

# **Tag Set Working Group Recommendations January 2007**

# Table of Contents

1. Tag Set Working Group Meeting .....	1
1.1 Attendees.....	1
1.1.1 NLM Staff.....	1
1.1.2 Secretariat (Mulberry Technologies, Inc.).....	1
1.1.3 Telephone Participants.....	1
1.2 Change Applicability .....	1
1.3 DTD Rationales .....	2
2. Decisions for the V2.3 Release.....	2
2.1 Currency attribute to <price>.....	2
2.2 Type Attributes .....	2
2.2.2 New attribute rationale.....	2
2.2.4 Add a “type” attribute .....	3
3.2 Layout/Display Related Requests .....	4
3.2.1 Format/Display Oct/Nov Decisions.....	4
4. Add namespace support .....	4
5. New Book Tag Set from the Suite .....	5
6. Documentation Suggestions .....	5
6.1 Document <sub-article> and <sec> .....	5
6.2 Incorporate PMC Guidelines in Tag Set Documentation .....	6
6.3 <hr> in table cells .....	6

# 1. Tag Set Working Group Meeting

The Working Group meeting took place by conference call on January 10, 2007.

This document contains a record of the advice of the Working Group when the group was able to reach consensus. Those items not addressed because of time limitations are not included here, but will be included in the list of discussion topics for the next Working Group conference call, on Tuesday February 6<sup>th</sup> 2007, at 11:00am EST.

## 1.1 Attendees

### 1.1.1 NLM Staff

- Jeff Beck (Moderator)
- Steve DeRose
- Marilu Hoepfner
- Laura Kelly
- Adeline Manohar
- Kim Tryka

### 1.1.2 Secretariat (Mulberry Technologies, Inc.)

- Deborah A. Lapeyre

### 1.1.3 Telephone Participants

- Mark Doyle (American Physical Society)
- Beth Friedman (DCL)
- Kathryn Heniss (Highwire)
- Sharon McCamant (Cadmus)
- John Meyer (Portico)
- Jules Milner-Brage (Highwire)
- Evan Owens (Portico)
- Bruce Rosenblum (Inera)
- B. Tommie Usdin (Mulberry)

## 1.2 Change Applicability

The following Tag Set icons appear with each suggested change to indicate to which of the Tag Sets a change applies:

 Archiving

 Publishing

 Article Authoring

 Book

## 1.3 DTD Rationales

Scope: **A** **P** **A** **B**

Rationales for the different tag sets were read and approved in general.

*Recommendations:*

Add a statement to the stated purpose of the Archiving DTD noting that the design for flexibility helps with differences in publisher granularity as well as publication order. Thus the flexibility is useful both for allowing conversion without extensive reordering and for matching various granularities in the publisher documents.

## 2. Decisions for the V2.3 Release

### 2.1 Currency attribute to <price>

Scope: **A** **P** **A** **B**

*Request:*

Add country code attribute to <patent> inside <citation>.

*Recommendations:*

A “currency” attribute will be added to <patent>. The ISO 3166-1 alpha-a codes will be referenced in the documentation and recommended as best practice.

### 2.2 Type Attributes

Scope: **A** **P** **B**

*Request:* The suggestion was made that the Working Group consider which of the elements that do not currently have "type" attributes might usefully take them.

#### 2.2.2 New attribute rationale

The Tag Library will document the rationale behind the addition of many new “type” attributes:

- Unless there was a conflicting rule, add a “type” attribute.
- Elements that cannot repeat (such as <body> and <title-group>) should not take a “type” attribute. This includes all the “-meta” elements and most top-level structural elements (such as <body> and <back>).
- Elements that can repeat by virtue of being part of a #PCDATA OR-group should take a “type” attribute.
- Elements that already have a “type” or “style” attribute should not get a second such attribute.
- Elements concerning glyph characters should not have type attributes.
- EMPTY elements (such as all the counts) should not have type attributes.

## 2.2.4 Add a “type” attribute

A “type” attribute will be added to each of the following:

<abbrev>	<access-date>	<ack>
<addr-line>	<address>	<aff>
<alt-text>	<annotation>	<app-group>
<app>	<array>	<article-title>
<attrib>	<author-comment>	<bio>
<book-title>	<chem-struct-wrapper>	<chem-struct>
<collection>	<collection-id>	<collection-member>
<comment>	<conf-acronym>	<conf-date>
<conf-loc>	<conf-name>	<conf-num>
<conf-sponsor>	<conf-theme>	<conference>
<contract-num>	<contract-sponsor>	<contrib-group>
<copyright-holder>	<copyright-statement>	<copyright-year>
<corresp>	<country>	<day>
<degrees>	<disp-formula>	<disp-quote>
<edition>	<elocation-id>	<email>
<fax>	<fig-group>	<fn-group>
<fpage>	<gloss-group>	<glossary>
<gov>	<graphic>	<inline-formula>
<inline-graphic>	<inline-supplementary-material>	<institution>
<isbn>	<issue-id>	<issue-title>
<issue>	<journal-subtitle>	<journal-title>
<long-desc>	<lpage>	<media>
<month>	<name>	<note>
<object-id>	<page-range>	<patent>
<phone>	<price>	<publisher>
<ref-list>	<ref> (special)	<role>
<season>	<self-uri>	<series-text>
<series-title>	<series>	<sig-block>
<sig>	<source>	<speaker>
<speech>	<statement>	<std>
<string-date>	<sub>	<subject>
<subtitle>	<suffix>	<supplementary-material>
<table-wrap>	<table>	<table-wrap-group>
<tex-math>	<time-stamp>	<title>
<trans-source>	<trans-subtitle>	<trans-title>
<uri>	<verse-group>	<verse-line>
<volume-id>	<volume>	<x>
<year>		

## 3.2 Layout/Display Related Requests

### 3.2.1 Format/Display Oct/Nov Decisions

Scope: **A** **P** *A* **B**

*Recommendations:*

1) **Expanded Emphasis Class.** The emphasis class to be expanded to the following elements (new ones are bold below). Note that “roman” was added in Oct/Nov but removed from the list below by the Working Group in January.

- bold | italic | monospace |
- overline | overline-start | overline-end |
- **sans-serif** | sc | strike |
- underline | underline-start | underline-end

*Recommendations:*

2) **HTML Style for Tables.** The HTML “style” attribute to be added back to the HTML table model for elements <th> and <td>.

## 4. Add namespace support

Scope: **A** **P** *A* **B**

*Request:*

In order to make it easier to use namespace-aware tools such as NVDL (Namespace-based Validation Dispatching Language), we suggest adding a namespace declaration to all the DTDs and schemas and an optional/variable namespace prefix to each element and attribute.

*Recommendations:*

- A namespace URI will be established and published in the documentation for each of the tag sets. The three book tag sets will be given individual URIs.
- No default namespace will be added to any of the tags sets at this time.
- The group members will continue to collect experience using namespaces with this data. Working Group members are encouraged to create websites that describe their uses of namespaced NLM tagsets and provide working descriptions and hints. The main Tag Set website will point to such documentation.

## 5. New Book Tag Set from the Suite

Scope: **B**

*Request:*

Some interest has been expressed in having a generic “book” archiving model made from the DTD Suite, in the same way that a generic journal article archiving model was made. The current NCBI Book model was written to specific NCBI requirements and is not generic enough for many would-be users.

*Recommendations:* Find the people and scope the project:

- A scope and goals statement for a generic book tag set should be written.
- Initially, such a tag set would focus on STM books such as books that are collections of articles and reference books. Later textbooks might be added. Technical books that are adequately covered by DocBook should not be considered.
- Such a tag set should be based on the Publishing (Blue) tag set with bookish metadata.
- Some book-specific publishers need to be invited to the Working Group.

## 6. Documentation Suggestions

In general, the discussion was that the Tag Libraries need many more examples of best practice and guidelines for use. People need to be told both what to do and what to avoid.

*Recommendations:*

As an **action item** for the entire group, we request that Working Group members send in documents and document fragments they consider well-tagged as well as those they have seen “in the wild” that they consider controversial. Please send such documents to:

<dalapeyre@mulberrytech.com>

### 6.1 Document <sub-article> and <sec>

Scope: **A P B**

*Request:*

The documentation needs to make a clearer distinction between what a <sec> is and what is more properly tagged as a <sub-article>.

*Recommendations:*

Additional examples from several journal types should help make this distinction more clear.

## 6.2 Incorporate PMC Guidelines in Tag Set Documentation

Scope: **P** **A**

*Request:*

Is there mileage in incorporating the PMC tagging guideline recommendations into the NLM Tag Set documentation? The PMC guidelines seem generic enough to be useful to all users of the DTD.

*Recommendations:*

The PMC Guidelines will remain a separate document focused on PMC processing.

However, it is the consensus of the group that DTD users want clear guidelines and samples of best practice. Even people who do not agree with the idea of PMC politically or philosophically want usage guidelines and will accept some PMC guidance. NLM and Mulberry will go over the guidelines carefully and incorporate as much into the Tag Library as seems general rather than PMC-processing-specific.

## 6.3 `<hr>` in table cells

Scope: **A** **P** **A** **B**

*Recommendations:*

Put in a footnote into the table element discussion that there are issues with browser implementation and the `<hr>`. Make reference to the description of floating and non-floating constructs in the CSS specification to better understand the observed typographic effects.

-- document end --