

A Xanalogical Collaborative Editing Environment

Angelo Di Iorio, Fabio Vitali
Department of Computer Science
University of Bologna (Italy)
E-mail: fabio@cs.unibo.it

Abstract

XanaWord is part of an ongoing research project at the University of Bologna aimed at creating a shared editing environment for the personal customization of the WWW. Its main objectives are the shielding of users from details about technologies, data formats, and access rights and privileges.

In particular, XanaWord allows users to create personal variants of web pages and edit them through a popular tool such as MS Word. The system manages versioning, change tracking and universal editing access in a manner rather similar to Xanadu, the pioneering hypertext project of Ted Nelson.

1 Introduction

Ever since the term 'hypertext' was invented by Ted Nelson, in the mid Sixties, researchers and computing professionals have been discussing which hypertextual functionalities to embed in information systems. Particularly important has always been the distinction between the roles of author and reader in the use of hypertext artifacts.

Indeed, in 1990, Conklin's seminal paper [6] singled out three different approaches in this regard: hypertext systems as idea-collectors, as browsing systems or as publishing/editing macro-systems. A famous example of the first approach is HyperCard [9]: the hypertext document is used by individuals or small groups to store, organize, and read their own information, and the tool suggests a complete overlap between authors and readers. The second approach heavily differentiates between authoring and browsing: for instance, in the case of on-site museum displays implemented with HyperTies [17], authors wrote content using authoring tools different and separated from browsing tools used by the actual museum visitors. Finally, by hypertext macro-system we mean an universal environment in which everyone can access, read, re-use, modify and comment any material of other users, tailoring it to his/her own purpose. Basically no such system ever existed, and the whole category was invented to describe Ted Nelson's Xanadu [15], an ambitious idea which unfortunately never came to exist.

Although Berners-Lee, the originator of the World Wide Web, was influenced by Nelson's writings, and although his first WWW client was indeed an editor, the World Wide Web rather early turned into a browsing system, where the tools for browsing are free, easy to use and rather limited, while the tools for creating pages and web

sites are expensive development environments mainly aimed at the professional market, extremely powerful but requiring sophisticated technical expertise for their use.

Today the World Wide Web is a powerful means to publish information and provide services but it lacks several hypertext functionalities [2]; in particular, a clear distinction exists today between readers and authors, whereby readers can only choose reading paths explicitly provided by the authors and cannot create new content, create new links or create personal variants of web pages during browsing.

There are a few exceptions to these problems that are gaining some support in recent times. For instance, weblogs [3] are tools for fast editing (mostly based on web forms) and publishing of personal diaries, addressed to individuals and small communities. Wikis [12] are collaborative tools for shared writing and browsing, allowing every reader to access and edit any page of the site, through simple web forms and a very intuitive text-based syntax for special typographical effects. Wikis come rather close to Nelson's initial idea of a global publishing medium open to customization and individual contributions: they are characterized by simple interfaces, raw layouts and an open editing philosophy ("*An open door is the best lock*" [20]). An ill-intentioned anonymous reader can in fact modify or even delete previous contents, but thanks to the internal revision tracking and differencing mechanisms, an army of self-appointed "wiki gnomes" can rapidly rebuild the damaged documents, making any destructions look futile and short-lived.

Wikis are limited in that the overall esthetics of the created pages is inherently meagre, that users have to learn yet another text-based editing language, which, although rather easy, is necessary anyhow, and that it only allows editing on a restricted pool of resources, i.e. on the pages of the wiki site.

At the University of Bologna we recently started a long-term and ambitious research effort whose goal is to develop a web editing environment that allows anyone, anywhere, anytime, to access and edit any web resource in an easy and intuitive way. Its main approach is to shield the user as much as possible from issues such as data formats, specifics of interfaces, and deployment and right managements.

A first result of this project is ISA, the Immediate Site Activator [18], which aims at allowing users to create graphically-sophisticated web sites using commercial applications: the idea behind ISA is that users may already know fairly well commercial applications such as

Photoshop or MS Word, and exploits the features of these tools to create sophisticated dynamic sites.

XanaWord, which we will describe in this paper, is yet another intermediate result of our project. XanaWord is a collaborative editing environment for the World Wide Web, that allows users to edit pages during browsing. The project has been driven by two main goals: permitting the overlap of author and reader roles on working with WWW pages and providing a tool that is easy to use by technically unsophisticated users.

XanaWord's main approach is to integrate MS Internet Explorer as the navigation browser and MS Word as the content editor. Since changing context from browsing to editing is completely automatic and transparent, there is no technical requirements on user when editing the Web pages. The name XanaWord reflects this feature: a tool providing typical functionalities of Xanadu, while using MS Word as editor. This paper explores a number of proposals deriving (directly or indirectly) from Xanadu's original project, and discusses XanaWord's own workings in this light.

2 Related works

Xanadu is probably the most influential *vaporware* of the computing history. A system never quite finished, never published, that only a few people had the chance to see running; later rendered obsolete by the birth of the World Wide Web, Xanadu is better known through *Literary machines* [15], and a few papers in specialized magazines..

Some keywords can be used to briefly describe Xanadu's vision. First of all the term *xanalogical storage*. Xanadu document is not stored as a whole block on a file system, but as a list of references to fragments combined in the final document on-demand.

This mechanism has a lot of advantages, mainly the *transclusions* [14]. "Transclusions" are the reuse in whole or in part of content from another document. This is different from pure copying, in that the document stores only a reference to the external material. The software is expected to fetch this content and place it inline with the main material. Transclusions implement complete change tracking and versioning of documents: a document in Xanadu is not a static resource but a container in continuous development, whereby each successive version refers to the content of previous ones through transclusions. Finally, Xanadu was meant to provide a powerful linking mechanism allowing users to connect every part of every document to any other document, and to store this information anywhere. Through these mechanism Xanadu was meant to become a global collaborative and pluralistic hypertext publishing system. Nelson's vision has in time inspired much hypertext research: while some researchers have continued to develop the original code [13], others have tried to enhance existing systems (firstly WWW) with Xanadu functionalities [10][16].

Other research efforts, not directly influenced by Xanadu, has been made to follow the original vision of the Web as an universal collaborative environment. The most important is certainly WebDAV [8], an HTTP extension proposed by IETF that allows users to collaboratively edit and manage files on remote web servers, supporting resources locking and versioning. The W3C project called Amaya [1] is a Web browser/editor that allow to create and update documents directly on the Web during browsing.

Finally we would like to mention other systems and protocols that allow personal intervention on external materials in the WWW. These include annotation systems such as CritLink [11] that allows users to comment every WWW page, adding annotations (through forms) inserted on-the-fly into the original document through a non-transparent HTTP proxy. In the same way, XLink-Proxy [4], exploiting the new W3c standard Xlink[7] and the use of an external linkbase, allows every user to add links from and to every Web page.

All these examples testify the efforts by researchers and professionals to move World Wide Web from a simple browsing system to a complete authoring and publishing system, to enhance, in a word, the set of hypertext functionalities [2] offered by the World Wide Web.

3 XanaWord

Some ideas behind XanaWord are based on a previous project called the Intermediate Site Activator (ISA) [18].

3.1 The ISA project

ISA is a novel approach to web page production, providing an intermediate solution between the advanced effects of professional tools such as Dreamweaver or Vignette, powerful but difficult to learn, and the simple but ugly results obtained when exporting HTML from a word processor.

The main idea of ISA is to exploit standard desktop tools for the creation of content and layout, and employ a server-side application for the delivery of the final web pages.

In the scenario of producing a web site with ISA, shown in fig. 1, a graphic designer creates the overall graphical aspect of the page using a desktop tool such as Fireworks or Photoshop. She will then use the slicing tool of the application (many graphical packages implement it) to draw the active areas (the zones where the ISA package will act), and specify their properties. After the layout has been created, the content producer (having no knowledge of HTML or any other markup language) can proceed to write the content documents. He will use either an HTML editor, or, more frequently, a word processor. ISA knows and handles MS Word files with ease. The content producer will thus create any number of Word files, using styles as instructed by the layout designer, and saving them on the site as Web documents.

By specifying the corresponding URLs in a browser

window, ISA will merge the layout and the content document to form a complete web page. In traditional web site generation there is the additional step of the web technologist joining the layout and content into the final

web page. ISA provides the services usually supplied by this intermediary. XanaWord is strictly connected with ISA because all documents created in this environment are handled with ISA techniques.

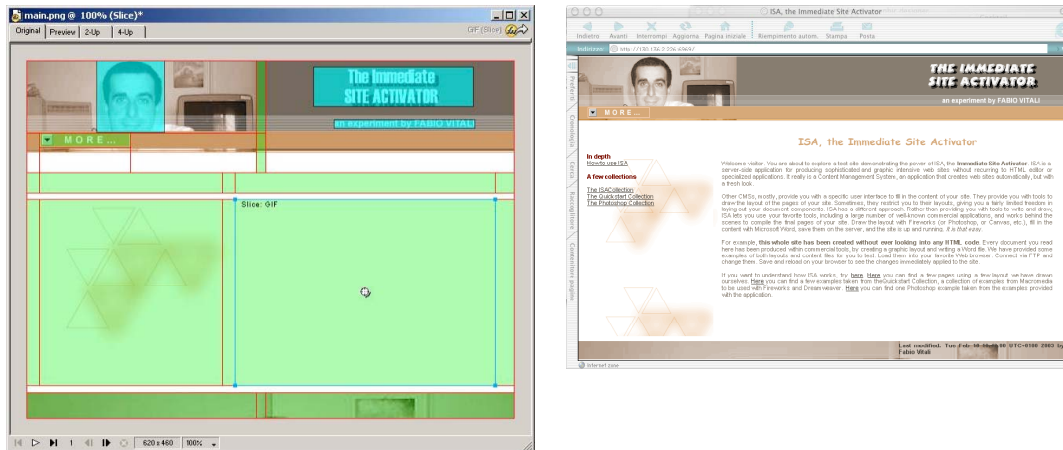


Figure 1: Drawing and slicing a layout and the final Web page created by ISA

3.2 Xanaword

While ISA still relies on web pages to be accessible through a web site under the control of the authors, XanaWord is meant to provide web customizability for all web pages, regardless of their authors and access parameters. Amaya[1], too, allows users to modify every accessed Web page, but, unless the user has write permission on the resource, the application creates a distinct resource stored on a preset server. XanaWord, on the contrary, allows the creation of a personal variant of the document strictly connected to the original copy. The basic scenario shows a user normally browsing WWW pages through a browser, then request the browser to edit the currently displayed page. An instance of MS Word is launched where the user edits the page and saves it in a specialized server. The system will then extract the changes introduced in the editing session, and creates a personal variant of the document (or a new version in case the user is also the original owner of the document). Anytime the user requests the same document the system will then add the changes again and will provide the user with the modified resource, although the original copy might still be unchanged. All web pages can be edited. All changes are stored separately from the original document, and are added on the fly when the user is requesting the document. Note that the user is not using some special tool, but can exploit his/her previous knowledge of web browsers and word processors, as with the experience of ISA. XanaWord is an environment coexisting with World Wide Web, allowing everybody to publish and modify pages during their browsing activity, through these well-known and widespread interfaces.

3.3 Versioning

The XanaWord system is based on versioning, i.e., the management of subsequent states of the same resource, through the imposition of a structure on them, and the ability of browsing and accessing every intermediate state. In [19] we pointed out some of the advantages of this mechanism, first of all parallel asynchronous collaboration, which is among the purposes of XanaWord. Furthermore, customization of public documents is also based on versioning. Since it is impossible to modify the main copy of the documents directly on their origin server, we rely, as CritLink [11] and XLinkProxy [4] do, on external anchoring and a non-transparent HTTP proxy to provide the required functionalities. When a document is requested, all changes introduced by each user (extracted through a forward delta based versioning engine, and stored in a separate database in XML format) are applied on-the-fly to the original document retrieved from the origin server. Note that this algorithm is not a generic differencing mechanism on XML data (see instead [5]) but it has been designed and implemented to handle and exploit specifically MS Word documents, which already include the change tracking information. Client side, through ad hoc browser menu, users can request all versioning features (client/server communication is based on WebDAV). An interesting option surely is the request "Show all changes": in this case a XanaWord server sends to the browser an unique document, containing a version of the document where all the previous changes are shown with different colors, as shown in figure 3.



fig. 3 Showing multiple version of a XanaWord document

4 Conclusions

At this moment, XanaWord is a complete editing and customization tool for web pages. It provides users with the possibility of accessing any web page, regardless of location, and modify it for their own purposes. Modifications are stored in an intermediate server that acts as a proxy and provides the modified version any time it is requested.

As mentioned, XanaWord is a partial result of a larger scale effort towards the implementation of a universal web editing environment allowing any user to access and edit any document in any data format in any moment. Among the improvements we are seeking is the abandonment of the proxy architecture, which is slow, unsafe and not applicable in intranets. Furthermore, we aim to widen the scope of application in terms of data formats: while XanaWord is specially aimed at MS Word documents, the whole project should allow any tool to be used, including web-based client editors and wiki plain-text syntaxes. We believe that XanaWord may show that, although impossible in the classical Web architecture, it is not difficult to provide a better integration between the author and reader's role in the World Wide Web, and that current technologies are sufficiently sophisticated already to allow this kind of editing without revolutionizing the basic architecture of the World Wide Web.

5 References

[1] Amaya, W3C, <http://www.w3.org/Amaya/>.
 [2] Bieber M., Vitali F., Ashman H., Balasubramanian V., Oinas-Kukkonen H., "Fourth Generation Hypertext: Some Missing Links for the World Wide Web", *International Journal of Human-Computer Studies*, 47, 1997, 31-65.
 [3] Blood, R. *Weblogs: a history and perspective*, http://www.rebeccablood.net/essays/weblog_history.html, September 2000, retrieved February 14, 2003.

[4] Ciancarini, P., Folli, F., Rossi, D. and Vitali F., XLinkProxy: external linkbases with XLink, *Proceedings of the 2002 ACM symposium on Document Engineering*, ACM Press, 2002, p. 57-65.
 [5] Cobena G., Abiteboul S., Marian A., "Detecting changes in XML Documents" in: *Proceedings of ICDE 2002*, San Jose, California, USA, IEEE, 2002, pp. 12.
 [6] Conklin, J. Hypertext, an introduction and Survey, *IEEE Computer*, 20(9), 1987, pp. 17-41.
 [7] De Rose S.J., Maler E., Orchard D., "XML Linking Language (XLink) Version 1.0", W3C, 2001, <http://www.w3.org/TR/2001/REC-xlink-20010627>.
 [8] Goland Y., Whitehead E., Faizi A., Carter S., Jensen D., "HTTP Extensions for Distributed Authoring – WEBDAV", RFC 2518, <http://www.ietf.org/rfc/rfc2518.txt>
 [9] Goodman D., *The Complete HyperCard Handbook*, New York, Bantam Books, 1987.
 [10] HTS, Keio University, <http://green.iprs.sfc.keio.ac.jp/~hts/>
 [11] Ka Ping Yee, "CritLink Mediator", Foresight Institute, 1997, <http://crit.org/>.
 [12] Leuf, B. & Cunningham, W. (2001). *The Wiki way*. New York: Addison-Wesley, 2001.
 [13] Miller S.M., *Udanax*, <http://www.udanax.com/>, 1998.
 [14] Nelson T.H., "Transcopyright: Dealing with the Dilemma of Digital Copyright.", *Educom Review*, 32(1), 1997, p. 32-35.
 [15] Nelson T.H., *Literary Machines*. Sausalito (CA), USA, Mindful Press, 1987.
 [16] Pam, A. *Fine-Grained Transclusion in the Hypertext Markup Language*, 1997, <http://xanadu.com.au/archive/draft-pam-html-fine-trans-00.txt>
 [17] Shneiderman, B. (1987). User interface design for the Hypertypes electronic encyclopedia, *1st ACM Conference on Hypertext*, November 1987, ACM Press, pp 189-194.
 [18] Vitali F., "Created sophisticated web sites using well-known interfaces" in: *HCI International 2003 Conference*, Crete (Greece), 2003. In print.
 [19] Vitali F., Durand D., "Using Versioning to Provide Collaboration on the WWW", *The World Wide Web Journal*, 1(1), 1995, 37-50.
 [20] *Wiki Philosophy FAQ*, WikiWikiWeb, <http://www.c2.com/cgi/wiki?WikiPhilosophyFaq>, retrieved April 14, 2003.