

Using Journal Article Tag Suite extensible markup language for scholarly journal articles written in Korean

Sun Huh¹, Tae Jin Choi², So-Hyeong Kim²

¹Department of Parasitology and Institute of Medical Education, College of Medicine, Hallym University, Chuncheon;

²National Research Foundation of Korea, Daejeon, Korea

Abstract

The Republic of Korea is the fourth ranking country for the number of PubMed Central (PMC) journals. As of September 2013, 75 journals from Korea are included in PMC. Starting in 2013, several research funding agencies for scholarly journal publications in Korea began to establish open access, full-text databases in the fields of medicine, science, and social sciences and humanities. In those databases, Journal Article Tag Suite (JATS) 1.0 is used so that articles written in the Korean language can be easily manipulated as full-text extensible markup language (XML). Editors or publishers must produce full-text XML files based on JATS 1.0. Thus this paper surveys the current state of the application of JATS 1.0 to Korean scholarly journals, including both those in English and those in Korean, focusing on the current technology, training programs, and the policy of the Korean government on open access, full-text XML. This experience in Korea can be a model for constructing mother-tongue, open access, full-text journal databases based on JATS 1.0 in other countries. The usefulness of JATS in scholarly journal publications not only of all fields, but also in all languages, is stressed.

Keywords

Database; Journal Article Tag Suite; Local language; Science and medicine; Open access journal

Received: October 25, 2013

Accepted: December 2, 2013

Correspondence to Sun Huh

shuh@hallym.ac.kr

ORCID

Sun Huh

<http://orcid.org/0000-0002-8559-8640>

Tae Jin Choi

<http://orcid.org/0000-0003-2363-4578>

So-Hyeong Kim

<http://orcid.org/0000-0002-3973-6782>

This paper was posted on the Journal Article Tag Suite Conference proceedings website available from: <http://www.ncbi.nlm.nih.gov/books/NBK159729/>.

Introduction

PubMed Central (PMC) has been the major gateway for medical journals from Korea to achieve worldwide dissemination, since it is the most accessible path to being indexed in PubMed, an important biomedical research database. Another pathway is inclusion in Medline, which-unlike PMC-accepts journals in various languages, but during the past 10 years, few medical journals of Korea have been approved to be indexed in Medline. When journals in Korea are rejected by Medline, they cannot help but change the language of their journals to

English and apply to PMC in order to reach an international readership. For medical journals in English, if the level of the extensible markup language (XML) quality fits PMC's criteria, it is easy to gain inclusion in PMC. Fig. 1 shows the number of journals from Korea included in PMC from 2008 to September 2013.

The increased visibility via PubMed and PMC have enabled many medical journals in Korea to increase their international readership, range and quality of submissions, and impact factor. This internationalization involves, first, being indexed in international citation databases such as Web of Science or Scopus; second, an increased number of submissions from outside of Korea; and third, increasing journal metrics year by year. When the first training for PMC XML production was offered in 2006 in Korea, the organizers could not imagine that medical journals from Korea would actually be able to achieve these things, but they are well underway for many such journals only 7 years later [1].

The Korean Association of Medical Journal Editors (KAMJE) has played a major role in producing PMC XML files. They constructed a full-text database based on PMC XML, entitled KoreaMed Synapse (<http://synapse.koreamed.org/>), in 2007. Scholarly journal editors in other fields began to pay attention to the activities of KAMJE. KAMJE's work motivated scientists in Korea to organize the Korean Council of Science Editors (KCSE), which was founded in September 2011. KCSE's main work involves education and training for science editors in editing, publication ethics, and publication of newsletters and journal on the international level. KCSE also has trained the editors and staffs of printing companies on adaptation of information technology to journal publishing.

Beyond the sciences, the National Research Foundation (NRF) of Korea has funded a full-text, XML-based, open access database for social science and humanities journals. Other

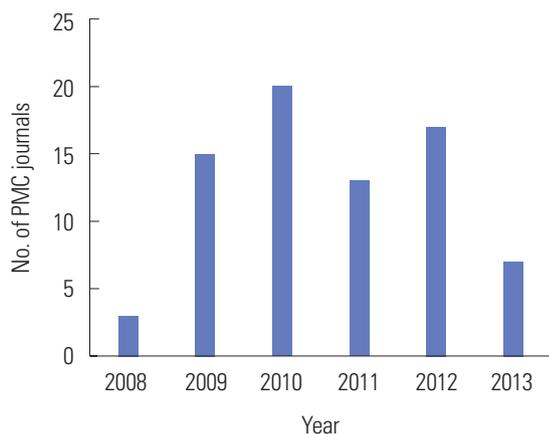


Fig. 1. Number of journals from Korea accepted for indexing in PubMed Central (PMC) each year.

new databases include Korea PMC, modeled on Canada PMC and Europe PMC, which is being constructed this year by the National Library of Medicine of Korea, an agency of the Korea Centers for Disease Control and Prevention. The Korean Federation of Science and Technology Societies (KOFST) is also testing an XML-based, full-text, free or open access database. The remainder of this paper will describe in more detail not only the Journal Article Tag Suite (JATS) XML-based, full-text databases in Korea but also the technology, training programs, and the policy of the Korean government on open access, full-text, XML-based databases.

KoreaPMC

The establishment of KoreaPMC was first proposed by KAMJE. In March 2007, KAMJE discussed the idea with the United States National Center for Biological Information (NCBI) that eventually became the contractor of Agreement for Participation in NIH PubMed Central (NIH Portfolio) Archive. However, this agreement was terminated in July 2007 when the NCBI established a new policy on the cloning of PMC by other countries. These clones, known as PMC international (PMCi), were to be managed as follows: first, PMCi should be supported by the governments of each participating nation; second, there should be a public access policy on government-funded biomedical papers. Because KAMJE is not a government-supported organization and the Korean government did not have a public access policy at that time, it was impossible for KAMJE to proceed with the project. Meanwhile, in 2011, the Korean government decided to establish a new organization, the National Library of Medicine (NLM) of Korea, under the Korea Centers for Disease Control and Prevention. Construction on the building began in 2012 and will be finished in 2014. Among the new databases managed by Korea NLM, KoreaPMC is also under construction. Owing to technical support from United States PMC, it is already well-established and in the testing period. KoreaPMC will index papers written in languages with non-Roman alphabets, such as papers in Korean. This project is completely supported by the Korean government. KoreaPMC will also be accepted as the third PMCi by US NCBI, since, in 2012, the Minister of Health and Welfare declared an open access policy for the publication of research funded by the ministry. KoreaPMC may appear at early 2014.

Social Science and Humanities Central

The Korea Citation Index (KCI) system is the citation database of more than 2,000 scholarly journals in Korea, including those in the fields of science, technology, engineering, medi-

cine, social science, and humanities and the arts, similar to Web of Science or Scopus. The NRF of Korea also asked the company to establish Social Science and Humanities Central (SSHC) for scholarly journals in Korea. Journals in English in the fields of social science and the humanities will be selected, and a JATS XML-based, full-text, open access database, in which the number of papers included is currently 1,000, much like PMC, will become available in 2014. No other international full-text, open access database in the fields of social science and the humanities exists in the world. SSHC will be the first step in this direction. SSHC will imitate PMC since PMC is the best model of full-text, free, or open access literature archives; however, it will accept journals in any local or regional language in the world, that is, in any language other than English. This work has been made possible based on a proclamation of the Korean Ministry of Education that stressed an open access policy for financial support of scholarly journal publication in December 2011. Through SSHC, the visibility of journals from Korea may increase dramatically. Out of the 34 OECD countries, Korea and Japan are unique in the level of financial and logistical support they provide for scholarly journal publication. The Korean and Japanese governments provided 8 million and 5 million US dollars, respectively, in 2013. KOFST and the NRF of Korea, both of which are funded by the Korean government, want entire scholarly journals, not just the individual papers, produced by the citizens of Korea to be competitive. In addition, those two agencies want to increase the knowledge of the public by providing free content. This is the reason why the Korean government supports scholarly journal publication and full-text databases.

ScienceCentral

ScienceCentral (<http://e-sciencecentral.org/>) is an archive of free or open access, full-text scientific scholarly journal literature at KOFST. It is in testing, with eight journals from Korea and one from Croatia as of September 2013. Out of the eight journals, two are in the Korean language. The Korean government supports scientific, technological, and medical journals through KOFST and social science, humanities, and arts journals through the NRF of Korea. Although KOFST is the federation of independent scientific societies in Korea, it has been funded by the Korean government for supporting societies' work. KOFST supports journal publishing through a number of initiatives for both local journals not indexed in international databases and international journals, usually in English. Special tailored support for limited number of journals is available, including consulting or funding to create a journal masthead and instructions for authors; publication ethics training; style and format development; design of the

cover page; layout of text pages; journal policy development; manuscript editing; English proofreading; a manuscript management system; a journal Web site; production of JATS XML and CrossRef XML; establishment of ORCID, CrossMark, and Fundref; and regular training for editors or editorial board members.

After extensive experience providing journal support, KOFST realized that a full-text, open access platform like PMC is necessary for scholarly journals from Korea to broaden their visibility; therefore, it decided to construct ScienceCentral, a clone of PMC for all the fields of science. The three prominent differences from PMC are as follows: first, it includes all scientific fields including technology, engineering, life science, and multi-disciplinary science; second, it includes journals in all languages in the world if the abstract, tables, figures, and references are also provided in English; third, it accepts only society or non-profit institutional journals [2]. It is a new experiment in full-text, open access databases. It is still uncertain how many open access, non-profit, society or institutional journals will participate in ScienceCentral. Since the cost of JATS XML production is very reasonable nowadays, the burden on publishers will be manageable. KOFST can fully fund JATS XML production for scientific journals from Korea. Another purpose of ScienceCentral is to help local or regional languages to survive in scientific writing. It is well-known that the language of science is English; however, if the full-text of other languages is presented in JATS XML format, readers with other mother tongues should be able to access the content as the technology for computerized translation develops, and thus the local language of both the authors and the readers can survive as tools for propagating scientific information for humanity.

How to display the alphabets of Korean and other languages via JATS XML

Korean authors usually write papers in English for international readers; however, when writing articles aimed at local readers, they write in Korean and submit to Korean-language journals. In 2012, JATS became the standard for academic journal XML accredited by the National Information Standards Organization (NISO). The most prominent advantage of JATS document type definitions (DTDs) is that they encompass any language in the world that can be encoded in unicode transformation format (UTF)-8, an ISO encoding standard that can represent every character of all the languages of the world as well as other special characters. Archiving Korean-language content no longer presents any major technical or economic challenges, owing to the availability of the JATS DTDs. The implementation of JATS in Korean language

publications is very simple since there is a language code structure in a JATS DTD. It can accommodate a variety of languages besides Korean very easily by adding the language attribute, “xml:lang Language.” For example, the following declaration added to the XML of an article can display Korean characters in that article.

```
< article
  article-type = “research-article”
  dtd-version = “1.0”
  xml:lang = “ko”
  xmlns:mml = “http://www.w3.org/1998/Math/MathML”
  xmlns:xlink = “http://www.w3.org/1999/xlink” >
```

If any languages are substituted, the language code can be changed, for example, xml:lang = “ja-Hani” (Japanese written in Kanji [Hanzi, Hanja, Han]), “mn” (Mongolian), “war” (Philippines), “vi” (Vietnamese), “zh” (Chinese), and “cs” (Czech). The language code can be inserted in many parts of the JATS XML to present multiple specific languages in a single article text. The DOIs for Korean-language references can also be displayed without difficulty when transformed from JATS XML to CrossRef XML.

Training programs for JATS XML

In August 2006, the first seminar on PMC XML production was organized by the present author, Sun Huh (former Chair of the Committee on Information Management, KAMJE) [1]. After that, the staff of KAMJE were able to produce PMC XML. Because JATS XML has a much larger number of tags than PMC XML, it is necessary to train the editors, manuscript editors, staff members of academic societies, and the staffs of printing and information technology companies. KCSE held a workshop entitled “Application of information technology to scholarly open access journal publication” on July 4 to 5, 2013 in Seoul with just under 100 participants. Three lecturers were affiliated with universities; two, with information technology companies; two, with manuscript editing companies; and one, with a university hospital library. The sessions and titles of each class for the two days were as follows:

Session 1: Journal XML

1. Journal Article Tag Set (JATS) 1.0: The new ISO standard for journal XML
2. How to produce full-text XML for journal papers in English as well as in Korean based on JATS 1.0
3. JATS 1.0 XML coding
4. How to construct XSL and/or CSS for the style sheet of XML files based on the data type definition (DTD)
5. Displaying HTML files from JATS XML files on the

Web

Session 2: CrossRef XML

6. Basics of CrossRef XML
7. CrossRef XML coding
8. How to convert JATS 1.0 XML to CrossRef XML
9. How to add hyperlink and cited-by functions using DOI
10. How to apply ORCID and FundRef via CrossRef XML
11. How to apply CrossMark

Session 3: Network and viewer tools

12. Adoption of social network services and QR codes
13. PubReader: A new journal paper reader tool convenient for Web, smart pad, and phone environments
14. How to produce mobile-friendly journal Web sites
15. N-Screen for accessing scholarly journals in a variety of device environments with one source

The purpose of this workshop was to promote each journal’s international competitiveness. Of course, the most important factor for reaching this goal is publishing papers with high quality content; however, most high quality papers end up in high impact factor journals. Therefore, local journals published by scholarly societies have difficulty in attracting the best quality papers. Even if the content itself is not the most rigorous, the visibility of such journals can be increased if they maximize other factors, such as the quality of the journal Web site and accessibility through international open access databases. In the pursuit of increased visibility, JATS XML is a basic factor to be implemented. From JATS XML, papers can be transformed to other XML formats very easily and can be formatted for display by viewers. The best example is PubReader.

To construct a full-text, open access journal Web site or database, the understanding of JATS XML is essential for editors. If the journal is published by an international commercial company, the editors do not need to learn it. However, most scholarly journals in Korea are published by societies, and their editors should understand the advancement of information technology applicable to their journals. If editors understand how to produce JATS XML and how to convert JATS XML for other XMLs or viewers, they will be able to have confidence to apply those technologies to their journals. This is a realistic expectation in Korea, as journal editors, like Koreans in general, have a relatively advanced understanding of information technology. It has become clear through ongoing training events that Korean journal editors have been able to understand the techniques and practice XML coding and conversion without difficulty from their own laptop. Participants have confirmed that CrossRef XML, the cited-by function through DOI, ORCID, CrossMark, and FundRef XML are easily produced from JATS XML. They learned how QR codes worked and the advantages of journal apps. PubReader was a

popular topic for the editors because it is the one of tools used for N-screen. N-Screen is a user-centered service, in which the user co-owns and executes multiple applications anytime and anywhere, without interruption through an advanced smart system. The linking services among devices such as smart TVs, smart phones, smart pads, laptops, and audio devices are easily realized through a variety of smart devices and cloud services. When editors understand these new standards and XML techniques, they can better manage engineers who will implement those services.

The Korean government's open access policy

KoreaPMC is not based on a requirement that journals based in Korea be open access. Rather, research that is funded by the Korean government must be published in open access papers. SSHC is still a preliminary project, so it is not mandatory that social science and humanities journals in Korea submit JATS XML-based full-text to the NRF of Korea. As for ScienceCentral, KOFST has a more authoritative policy, in that it plans, to favor the funding of open access journals. Since KOFST provides operating funds to scholarly journals, it can consider the use of JATS XML among its criteria for evaluating journals competing for limited funds. Thus, it is expected that most editors of science journals will make a sincere attempt to fulfill all the criteria recommended by KOFST to obtain funding. This may be the first case in the world of an open access journal policy established by a national government. Until this policy is established, however, depositing JATS XML files in ScienceCentral will remain optional.

Conclusion

This paper surveyed the current state and future directions of JATS XML-based, full-text databases in Korea; as well as the

technology, training programs, and Korean government policy on open access, full-text, XML-based databases. Based on the information presented here, it should be clear to the reader why such databases are being established in Korea now. The training program for editors described here may be a helpful model for other countries or other fields wishing to encourage accessibility to local research. It will be worthwhile to continue observing how the three new journal article databases in Korea—KoreaPMC, SSHC, and ScienceCentral—perform in the future.

Conflict of Interest

No potential conflict of interest relevant to this article was reported

Acknowledgments

This work is supported by the research grant of the National Research Foundation of Korea (policy research-2013-003-academic infrastructure promotion) and the Korean Federation of Science and Technology Societies (KOFST-2013), Government of the Republic of Korea.

References

1. The Korean Association of Medical Journal Editors. KoreaMed technical workshop, 2006 [Internet]. Seoul: The Korean Association of Medical Journal Editors; 2006 [cited 2013 Oct 25]. Available from: <http://kamje.or.kr/workshop/2006/0811/1.pdf>
2. Huh S. ScienceCentral: open access full-text archive of scientific journals based on Journal Article Tag Suite regardless of their languages. *Biochem Med* 2013;23:235-6. <http://dx.doi.org/10.11613/BM.2013.029>