



The American Institute for Conservation of Historic & Artistic Works

## FINAL REPORT

Project MT-2210-9-NC-32: JAIC-Online: Searchable Access to the Complete Text of the Journal of the American Institute for Conservation (JAIC) on the World Wide Web.

The American Institute for Conservation is the largest conservation membership organization in the United States, and counts among its more than 3000 members the majority of professional conservators, conservation educators and conservation scientists worldwide. The Journal of the American Institute for Conservation (JAIC, or the Journal) is the primary vehicle for the publication of peer-reviewed technical studies, research papers, treatment case studies and ethics and standards discussions relating to the broad field of conservation and preservation of historic and cultural works. Subscribers to the JAIC include AIC members, both individuals and institutions, as well as major libraries and universities.

The JAIC began as the Bulletin of the International Institute for Conservation - American Group (IIC-AG), in April 1961. In 1975, three years after the IIC-AG became independently incorporated as the American Institute for Conservation (AIC), the last issue (Volume 16, Number 1) of the Bulletin was printed. Two years later, the publication was completely redesigned as the JAIC, and the first issue (Volume 16, Number 2) was released to the membership. Since that time, the JAIC has been published at first twice and now three times a year, and between Volume 16:2 (February, 1977) and Volume 39:3 (Fall, Winter 2000) there are 56 issues containing approximately 5000 pages of text and almost 2000 illustrations. These issues include a broad range of peer-reviewed articles on conservation treatment case studies; issues of conservation history, philosophy, or method; conservation research; and technical studies aimed at addressing questions in allied fields. Given the progressively incremental development of conservation theory and practice, the great majority of these benchmark articles continue to be relevant and useful; and in fact represent the greatest single resource for the core body of knowledge relating to preservation and conservation activities.

Unfortunately, back issues of this essential publication were not always easy to find. Because the JAIC is essentially an AIC member benefit, many museum conservation libraries may only contain copies subsequent to their original affiliation with the organization. In the past two decades individual membership in AIC has tripled, and institutional membership has increased five times what it was in 1977. Clearly, the majority of conservators practicing today would rarely have access to a complete set of AIC Journals. And students, museums without conservation departments, and the majority of allied professionals would find access even more difficult.

Nor was there an unlimited supply of back issues available from the AIC. Volumes from 1977

to 1989 exist only in a small number of original hard-copies, since either no digital version was ever produced or was never saved. Issues published after 1989 were written in a variety of mutually incompatible software formats, and stored on hardware ranging from 5 1/4" PC floppies to Macintosh Zip disks. Notwithstanding the realization that archiving materials on magnetic media with its limited service life is nowhere near 'archival', rapid changes in the computer industry have also rendered many older program files and disk formats unreadable. Furthermore, substantial corrections were often made to the original files at the printer. And none of the illustrations, photographs and line art are available in digital form, since they were usually stripped into the 'blue line' copies prior to printing.

In 1998, the AIC Board decided to make this unique resource available via the world wide, both to preserve the data contained in the JAIC, and to increase its availability by allowing the files to be publicly accessed. At the same time, it was decided to only include issues more than three years old, thus retaining the members-only benefit of recently printed editions. This website, which went online in April, 2001, is the result. Subsequent issues of the JAIC, as well as other AIC publications, will be periodically added to further develop this valuable resource.

#### Technical Details

To create this site, the text of each article in the JAIC was digitized and marked up using XML, or eXtensible Markup Language. Like its brother HTML, XML is a subset of SGML, the generalized markup language that is the International Standard (ISO 8879:1986) for platform-independent markup of structured documents. Unlike HTML however, XML incorporates most of the power of SGML and is designed to be extensible, thus making it an ideal language for marking up semi-structured documents, such as the JAIC. Unlike HTML, which tends to mark up primarily stylistic elements such as headings and paragraphs, XML marks up document-specific syntactic elements such as detailed bibliographic information including, for example, an element for the address of a conservation material supplier. XML can also facilitate the generation and management of meta-data, or embedded information about the information in the text. This can result in a high level of automation once the data is held in this format, and can enable fully automated publication, both on paper and via the internet, and a great increase in structured information retrieval.

The data conversions were done by Techbooks, a company experienced with technical journal formats and SGML markup. Techbooks was provided with every copy of the Journal between Issue 16:2, February 1977 (the first issue) and Issue 38:, 1999, either in hardcopy format or on disk. Where computer disks were available, hardcopies were also provided because of possible changes in the final published version.

Technicians at Techbooks analyzed the contents of the Journals and, in collaboration with members of the project team, developed a custom DTD, or Document Type Definition, which describes the content of the JAIC, and defines the various markup elements and entities that comprise the logical and physical structure of a JAIC article.

Once the DTD had been created and accepted, the process of data conversion and markup

took place. Older journals that exist in only paper format were manually retyped; and the wide variety of later digital versions were translated from their proprietary formats. The digital files were checked against the printed versions and manually corrected as necessary. Images, tables, charts, and graphs were scanned, and saved in digital form. Finally, all the files were marked up in XML according to the JAIC DTD.

After the JAIC articles had been converted into XML format, publishing of the data was simply a matter of writing a script using XSTL, or eXtensible Style Language Transforms. The JAIC XSL transform script parsed the syntactical information embedded in the JAIC XML files to create the correctly formatted HTML documents for web delivery. Similar XSLT scripts can produce cumulative bibliographies, material lists, and other data extractions from the JAIC XML database. Finally, a graphical user interface (GUI) was developed to facilitate browsing, and the entire website was indexed to enable keyword searches.

The following programs and utilities were used to produce this website:

The JAIC DTD and markup were written using XML Spy (for information about XML Spy see <http://www.xmlspy.com> ).

The JAIC XSL transform was written in XEmacs, using XSLide mode (for information about Xemacs see <http://www.xemacs.org>; for information about XSLide see <http://www.menteith.com/xslide> ).

The XSL transform was processed in XEmacs using xt (for information about xt see <http://www.jclark.com/xml/xt.html> ).

The table of contents was generated using XSLT and Perl (see <http://www.perl.com>)

The graphical user interface was created in Adobe GoLive (see <http://www.adobe.com/products/golive/main.html> ).

Additional HTML was written in XEmacs using psgml mode (see [http://www.lysator.liu.se/~lenst/about\\_psgml](http://www.lysator.liu.se/~lenst/about_psgml) )

The search index was created using ht://Dig (see <http://www.htdig.org> )

Project Team:

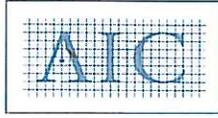
JOHN BURKE, Head of Conservation at the Oakland Museum of California, conceived of this project in 1998 while serving as Director for Professional Education on the Board of AIC. John wrote the grant application and acted as Principal Investigator, directing the development of the DTD and the data conversion process. John also wrote the XSLT and Perl scripts, generated the final HTML files, and created the graphical user interface of the website.

WALTER HENRY, Lead Analyst in the Preservation Department at Stanford University Libraries and Webmaster of CoOL, provided technical oversight in the development of the grant application, and served as SGML content expert during the DTD development. Walter also

performed the image translations for web delivery, installed and created the site search index, and continues to manage the overall server operations critical to delivery of this resource.

PENNY JONES, Executive Director of AIC, along with her staff administered the financial aspects of the project, and managed the process of gathering together the numerous original printed editions and the variety of digital media sources for delivery to the data conversion specialists.

TECHBOOKS, a provider of internet application development and content preparation services, analyzed the Journal structure to create the JAIC DTD, rekeyed approximately 4500 pages of the text into XML format, and scanned approximately 2000 images to provide the raw material for this resource.



The American Institute for Conservation of Historic & Artistic Works

## EXECUTIVE SUMMARY

**Project:** MT-2210-9-NC-32

**Title:** JAIC-Online: Searchable Access to the Complete Text of the Journal of the American Institute for Conservation (JAIC) on the World Wide Web.

The Journal of the American Institute for Conservation (JAIC) is the primary vehicle for the publication of peer-reviewed papers relating to the field of conservation of historic and cultural works. Since its first issue in February, 1977, and Volume 39:3 (Fall, Winter 2000) there have been 56 issues containing approximately 5000 pages of text and almost 2000 illustrations. In 1998, the AIC/FAIC Board decided to make this resource available via the world wide, both to preserve the core data, and to increase its availability by allowing the files to be publicly accessed. The resulting website ( <http://aic.stanford.edu/jaic> ) went online in April, 2001.

To create this site, each article was manually re-entered, since digital versions were either never produced, or published in a variety of software formats and often changed at the printer. After conversion, a custom Document Type Definition (DTD) describing the logical structure of a JAIC article was developed, and the articles were marked up using XML, or eXtensible Markup Language. Images, tables, charts, and graphs were scanned as digital images. A transform script was written in eXtensible Style Language (XSL) to parse the XML files and create formatted HTML documents for web delivery. Finally, a graphical user interface (GUI) was developed to facilitate browsing, and the website was indexed to enable searches.

Currently, the site includes only issues over three years old, in order to retain the members-only benefit of recent editions. Subsequent issues of the JAIC will be periodically added.