JATS4R - working together to apply the standard standardly

Melissa Harrison*

JATS4R Chair, Head of Production Operations, eLife Sciences Publications Ltd., Westbrook Centre, Cambridge, UK

Abstract. JATS4R is a volunteer-run organisation that produces recommendations for how people should use the Journal Article Tag Suite (JATS XML) to aid the exchange, reuse, and mining of content. This paper briefly describes this overlay to the NISO Standard Z39.96-2019 which itself defines a set of XML elements and attributes for tagging journals articles and provides three models. JATS4R brings a level of normalization to the use of the basic standard to ensure interoperability across vendors. This paper discusses what JATS4R does, how the standard is maintained and updated, what the oversight group has achieved since it was established in 2013, and what the future may hold.

Keywords: JATS, XML, JATS4R, DTD, Community, Journal Article Tag Suite

1. What is JATS4R?

JATS4R is a volunteer-run organisation that receives no formal funding. It exists and continues to exist because of the passion and commitment of a few people, coupled with the voluntary time commitment of others within the JATS "community".

JATS4R was established in 2013 and is all about optimising the reusability of scholarly content by developing best-practice recommendations for tagging content in JATS XML. A JATS XML file must be valid according to the Document Type Definition (DTD), but the JATS models allow multiple ways to tag the same information; a document that works in your system might not be understood by users outside of your system. However, if you follow the JATS4R recommendations, others can understand how your content is tagged and exchange content with you much more easily because the tagging is normalized and expected. This gives you the benefits of exchanging your content with vendors and hosts and enables more flexibility in your workflows and the potential to, for example, change your vendors more easily. Expanding this concept even wider, it is worth considering that 60% of a data scientist's time is taken up with clean-up of the data. JATS4R aims to reduce this wasted effort. Therefore, there are multiple benefits to adhering to JATS4R recommendations, whatever type of publisher you are.

---

*Tel.: +44 1223 855340; E-mail: m.harrison@elifesciences.org.

2. What's the difference between JATS and JATS4R?

President of Mulberry Technologies, Inc., a consultancy specializing in design of XML vocabularies for prose documents and publishing has changed over this time. Therefore, JATS4R acts as an overlay to help create more standardisation in application of the JATS models; JATS4R is a NISO working group and the output is recommended best practices.

3. What does the Chair of JATS4R do?

As chair of the organisation I see my main responsibilities in two ways. Before going into that though, it's important to point out that I effectively chair the organisation without being a JATS or XML expert, a developer/technologist, a designer, or a developer. This has the benefit that I know I am not the expert that can come up with the right recommendation. My role is primarily to pull together and facilitate the groups of people gathered in order to do this. My paying job facilitates what I see as my secondary role in JATS4R.

I work for a forward-thinking and innovative publisher, which means I have the opportunity to be aware of and part of new initiatives in publishing. I am involved in cross-publisher and cross-organisation working groups so I have the opportunity to promote JATS4R and also put it at the centre of some of the requirements for change. Editorial policies can develop, but without thinking about the structural document changes underlying this, and the interchange and reuse of this information, in an increasingly open information world, this is detrimental to the overall effort. Therefore, I greatly enjoy this role of rallying the wider community to contribute their time and expertise to subgroups in order to produce recommended practices that should create interchange more easily at the start of a new initiative. This also reduces the effort for individual publishers to have to think through how they will tag new content models; we've done some of the heavy lifting for them.

4. What has JATS4R achieved?

Although JATS4R came about as a result of a number of open-access publishers working together, it has quickly expanded and we now have input and expertise from many different publishers, submission systems, hosting platforms, and other contributors to the expanding ecosystem of publishing.

JATS4R is headed up by a Steering Committee of five people, who are collectively responsible for the sign-off of recommendations before they go out for public review, and final sign-off before formal publication. This group also heads up subgroups as they are formed. We became a NISO working group in 2018. JATS4R has, as of May 2020, produced sixteen recommendations:

- Eleven have been formally completed
- Two are due to be published imminently
- One is out for public comment
- One is still being worked on by a subgroup
- One is awaiting the formation of a subgroup to complete the recommendation post JATS update

JATS4R focuses on new initiatives as well as addressing the variability in practices that have evolved over years. Both the new and the old present challenges, although they are sometimes different. For instance, the working group on authors and affiliations had to deal with very conflicting opinions on

long-established practices in tagging this information. The current peer review materials working group has been working for more than a year on this recommendation partly because very few publishers are as yet publishing peer reviews, and the emerging models are not yet fully-specified and embedded though the recommendation is trying to include them.

5. What is the future for JATS4R?

The JATS4R website is currently being redesigned, and we hope to have a new version out this summer (2020). With the new design, we hope to pull together under one roof different sources of information about the group. For instance, I receive emails from users with feedback on the recommendations and currently these are tracked via a GitHub project board. Ideally, we'd like users to provide feedback on the site and we want to host the roadmap on the website so that we maintain openness in the most accessible way.

The JATS4R steering committee members all have full time jobs, and managing a subgroup is not always within their labor bandwidth, therefore, we'd like to open this task out to other members of the community. This will involve support and good guidance, which again we hope the new website will help facilitate. Finally, we'd like more contributions from the community as to the issues/problems for which we should actually be creating recommended practices.

All of this is a big To-do list for a small group of volunteers, who are mainly XML geeks with little marketing experience, so we appreciate all the help and input we can get from others. Considering the limitations, we're proud of how far we've come and what we've achieved and look forward to making JATS4R bigger, better, and more important for creators and users of JATS XML in the future. As of December 2018, JATS4R became an official NISO Working Group. To learn more about how you can participate go to: https://jats4r.org/who-participates.