

**Standardized Markup for Journal Articles
Based on the NLM's Journal Archiving and Interchange Tag Suite:
A Proposed NISO Work Item**

Proposal for Consideration by the NISO Voting Membership
Approval Ballot Period: August 4 – September 2, 2009

The following proposed work item is submitted by
Jeff Beck, Technical Information Specialist, National Center for Biotechnology Information (NCBI),
U.S. National Library of Medicine, National Institutes of Health
beck@ncbi.nlm.nih.gov

Proposal Last Modified: June 18, 2009
Approved by the NISO [Content & Collection Management Topic Committee](#) July 29, 2009

Work Item Title:

Create a Standardized Markup Language for Journal Archiving and Interchange based on the NLM's Journal Archiving and Interchange Tag Suite (proposed short title: NLM Tag Suite)

Background and Problem Statement:

The [National Center for Biotechnology Information \(NCBI\)](#) of the [National Library of Medicine \(NLM\)](#) created the Journal Archiving and Interchange Tag Suite with the intent of providing a common format in which publishers and archives can exchange journal content.

The Suite was created in response to an early draft of the PubMed Central DTD¹, which was intended for use with life sciences journals. Input and support from Harvard University Libraries, as well as support from The Andrew W. Mellon Foundation and collaboration with Inera, Inc. and Mulberry Technologies, Inc., allowed the scope of the project to be broadened and resulted in the Journal Archiving and Interchange Tag Suite as it currently exists.

The intent of this Tag Suite is to preserve the intellectual content of journals independent of the form in which that content was originally delivered. The Suite has been written as a set of XML schema modules, each of which is a separate physical file. No module is an entire schema by itself, but these modules can be combined into a number of different schemas. The modules define elements and attributes for describing the textual and graphical content of journal articles as well as some non-article material such as letters, editorials, and book and product reviews.

Although the Suite was developed to support electronic production, the structures should be adequate to support print production. The Suite can be used to construct schemas for authoring and archiving journal articles as well as transferring journal articles from publishers to archives and between archives.

Three schemas for journal articles have been created from the Suite and are maintained by NLM as well: NLM Archiving and Interchange Tag Set, NLM Journal Publishing Tag Set, NLM Article Authoring Tag Set.

¹ Document Type Definition (see http://en.wikipedia.org/wiki/Document_Type_Definition).

These schemas are maintained in DTD format; however, W3C and RELAX NG versions are available for each release.

The Tag Suite and associated schemas are in version 3.0, which was released by NLM in November 2008. Details can be found on the NLM website: <http://dtd.nlm.nih.gov/>.

Statement of Work:

The goal of this work item is to take the currently existing Journal Archiving and Interchange Tag Suite version 3.0, the three journal article schemas, and the documentation and shepherd it through the NISO standardization process. We intend for the Tag Suite to be the standard and each of the schemas to be a “sub-standard” or appendix to the Tag Suite standard.

Because of the maturity of the Tag Suite, we believe that, once a Working Group is created, it will be possible to propose that the Tag Suite, as it currently stands, be accepted as a Draft Standard following a brief period of review within the group.

Partners and Participation:

The NLM Tag Suite currently has an advisory committee consisting of representatives from:

- NCBI/NLM
- Commercial and Society Publishers
- Tagging vendors
- Markup Consultants
- Journal Archives

Those groups should continue to be represented on the NISO Working Group, with as much continuity as possible with the currently serving members. Currently there is no representation from national and research libraries (beyond the NCBI participants representing PubMed Central); this is certainly a group that should be involved.

Timeline:

Milestone	Timeline
Appointment of working group.	Month 1
Approve initial Work Plan (which will include determining how to package the Tag Suite as a Draft Standard).	Months 1-3
Submit Tag Suite as Draft Standard; comment period.	Months 3-9
Completion of final standard document (including review of and response to comments received during the draft standard phase).	Months 10-12
Approval and Publication	Month 13

Funding:

N/A