

# How JATS Empowers Scholarly Communication

Advantages of JATS for journal article publishing, production,  
and management

**Deborah A Lapeyre**

Mulberry Technologies, Inc.  
17 West Jefferson Street  
Rockville, MD 20851  
USA  
dalapeyre@mulberrytech.com

Version v1.0 (September 2018)

Copyright © 2018 Mulberry Technologies, Inc.





# How JATS Empowers Scholarly Communication

JATS: a standard model for journal articles* . . . . .	1
Some advantages of using JATS are really advantages of using XML . . . . .	1
Why JATS? . . . . .	2
Synergy in numbers: the world publishes journals in JATS . . . . .	2
JATS model fits journal production . . . . .	3
For publishers: faster and less expensive . . . . .	3
Also improved lifecycle and quality . . . . .	4
More intelligent QA than . . . . .	4
Enables new QA techniques . . . . .	4
Also advantages for scholarly community . . . . .	5
JATS supports extensive metadata that travels <i>with</i> the XML . . . . .	5
Metadata for mining and extraction . . . . .	6
Detailed references . . . . .	6
JATS enables interchange and communication! . . . . .	6



## **JATS: a standard model for journal articles\*** **(Journal Article Tag Suite)**

JATS is a named collection of XML elements describing  
*one journal article*  
called a “model” of the article

- Originally modeled STEM articles (Scientific, Technical, Engineering, Medical)
- Now covers all journal types: (sociology, economics, soft sciences, etc.)
- Models only journal *articles*, not books, issues of journals, magazines, etc.

\*ANSI/NISO Z39.96-2015 JATS: Journal Article Tag Suite (JATS V1.1)

## **Some advantages of using JATS are really advantages of using XML**

- Platform and vendor independence
- Content computer-manipulable and human-readable
- Full Unicode (internationalization)
- Mark up data *once* and make multiple output products
- Validate structure with a schema
- Machine-readable context for better searching/data-mining

## Why JATS?

- **Designed** for modern journals (the 80/20 case)
- **Price is right:** models, documentation, some tools are *free*:
  - US National Library of Medicine (NLM)
  - github (JATS4R et al.)
- **Documented:** extensive Tag Libraries, examples, recommendations
- **Highly customizable**
- **Semantic and declarative markup**, not presentational/behavioral

## Synergy in numbers: the world publishes journals in JATS

(many languages and many multilingual documents)

- SciELO schema is a JATS superset
- Almost all middle-sized/small journal publishers use JATS
- All of the huge publishers *can* (and do) make JATS for interchange
- Archives accept (or require) JATS  
(US and UK PubMed Central, British National Library, US Library of Congress, Australian National Library, ITHAKA/JSTOR, many more)
- Web-hosting and service vendors require or support JATS
- JATS used world-wide  
(including: All the SciELO countries, Australia, Belgium, Canada, Egypt, Finland, France, Germany, Italy, Japan, Norway, Russia, South Korea, Sweden, US, UK, and more)

Whole communities can agree on data content

## JATS model fits journal production (structure of a JATS journal article)

- **Front Matter (<front>)**
  - journal-level metadata
  - article-level metadata
- **Body Matter (<body>)**
  - paragraphs
  - sections
  - figures and tables (XHTML, CALS)
  - equations (MathML)
- **Back Matter (<back>)**
  - appendices
  - citations with detailed metadata

## For publishers: faster and less expensive

- Faster time to publication
- Single format for accepting content  
(publishers create, convert on receipt, authors create)
- Single source publishing: many output products from one file:  
(HTML5, PDF, other print, eBook, web apps, accessible files)
- Interoperability
- Reduced cost of article production

Faster publication and lower cost make open access more sustainable!

## Also improved lifecycle and quality

- Editors, reviewers, and production get the advantages of XML
- Improved communication throughout lifecycle
- Investment in tools, training, process pays benefits throughout lifecycle
- Better control of repository or archive
- Improved Quality Assurance

## More intelligent QA than

### “Does it look right?”

- Checking:
  - business rule compliance and irregularities
  - empty titles and paragraphs
  - internal cross-references
- References checked with deposit authority
- Order or numbering of references corrected

## Enables new QA techniques

- Schematron (the rules-based XML validation language)
- False color proofs  
(turn semantic components bright colors and scan by eye)
- Lists of anything imaginable, for human checking
- Code values checked against authority files  
(journal abbreviated titles, keywords, surnames)



## Also advantages for scholarly community

- Production advantages for publishers and archives
- Manageable and searchable text of an article
- Recording metadata essential for research communication

## JATS supports extensive metadata that travels *with* the XML

Article-level metadata for citation and credit

- Article title (including alternative languages)
- Article identifiers (DOI, etc.)
- Copyright and permissions
- Finding aids (abstracts, keywords, subjects)
- Contributors with
  - alternative names (both Japanese characters and romanized)
  - multiple affiliations (with alternative names for multiple languages)
  - multiple roles (possibly linked to CRediT taxonomy)

## Metadata for mining and extraction

- Machine-processable external identifiers
  - for authors (e.g., ORCID)
  - for institutions (e.g., Ringgold, ISNI)
- Detailed funding description
  - both monetary and non-monetary funding
  - award and grant identifiers
  - funding sources and principals
- Labeled Data Availability Statements
- Terms and subjects linked to ontologies/taxonomies
- Links to companion articles (different language versions), related books or web pages, errata

## Detailed references

To allow publication tracking

- For journal articles named elements include journal title, article title, journal issue, first and last page numbers, author surname, day-month-year of publication, and DOI
- sufficient named elements to cite books, patents, conference papers, standards, and other works
- Allows for data citations (including access times)

## JATS enables interchange and communication!