



An Augmented Wiki for Interactive Scientific Visualization and Evolutionary Collaboration

Francis T. Marchese, Pace University, *Computer Science Department, 163 William Street, 2nd Floor, New York, NY 10038*, URL: <http://csis.pace.edu/~marchese>, Email: fmarchese@pace.edu

Abstract

Wikis may be used to expand the scientific research process beyond the traditional publication of a communication within a journal. It will be shown that the tools used during the data analysis phase of scientific research may be integrated into a wiki for interactive visualization. This approach binds together two disparate segments of the scientific visualization process – the visualization of complex experimental data with the publication of results and analysis. The ability to add new insights beyond the original research means that the document can evolve over time and be a focal point for collaborative extension of the scientific narrative.

Introduction

Journals are the *de facto* medium for communicating results of scientific research. Yet, they exhibit a number of limitations. At their core resides the scientific paper, a concise text-based narrative that is responsible for the bulk of communication. Imagery is relegated the subservient role of illustrating the essential points set forth in the report. Within this context, readers must contend with static imagery and charts scattered about the text. Images are typically small, of excessive information density, and often poorly positioned within a page layout. The result is a minimization of the linkage between image and narrative. Moreover, many illustrations are snapshots from complex visualizations over multiple spatial and temporal dimensions that, by their nature, are incomplete representations of the systems studied. In contrast, the interactive visualization software that created this imagery may have been run on multi-screen displays or on special output devices such as video walls [NI061], 3D stereographic systems [MAR02], or immersive environments such as CAVES [NEI93] that provide a level of visual description and detail not found in print media. Conversely, scientific visualization systems are not designed to accommodate, annotate, and format the large body of descriptive text required for a complete document. Moreover, journal articles themselves are closed, unable to be annotated or appended by other scientists. This is unfortunate, given the fact that a publication is one milestone out of many in an open-ended asynchronous collaborative research process that is science. And journals are the major repository of a field's group memory or knowledge!

This paper attempts to address these issues of scope, interaction, and collaboration by creating a prototype asynchronous collaborative visualization system as a mash-up between two technologies – a wiki and an interactive scientific visualization program. In the next section we discuss the problem of binding a journal article to visualization. A discussion on building collaboration will follow, ending with a final discussion.

Binding

The purpose of this section is to discuss some ways in which a better connection may be made between journal article and visualization software. We have selected molecular visualization as the application domain for our discussion because molecular visualization environments typically integrate data generation, visualization, and analysis [MAR03]. In addition, these systems must render a vast array of graphical representations not found in mainstream visualization systems in order to support visual analysis of data generated for molecular structure and dynamics by both theory and experiment.

An Augmented Wiki for Interactive Scientific Visualization and Evolutionary Collaboration

There are a number of approaches for binding a scientific publication with an interactive visualization system. The simplest, most direct, and standard practice followed by many scientific journals is to provide source data as supplementary material. The scientist can download the data, read it into his or her visualization program, and follow the paper's commentary. Since the scientist who wrote the paper and the scientist who reads the paper are assumed to be domain experts and use the same or similar visualization systems, it should be relatively easy for the reader to translate the paper's descriptive text into interactive visualization.

The major problem with this approach is that *raw* data file formats typically allow only sparse annotation. And the comments are only related to the data itself (e.g. the authors, data history, experimental conditions, data collection methods, etc.); nothing specifying how the images were created that are contained in the paper. As a result, all information about the visualization resides in the publication.

An early approach to delivering augmented data was pioneered by the journal *Protein Science* in 1992 when it distributed annotated source data files called kinemages (kintec images). Kinemages are ASCII files that contain source data, annotated display lists for visualization scenarios, and accompanying descriptive comments. They were designed by David and Jane Richardson of Duke University as part of their interactive molecular visualization system called Kinemage [RIC92]. Kinemages are created by a scientist author with a helper program that embeds script commands for constructing views of the images displayed, and extended comments or annotations. These files are then made freely available at the journal's website. A scientist then downloads a kinemage into a local copy of the Kinemage molecular viewer, selects a script for each three-dimensional view, manipulates it, and reads the associated documentation.

Most recently, a number of publishers have added direct interactive visualization support for data files. This is currently done for three-dimensional biomolecular structure data at the American Chemical Society's ACS Chemical Biology, the Royal Society of Chemistry's Molecular Biosystems [MOL06], and Nature Structural and Molecular Biology [EDI06]. These sites use a program called FirstGlance in Jmol [FIR07] that is a user interface to the Jmol molecular visualization applet [HER06]. First_Glance works within a web browser to enable readers of scientific journals to visualize the main features of newly published 3D models. It parses data files to extract out their essential structural features and produces a web page that contains the embedded Jmol applet and a set of buttons that control the display of a standard collection of biomolecular representations that flesh out the contents of the data file.

Jmol is the essential visualization component of these webpages. It is a free, open-source molecule viewer for chemistry and biochemistry written in Java. Jmol can read a wide variety of molecular structure files, and allows users to directly manipulate the three dimensional molecular data with mouse control over geometric transformations such as rotation, scaling, and translation; and change the rendering style for all or part of the molecule. Besides its extensive visualization capabilities, Jmol has two attributes that are essential to First_Glance's operation and its use in linking source data to text: it can read command scripts that control how the data is displayed; and it uses LiveConnect [LIV07], the API that furnishes JavaScript with the ability to call methods of Java classes using the existing Java infrastructure. Hence, Javascript provides the means to connect Jmol's functionality with user interface objects (e.g. buttons, check boxes, etc.) that are embedded within HTML. With a button click, a script command is sent to the applet that in turn updates the graphical representation.

Demonstrations of Jmol's power as an embedded visualization tool and a means for tightly binding textual communication with visualization may be seen in the HTML pages created by the students of the Biology Department of Kenyon College [BIO07]. The web page displayed in Figure 1 is one such example. It presents the structure of HIV-1 reverse transcriptase complexed with FAB-28 monoclonal antibody fragments. The HTML page is segmented into two frames: the left frame containing an embedded link to a chemical data file that is read by Jmol, and the right frame containing the descriptive text. When the web page is initially loaded into the browser, text and graphical descriptions of the data are synchronized. The viewer may change the visual representations at any time through Jmol's pop-up menu or moving the mouse over the image; but, in so doing, the dynamic image may no longer accurately reflect the contents of the static text. Jmol can correct for these changes with the use of its scripting capabilities. By clicking on embedded buttons within the text (c.f. Figure 1), Jmol receives a sequence of commands via its Javascript interface and executes the Java methods associated with those commands to change structural representations and invoke animation commands. And in this web page, these buttons are used to synchronize imagery to text by invoking scripts that actively display what the text is communicating.

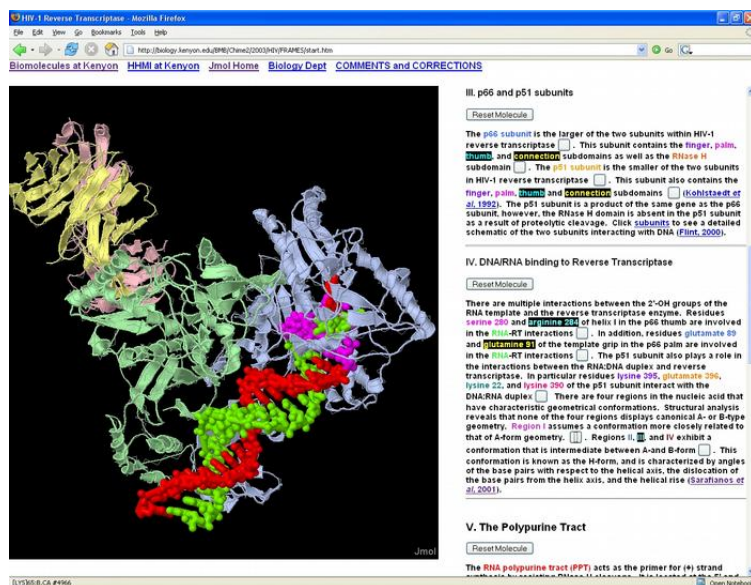


Figure 1. Jmol in webpage with embedded buttons that launch Jmol scripts.

Collaboration

Although rich in content, the static web pages discussed above are difficult to annotate and share. A natural approach is to use Wiki website technology that supports the addition and editing of hypertext. As such, a wiki sustains collaborative authoring of dynamic hypertext documents that can evolve over time. In so doing, a wiki can be used to extend the scientific publishing and communication paradigm by opening up a closed scientific paper to augmentation by the scientific community as new data becomes available or new insights are realized.

There are hundreds of wiki engines from which to choose that have been written in a wide variety of programming and scripting languages [WIK07]. JSPWiki [JSP07], a WikiWiki engine built around standard J2EE components (Java, servlets, JSP), was selected for this work. It is a mature and stable Wiki with simple formatting rules, installs easily, provides a simple authentication mechanism, supports RCS-based version control, allows file attachments to a page, and can be extended using plug-ins.

In order to create the collaborative scientific narrative, the Jmol applet must be inserted into JSPWiki. There are a number of ways of doing so. Creating a plugin is one possibility. Another is to allow the wiki to execute HTML code including Javascript. Since Jmol uses Javascript embedded within HTML this option becomes the path of least resistance. However, this method is not without its dangers, given that the insertion of insecure Javascript could lead to system corruption. That said, we will use this option for the purpose of demonstration.

Figure 2 displays the Jmol/JSPWiki combination, the starting point for collaboration. A scientist has created a wiki entry containing and descriptive text with embedded Jmol commands linked to interface buttons. It was inserted into a new wiki page by means of a copy and paste. This was followed by the uploading of a file containing the molecular coordinates as an attachment to the wiki page. A two column HTML table was used to arrange Jmol and the body of text within an HTML in-line frame. This approach mimics the arrangement found in Figure 1. Other schemes for partitioning the document space are possible, such as using the HTML div tag. But the essential issue is that the page should be organized in such a way as to maintain the close proximity between narrative and visualization so as the text is read, the graphical representations may be manipulated.

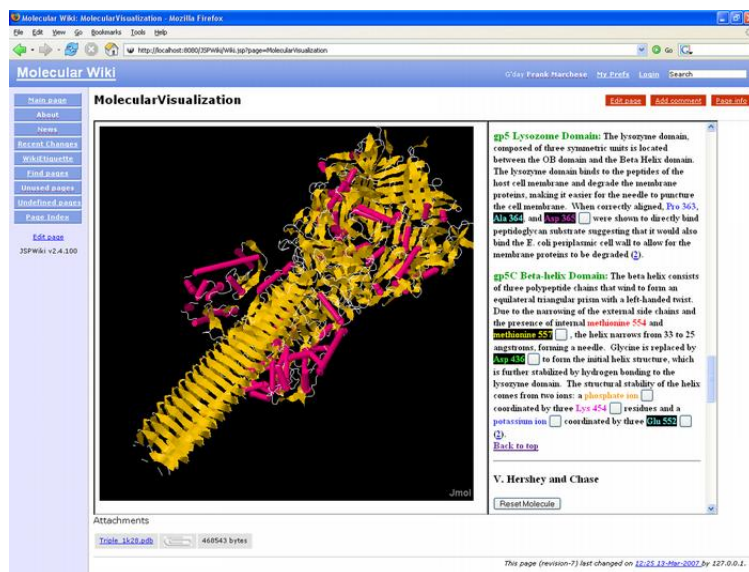


Figure 2. Jmol in with enhanced text embedded within wiki.

At a later time, a collaborator explores the wiki page by using the embedded Jmol scripts or using Jmol's interactive pop-up console to input script commands on-the-fly. When done, the collaborator may either enter a comment, edit the document to include the new scripts, or download the molecule structure file to refine the data analysis on a local machine. After analysis has been completed, a new commentary may be uploaded to the wiki. Over time each scientist can review the document's history by using the wiki's version control system. Specifically, it can display how the document has evolved as each successive change and addition has been made. Each version in the history can be loaded and explored at any time.

Conclusions

In this paper we have discussed a number of issues that arise with respect to the use of documents as mediators in the scientific research process. The first issue is the creation of scientific publications as hypermedia with embedded interactive visualization software. A scientific paper formatted in this way is far superior to a document delivered currently by publishers because this format makes a connection between the visualization software from the analysis phase of the research process with that of the communication phase. In so doing, the complex static images that are normally rendered in a publication are mapped back into the interactive graphical states as originally viewed by the document's authors. In addition, if the visualization software runs on special display hardware such as video walls and 3D stereographic systems, then these systems become integrated into the document, enriching the presentation. Finally, the imagery can be synchronized with the descriptive text. This synchronization adds a complementary redundancy to the scientific narrative so that not only is the information described in detail but also is shown in detail [MAR04]. This is particularly important where complex data sets are analyzed.

The second issue is document delivery. It is possible for scientific journals to implement what we have described in concept. Many publishers (e.g. the American Chemical Society) deliver journals as either PDF or HTML files. The latter format already contains rich content with links to images and other documents. Hence, embedding a visualization applet becomes a real possibility. In addition, authoring tools for creating these documents can be developed. For example, two such tools have recently appeared for preparing Jmol-based webpages: SchiSM [CAM07] and Jmol_Web_Page_Maker [JMO07]. SchiSM is a web-based client-server program that combines HTML forms as a user interface with pre-defined authoring templates to construct an HTML document. A user enters Jmol scripts and associated text. When done, the server displays the result and delivers the document as a compressed file ready for downloading and incorporation into a website as shown in Figure 3. The results are reasonable, and may be enhanced with additional editing.

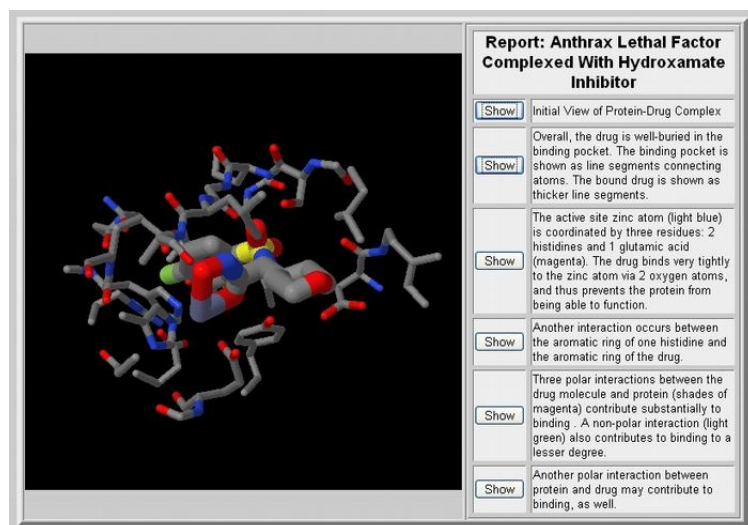


Figure 3. Sample output by SchiSM template.

Publishers could advance the authoring of these kinds of documents by providing templates for interactive webpage construction. This would foster the creation of more meaningful scientific commentary. Since scientific narrative is formally structured, it should be possible to generalize the kinds of templates that already exist and adapt them to the narrative form. Many publishers currently require templates to be used by technical authors. For example, Springer-Verlag has a long history of designing and using TeX document templates for monographs, multi-authored works, lecture notes, and conference proceedings; and computer scientists are required to author their documents employing ACM and IEEE publication templates for publication in conference proceedings.

Third, we discussed how these interactive documents may be placed within a wiki to support collaborative asynchronous visualization in which multiple authors contribute to a document's contents allowing it to evolve over time, thus expanding the scientific publishing and communication paradigm. In addition, these wiki systems support emergent collaboration [MAI94]. Emergent collaboration occurs when individuals who have no formal connection to the original document create modifications that are integrated into the document. This is an empowering notion that opens up the research process so that the document can grow as other individuals become interested in the research, possibly taking it far beyond its original scope.

Fourth, from a technical perspective, the ability to easily insert a graphical application into a wiki in order to create a collaborative visualization system means that the need for specialized software designed to support collaborative visualization is mitigated, because the wiki becomes the basis for sharing. With a wiki as foundation, it then becomes easier to construct more complex collaborative visualization systems. For example, this system could have been extended by adding an interactive charting applet to display xy graphs and statistical plots - complementary data representations that support scientific research.

Finally, because wiki technology is open, and it may be modified either outright or adapted to accommodate new features, there are opportunities for further active development of collaborative visualization and publication systems. Specifically, wikis may be enhanced to better accommodate applets, and applets may be extended to better accommodate hypermedia communication. Development in these areas will be welcome.

Bibliography

- [BIO07] Biomolecules at Kenyon, [cited Mar. 30, 2007], available from World Wide Web: <<http://biology.kenyon.edu/BMB/chime.htm>>.
- [CAM07] S. Cammer, SChISM2: creating interactive web page annotations of molecular structure models using Jmol, *Bioinformatics* 23, 3 (Feb. 2007), pp. 383-384.
- [EDI06] Editorial, We're living in a 3D world, *Nat Struct Mol Biol.* 13, 2 (Feb 2006), p. 93.

An Augmented Wiki for Interactive Scientific Visualization and Evolutionary Collaboration

- [FIR07] First Glance in Jmol, [cited Mar. 30, 2007], available from World Wide Web: <<http://molvis.sdsc.edu/fgij/index.htm>>.
- [HER06] A. Herráez, "Biomolecules in the computer: Jmol to the rescue, *Biochemistry and Molecular Biology Education* 34, 4, 2006, pp. 255-261.
- [JMO07] Jmol_Web_Page_Maker, [cited Mar. 30, 2007], available from World Wide Web: <http://www.uwosh.edu/faculty_staff/gutow/Jmol_Web_Page_Maker/Jmol_Web_Page_Maker.shtml> .
- [JSP07] JSPWiki, [cited Mar. 10, 2007], available from World Wide Web: < <http://jspwiki.org/wiki/>>.
- [LIV07] LiveConnect, [cited Mar. 30, 2007], available from World Wide Web: <<http://devedgetemp.mozilla.org/library/manuals/2000/javascript/1.3/guide/lc.html>> .
- [MAI94] C. Maioli , S. Sola, F. Vitali, The support for emergence of collaboration in a hypertext document system, In *ACM CSCW'94 Workshop on Collaborative Hypermedia Systems*, Chapel Hill (NC), GMD Studien n. 239, ACM, 1994.
- [MAR02] F.T. Marchese, A stereographic table for biomolecular visualization, In *Proceedings of the Sixth International Conference on Information Visualization: IV'02* (London, July 2002), IEEE Press, 2002, pp. 603-607.
- [MAR03] F.T. Marchese, J. Mercado, and Y. Pan, Adapting single-user visualization software for collaborative use, In *Proceedings of the Seventh International Conference on Information Visualization: IV'03* (London, July 2003), IEEE Press, 2003, pp.252-257.
- [MAR04] F.T. Marchese, Dynamically binding image to text for information communication, In *Proceedings of the Eighth International Conference on Information Visualization: IV'04* (London, July 2004), IEEE Press, pp. 707-712.
- [MOL06] Molecular Biosystems News, 3D View adds a new dimension, *Molecular Biosystems*, (November 21, 2006).
- [NEI93] C. Cruz-Neira, D.J. Sandin, T.A. DeFanti, Surround-screen projection-based virtual reality: the design and implementation of the CAVE, *Computer Graphics* (Proceedings of SIGGRAPH '93), ACM SIGGRAPH, (August 1993), pp. 135-142.
- [NI06] T. Ni , G.S. Schmidt, O.G. Staadt, M.A. Livingston, R. Ball, and R.A. May, II, A survey on large high-resolution display technologies, techniques, and applications. In *IEEE Virtual Reality Conference 2006*, IEEE Computer Society, Washington, DC, 2006, pp. 223-236.
- [RIC92] D.C. Richardson and J.S. Richardson, The kinemage: a tool for scientific communication, *Protein Science* 1, 1 (Jan. 1992), pp. 3-9.
- [WIK07] Wiki Engines, [cited Mar. 10, 2007], available from World Wide Web: <<http://c2.com/cgi/wiki?WikiEngines/>> .