRABLE ATTRIBUTES

First, we list some desirable attributes of interactive publications. Note that this list is a subset of a larger collection of desired features and functions, many of which can be found in modern document browsers and authoring tools and are thus ignored in this discussion. Examples of the latter are features inherent to document creation such as authoring, styling, layout controls, table creation, and referencing, among others that are commonly found in present-day tools. In addition to platform-independence and ease of creation, we consider the following attributes necessary for an interactive publication.

- Appearance
  - Paginated view of the document should be similar to that of a traditional article, implying the availability of a large variety of fonts, weights, styles, paragraphing, multi-column formatting, etc.
- Page transitions
  - Traditional use of keyboard keys (page up/down) and mouse (scroll bar) should be possible.
  - Also desirable would be additional page forward and back links in the document.
- In-page navigation
  - Traditional use of keyboard keys (cursor up/down/left/right) and mouse, as well as additional use of control keys (as shortcuts).
- Image browsing

    - Images should be natively supported, especially all the commonly used formats such as JPEG, TIFF, DICOM, GIF, JP2, BMP, PNG, etc.
    - Should be easy to encode some degree of interaction with these into the document model.
- Navigating to an embedded / linked media object
    - Mouse-click (or keyboard) activation of audio, video and other objects.
    - Embedded or linked media objects should be able to invoke appropriate viewers or players.
- Native support for interactivity
    - The document model should provide native support for adding interactivity to tabular data, images and other multimedia data.
    - The document model should allow authors to define metadata needed to control interactivity with multimedia data, e.g., start-frame and end-frame numbers for video, row-column selections in a table, etc. These metadata could enhance the reader's interaction with the document.
    - Data in specialized and proprietary formats should be viewable using appropriate supporting application software.
- Execute code
    - The document model should support the inclusion of software programs, and their secure execution. This would enable peer reviewers, for instance, to test algorithms in the document. The programs could use input data from other multimedia components in the document or accept data provided by the reader, but would not write any data to the reader's computer.
- Transmission
    - The document model should support a reader-controlled order of transmission for data intensive multimedia-rich documents for convenient usage.
- Embedding and linking of multimedia/interactive objects
    - The document model should support both embedding and linking of multimedia and other interactive data such as dynamic tables or active images.
- Document integrity and structure
    - It is imperative that the document be self-contained. That is, the multimedia components should exist within the document (whether embedded in a file, or as files within a folder), and not simply exist in remote databases at, say, publishers' Web sites.
    - In both embedded and folder formats, the document model should support document integrity by closely linking the text document to the multimedia components.

We conclude that many of the desirable characteristics listed are found in present-day file formats or published standards or recommendations. For example, it is possible to embed multimedia components within the Microsoft Office or Adobe Acrobat frameworks. However, there are shortcomings: current file formats do not permit an instantiation of an interactive publication as envisioned above. The lack of a document framework that addresses all of the desired characteristics provides impetus for our research as well as next steps discussed in Section 5. In the next section we present our approach to the development of prototype documents and the design of a tool necessary for incorporating dynamic tables in an interactive publication.

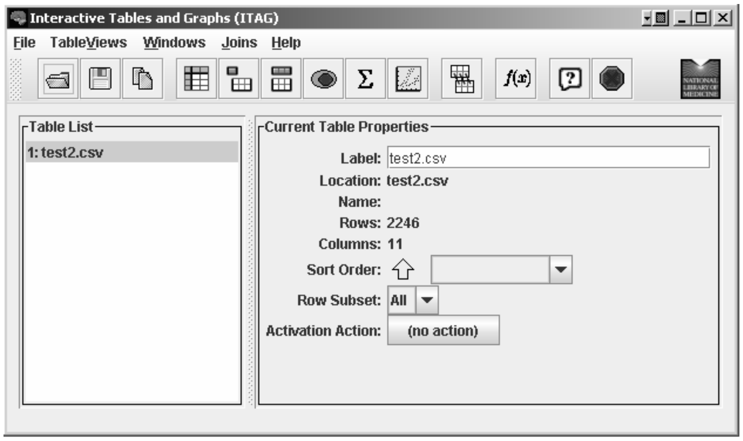
3. DEVELOPING AN INTERACTIVE PUBLICATION PROTOTYPE

Our approach to creating prototypes is to use common document authoring tools in existing software environments or formats, viz., Adobe Acrobat, Microsoft Word, Flash and HTML. Using these tools, documents were created in both the 'embedded' and 'folder' versions, and authoring procedures were recorded to possibly serve as a guideline for publishers and authors[11]. In these prototypes, we include a range of media types encountered in the biomedical literature, such as: paginated text; 3D animations of anatomic structures; a microscopy video of cell evolution; and clinical images such as CT, MRI and ultrasound in the DICOM format[12].

We also include large tables of data, graphs and charts. Enabling a reader to interact with, and analyze, tabular or graphical data in a document is necessary to fully realize the promise of an interactive publication (IP). Most biomedical journal articles present data as tables or graphs which usually show only a part of the data analyzed in the study/experiment on which the paper is based. Often this is due to size limitations (e.g., number of allowable rows) in certain proprietary though commonly used spreadsheet programs such as MS Excel. Furthermore, the presentation is static, i.e., the reader does not, as a rule, have a ready way to manipulate or analyze the data. In an interactive document, the reader should be able to view the same data as a table or a graph (convert from one to the other), sort columns in ascending or descending order, create new tables from subsets of the data in an existing table, calculate statistical/mathematical quantities from the data, zoom into the graphical data for more detail, and save the results of analysis in formats compatible with Excel or other common analysis tools. The challenge is to enable a reader to do any of these functions with a mouse-click or two. To enable a reader to interact with this information, we need to make the tabular or graphical data dynamic and then to incorporate this 'active' object into the IP.

To effectively incorporate large quantities of numeric data, and to make the tables and graphs dynamic for the reader, we created a tool, ITAG standing for *Interactive Tables and Graphs*, which we based on an open source software package called Starlink TOPCAT (http://www.star.bris.ac.uk/~mbt/topcat/), funded by the Particle Physics and Astronomy Research Council, UK and developed for the British astronomy community. Though TOPCAT has much of the desired functionality, we customized it by eliminating functions specific to astronomy, and extended it to generate line graphs and to include command line arguments to incorporate tables (or graphs) into an IP. This last modification allows an author to make a dynamic table or graph part of the document being created, and a reader to invoke them for viewing and analysis.

The ITAG software has the capability of generating and displaying interactive tables and graphs from raw data in formats such as *CSV* (comma separated values). The software can generate an interactive table from raw data and then display its plot. It can also create an interactive graph from raw data and then generate a table from it. Figure 1 shows a snapshot of the main control window of ITAG.



1: test2.csv

Label: test2.csv
Location: test2.csv
Name:
Rows: 2246
Columns: 11
Sort Order:
Row Subset: All
Activation Action: (no action)

Figure 1 – Main control window in ITAG.

The steps to make a table or graph dynamic and to include it in an IP are specific to both document format and operating system. We outline the steps needed to incorporate a table or graph into a PDF document, and the creation of a Table Invoker module. Since this batch/script file is necessary for the reader to interact with the table, it must be included in the document by the author.

The author starts with raw data in the CSV format. The next step is to select a document format, say Microsoft Word, and start writing the document using its text editor. The raw data can be represented in the document using a static table showing a selected subset, which can be created using a custom menu for drawing tables. By invoking ITAG, a reader would access the entire dataset in the table. An author wishing to show a pictorial representation of the raw data may use ITAG to create a graph and then use the ITAG menu to save it as a .gif file, which then may be inserted at the appropriate place in the document. Assuming that the final document is to be in PDF format, the Microsoft Word document should then be converted to PDF format using Acrobat. The static tables and graphs in the document should then be linked to the ITAG software and the corresponding raw data (CSV) file in order to make them interactive. This is done in one shot by creating a Table Invoker for each interactive table or graph. The following are the steps to create a Table Invoker:

1. Verify that the raw data is in proper CSV format as specified by ITAG.
2. Open a text editor, say Notepad, and type in the following text phrase:

   java -jar itag.jar -table -f csv <filename.csv>
   -table is the option specified by ITAG for generating interactive tables, and **–f** stands for format. [Interactive graphs could be created by using other options: **–graph, -x** and **–y**; A combination of **–table** and **–graph** may be used for generating both. <**filename.csv**> should be replaced by the corresponding CSV file (raw data) supporting the table or graph.]

3. Save the file in the format '**filename.bat**'. [It is recommended that the filename reflect the corresponding table name.]
4. Verify that the resulting saved file has a **.bat** extension and not a **.txt** extension. This file with **.bat** extension is called the Table Invoker.

ITAG has been implemented in Java and is thus platform independent, i.e., it can work under Linux, Macintosh, Microsoft Windows, or any flavor of Unix. However, a batch file equivalent on non-Microsoft Windows systems is a shell script. While no changes are necessary in the script itself, for Unix, Linux, or Macintosh (OS X) machines, the batch file must be set to be an executable, i.e., the command **chmod ugo+x filename.bat** needs to be executed. Further development is necessary in ITAG to enable this in an automated way.

Now that the Table Invokers for any tables and graphs have been created, we can proceed to creating an interactive publication. The procedure to create a folder-type PDF document using Acrobat is as follows:

1. Create and name a folder (this will be the document.) Start with the text as an ordinary MS Word document. Assemble all the supplemental material (bitmapped images, video, clinical data, tabular data). Place into the folder the text file, the supplemental files, the ITAG software and Table Invokers for each table or graph.

2. Convert the text document to PDF format using Adobe Acrobat, following the steps File>Create PDF>From File. Navigate to the file to be converted, then click Select. The result is a PDF document identical to the original Word document, including all its images.

3. Create bookmarks for the document (to allow readers to jump to any point in the document, much like an interactive table of contents) by selecting the Bookmarks tab that appears to the left of the document. Move the document to the position for the bookmark, and select Options>New Bookmark. In the highlighted area, enter the name of the new bookmark.

4. Use the Button tool in Adobe Acrobat to define interactive regions of the document. (Typical interactive regions are tables that when clicked will open ITAG, or images that will open a video when clicked.)

5. On defining an interactive region a dialogue box opens that permits changing the appearance of the Button, as well as altering action settings. For appearance, set the Border and Fill to No Color, so that no additional graphical elements clutter the document when printing.

6. Next, select the Actions tab. In the Add an Action box, choose Mouse Up for the Trigger, and Open a File for the Action, then click Add. When a dialogue box opens, navigate to the File Location, select the file (video, DICOM image, etc.), then click Select. (This will ensure that when the interactive area is clicked, the selected file will be launched.)

7. To assist the reader, place links on first and last page to software needed to view or use the multimedia objects in the interactive publication, e.g., QuickTime, Windows Media Player, DICOM Works.

4. COMPARISON OF PROTOTYPES

The prototype documents created were compared against the attributes listed in Section 2. With respect to appearance, PDF and Word display paginated views typical of traditional documents. HTML would not typically, though the use of cascaded style sheets could allow this. Flash could provide this also, although significant programming effort would be required on the part of the author. Page transitions and navigation within pages in a Flash document would similarly require considerable programming for appropriate response to key strokes, while PDF, Word and HTML documents allow traditional use of keys and mouse for such control.

Multimedia and tabular data may be embedded or linked in Word and PDF since both support object linking and embedding, though Word requires the supporting OLE application. Again, Flash requires special programming to achieve this, and HTML supports limited embedding but does allow linking. All four platforms supported linked objects but failed in various degrees when tested with embedded objects.

With respect to interactivity with tabular data, none of the formats provide native support for this, though Word allows interactivity with Excel from its Office suite of applications. None of the formats provide native support for interactivity with DICOM images. Moreover, these formats do not support elegant and user-sensitive downloading, a challenge for data intensive multimedia-rich publications.

We compared the four formats subjectively from the point of view of authoring, reading and required systems and software, as shown in Table 1.

	PDF	MS Word	HTML	Flash
Authoring				
Ease of creation	Easy	Easy	Moderate	Difficult
Specialized skills required	No	No	Yes	Yes
Skill level needed	Low	Low	Moderate	High
Reading				
Ease of use	Easy	Moderate	Easy	Easy
Speed moving from one media type to another	Same across all formats			
Speed for invoking media types	Same across all formats			
Need special client tools for viewing	Yes	Yes	Yes	Yes
Need special client tools for manipulating	Yes	Yes	Yes	Yes
System/Software				
Cross Platform	Yes	No	Yes	Yes
Open source	No	No	Yes	No

Table 1 – Comparison of four document prototypes.

5. FUTURE WORK

While prototype documents have been developed and demonstrated, further research in enabling interactivity in scientific publications requires a comprehensive approach which includes decisions in the choice of the appropriate document format, tool development for authoring, downloading, and viewing, and a thorough evaluation of the technology at various stages to ensure that the participating groups, viz., authors, publishers, researchers, and archivists benefit from the ubiquity of this type of publication. These topics are reflected in our next steps.

5.1 Enhancing ITAG

The current version of ITAG, though useful, needs to be enhanced to allow greater user interaction with the data. These enhancements include the ability to parse, display, and interact with complex (hierarchical) tables, enriched charting techniques, greater interoperability with popular statistical analysis software, ability to install ITAG as a client application, automatic startup of ITAG software without the use of authored scripts, and not requiring separate files for data subsets. Next steps on this topic are guided by these desirable features.

5.2 Authoring and formatting tools

At the present state of development, the authoring tools proposed are primarily popular editors for document development accompanied by step by step procedures as instructions for the author. The author introduces interactivity for each media object (movie, DICOM image, table, etc.) by following a series of steps and selections, and has to keep almost every use situation in mind. This process can be tedious, repetitive, and possibly exasperating. Easy to use tools with wizard-like interfaces would be necessary to encourage widespread use.

Additionally, for most linked media objects, the interactivity is limited to the format in which the document was authored. A change of that format to another one for purposes of dissemination, for example, would not guarantee the interactivity originally designed. The author would need to recreate the interactivity in the target file format. A typical example is often seen in use today: most articles are created using MS Word or WordPerfect, but are often placed on publishers' Web sites as PDF documents. The problem is that the interactivity introduced in the originals may not be retained automatically after format conversion to PDF. The solution is for tools that analyze source and target formats, correcting formatting problems as needed.

5.3 Efficient downloading

Whether the interactive publication is of the embedded or folder type, its size can be very large due to data-intensive media objects such as video, DICOM images or tables with tens of thousands of rows. Since the very concept of an IP is rooted in encouraging the inclusion of as much research data as necessary to promote understanding, one can conceive of documents ranging in size from tens to hundreds of megabytes. Such large sizes can pose a serious barrier to widespread dissemination and use.

Our approach is to design an intelligent Download Manager utility to download the textual portion of the document first (allowing the reader to start perusing the text), while the data-intensive media objects arrive in the background. Moreover, the order in which these media are downloaded should be controllable by the reader based on his/her interest. For example, if a reader wishes to forego the introduction or methods sections typically found in research articles in favor of a dynamic table further down in the article, the Download Manager should be able to deliver the table first.

5.4 Evaluation

In order to understand the value of the interactive publication, its shortcomings in its current form and practical directions for the future, a comprehensive evaluation needs to be conducted in collaboration with significant players in this enterprise. These would include publishers, authors and readers.

First, we intend to recruit one or more biomedical publishers with whose help we will establish a plan containing evaluation criteria for interactive publications: e.g., improvement of comprehension, learning, and degree of assessment of research reported in the publication. Next, we will provide our tools and procedures as they are developed to the publishers, their contributing authors and designated peer-reviewers, and train them in using our tools. In the course of authoring the publications, our collaborators will be expected to evaluate the tools and procedures provided on the following grounds: ease of creation, whether specialized skills are required and the skill level necessary. Reading the publications will be evaluated on the following: ease of use, speed of moving from one media type to another in the publication, speed in invoking the various media objects, and whether viewing and manipulating the objects require additional client tools. The results of this evaluation will inform the next stages of tool development.

5.5 Preservation

The long term preservation of all significant material in biomedicine is a mandated task for the NLM, irrespective of the media or formats they come in. This will be the case for interactive publications as well. At present we are engaged in the design of systems for archiving scanned images and Web resources from various collections[13, 14]. This activity will be expanded to investigate factors relevant to the preservation of interactive publications, e.g., design of suitable archival systems, extraction of descriptive and technical metadata, and bulk migration of file formats.

ACKNOWLEDGMENTS

We gratefully acknowledge receiving CT, MRI and ultrasound images in the DICOM format from Eliot Siegel, M.D. of the VA Maryland Healthcare System, Baltimore, MD, and a cell development video from Dr. Alexey Khodjakov of the Wadsworth Center, Albany, NY. These media were used in the development of our prototype interactive publications. We also thank Dr. Mark Taylor of University of Bristol, U.K., and Timothy E. Thate (Emerging Leader Fellow, U.S. Dept. HHS) for discussions useful in our development of the ITAG tool. This research was supported by the Intramural Research Program of the National Library of Medicine, National Institutes of Health.

REFERENCES

1. D.A.B. Lindberg, B.L. Humphreys, *2015 – The future of medical libraries*, N. Engl. J. Med. (2005) 352; 11:1067-70.
2. D.A.B. Lindberg, *Research opportunities and challenges in 2005*, Methods Inf Med 2005; 44:483-6.
3. Synchronized Multimedia Integration Language (SMIL) 1.0 Specification. http://www.w3.org/TR/REC-smil/
4. P.N.M. Sampaio, J.P. Courtiat. *Providing consistent SMIL 2.0 documents*, Proc. IEEE International Conference on Multimedia and Expo, ICME 02, August 2002, Vol. 2, 337-40.
5. Hypermedia/Time-based Structuring Language (HyTime): ISO/IEC 10744. http://www.y12.doe.gov/sgml/wg8/document/n1920/html/n1920.html
6. J.F. Buford, L. Rutledge, J.L.Rutledge, *Integrating object-oriented scripting languages with HyTime*, Proc. International Conference on Multimedia Computing and Systems, May 1994, 425-34.
7. S. Boll, W. Klas, U. Westermann, (1999) *A comparison of multimedia document models concerning advanced requirements*, Technical Report – Ulmer Informatik-Berichte. No. 99-01, Dept. of Computer Science, University of Ulm, Germany.
8. S. Boll, W. Klas, *ZyX – A multimedia document model for reuse and adaptation of multimedia content*, IEEE Trans. Knowledge and Data Engineering, Vol. 13, No. 3, May/June 2001.
9. R. Friedl, M. Preisack, M. Schefer, W. Klas, J. Tremper, T. Rose, J. Bay, J. Albers, P. Engels, P. Guilliard, C.F. Vahl, A. Hannekum, *CardioOp: an integrated approach to teleteaching in cardiac surgery*, Stud Health Technol Inform. 2000; 70:76-82.
10. R. Friedl, W. Klas, U. Westermann, T. Rose, J. Tremper, S. Stracke, O. Godje, A. Hannekum, M.B. Preisack, *The CardioOP-Data Class (CDC). Development and application of a thesaurus for content management and multi-user teleteaching in cardiac surgery*, Methods Inf Med. 2003; 42(1):68-78.
11. G.R. Thoma, S. Antani, G. Ford, M. Chung, K. Vasudevan. *Interactive Publications Research -- A Report to the Board of Scientific Counselors,* Technical Report LHNCBC-TR-2005-005, September 2005. http://archive.nlm.nih.gov/pubs/thoma/tr2005005.pdf
12. Digital Imaging and Communications in Medicine (DICOM). http://www.psyc.nott.ac.uk/staff/cr1/dicom.html
13. G.R. Thoma, S. Mao, D. Misra. *Automated metadata extraction to preserve the digital contents of biomedical collections*, Proc. 5th IASTED International Conference on Visualization, Imaging and Image Processing (VIIP 2005). September 2005. Benidorm, Spain; 214-19.
14. S. Mao, D. Misra, J. Seamans, G.R. Thoma. *Design strategies for a prototype electronic preservation system for biomedical documents*, Proc. IS&T Archiving 2005 Conference, Washington DC; April 2005; 48-53.