Digital publishing for STM editors: JATS XML, DOI, CrossMark, FundRef, Text and Data Mining, ORCID, manuscript management system, and homepage

- 허선(한림대학교 의과대학 기생충학교실 교수, 과편협 기획운영위원장)

2015. 12. 4. 14:30 - 15:10
과총 지하 중회의실
과편협 편집인 워크숍

학습 목표: 편집인이 학술지 디지털 출판에 필요한 내용을 5 가지 이상 나열할 수 있어야 한다.

구체 목표
① 웹에서 학술지 제작 표준이 무엇인지 설명할 수 있어야 한다.
② 논문마다 고유 식별자를 붙이는 배경을 설명할 수 있어야 한다.
③ CrossMark, FundRef, Text and Data Mining을 정의내리고 구현 방법을 기술할 수 있어야 한다.
④ 연구자 개별식별자에는 어떤 종류가 있는 지 2 가지 이상 기술할 수 있어야 한다.
⑤ 투고관리시스템 도입의 장단점을 나열할 수 있어야 한다.
⑥ 학술지 누리집에 필수 요소를 5 가지 이상 기술할 수 있어야 한다.

머리말

우리나라에서 학술지 누리집이 처음 등장한 것은 1993년도 대한안과학회에서 발행하는 '대한안과학회지'이며 같은 해에 대한기생충학회에서 발행하는 Korean Journal of Parasitology 누리집이 등장하였다. 이후 국내 인터넷 환경이 세계에서 최고 수준으로 가장 빠르게 발전함에 따라 학술지 누리집은 물론 디지털 출판이 다양하게 등장하였다. 그러나 우리나라가 온라인에서 가장 빠른 속도를 자랑하는 네트워크를 지니고 있지만 디지털 출판에서 국제 표준은 국제 상업회사가 주도하고 우리나라 편집인은 따라가기 바쁘다. 계속 새 내용이 나오고 그것을 우리 학술지에 빠르게 적용시키는 것이 편집인의 임무이다. 우리가 언젠가는 국제 표준을 제안하고 주도할 수 있을 것이나 지금은 대부분이 학회지이고 편집은 모두 다 자발적인 봉사인 현실에서 학술지에서 새 기술을 선도하기가 쉽지 않다.

이 시간에는 참석자가 모두 novice editor 라고 여기고 각 학술지마다 반드시 구현하여야 할 디지털 출판 관련 내용을 정리하는 시간이면 충분하다. 또한

이런 내용을 대부분 알고 있으므로 자신의 지식이 어느 정도인지 점검하는 시간으로 삼으려고 한다. 강의가 아닌 team-based learning 으로 진행하여 개인 역량을 점검하고 나서 팀에서 문제를 해결하도록 한다.

Individual Readiness Assurance Test

1. 정기간행물의 고유 식별자는?

1) ISSN
2) ISBN
3) DOI
4) CrossMark
5) ORCID

2. 학술문헌의 고유 식별자는?

1) ISSN
2) ISBN
3) DOI
4) CrossMark
5) ORCID

3. 현재 CrossRef에서 권장하는 DOI 표기법은?

1) http://dx.doi.org/10.3352/jeehp.2015.12.52
2) dx.doi.org/10.3352/jeehp.2015.12.52
3) doi/10.3352/jeehp.2015.12.52
4) doi: 10.3352/jeehp.2015.12.52
5) 10.3352/jeehp.2015.12.52

4. 국제적인 연구자 개별 식별 체계 가운데 톰슨로이터사에서 운영하는 것은?

1) ResearcherID
2) Scopus author ID
3) DOI
4) CrossMark
5) ORCID

5. 학술 문헌은 온라인에 등장한 뒤, errata, retraction, corrigendum 등 다양한 출판후 변경 과정을 거칠 수 있다. 이런 경우 문헌의 최신 정보를 제공하는 체계는?

1) DataCite
2) FundRef
3) DOI
4) CrossMark
5) ORCID

6. 연구비 지원기관의 개별식별자는?

1) DataCite
2) FundRef
3) DOI
4) CrossMark
5) ORCID

7. 학술지 JATS XML 파일을 웹에 보여줄 때, HTML 요소인 폰트 크기나 색깔 등을 변경하려고 할 때, 어느 부분을 변경하여야 하나?

1) database schema
2) eXtensible Stylesheet Language Transformer
3) eXtensible Stylesheet Language
4) Data Type Definition
5) Cascading Style Sheets

Group Readiness Assurance Test

Individual Readiness Assurance Test를 팀이 같이 해결한다.

Further application

1. 우리 학술지는 다음과 같은 디지털 출판 관련 사항을 구현하는 지 점검하여 보고, 구현하지 못한 것이 있으면 방안을 제시한다.

항목	구현 여부	도입연도	구현하지 않은 경우 구현 예정 여부
JATS XML			
PubReader			
epub 3.0			
DOI			
Cited-by function			
CrossMark			
FundRef			
Text and Data Mining			
ORCID			
투고관리시스템			
학술지 누리집 독립 도메인			

2. 다음 중 우리 학술지 누리집에서 차림표(menu)나 작은차림표(submenu)로 제공하는 영문 정보는? (국문으로 제공하는 것은 해당하지 않음)

항목	구현 여부	구현하지 않은 경우 구현 예정 여부
Aims and scope		
Editorial Board		
Current issue		
Archiving (Back issue)		
Instructions to authors		
Research and publication ethics		
Copyright policy		
Article processing charge or submission fee		
Subscription info		
Contact info		
Manuscript management system		

Journal article tag suite (JATS) XML

미 국립의학 도서관, National Center for Biological Information에서 제정한 ISO 웹 문헌 표준이다. 학술지 출판과 학술 단행본 출판에서 2013년부터 JATS XML은 ISO 표준으로 자리 잡았다. 우리가 브라우저 통하여 읽거나 보는 것은 HTML 형태이나 XML에서는 문서의 형태 뿐 아니라 속성을 보여준다. 즉, 어느 부분은 제목이고 어느 부분은 저자, 소속이며 어디부터 서론이고 참고문헌인지를 일일이 tag를 주어 속성을 부여한다. 사람의 예를 들면 검은 머리, 갈색 눈동자, 키와 같은 보이는 정보 이외, 국적, 학력, 경력 등, 개인의 눈으로 보이지 속성을 제공하는 방법이다. 학술지에는 다양한 요소가 있으므로 개별 요소를 일일이 tag 로 구별한다. 예를 들어 소속은 <affiliation>Hallym University</affiliation> 이렇게 표기한다.

JATS XML로 archiving 하는 데서 나아가 PubReader, epub 3.0 과 같은 다양한 형태로 보여줄 수 있고 나아가서 PDF로 쉽게 전환도 가능하다. 이런 다양한 형태는 단순히 컴퓨터 화면에서 나아가 손전화나 smart tab 에서도 쉽게 볼 수 있다는 점에서 이제 모든 편집인과 발행인, 출판사, 정보회사에서는 필수로 다루어야 하는 언어이다. XML이 web 표준이 된 것이 1996년, PMC XML이 등장한 것이 2000년, 우리나라에서 전문 XML 제작을 도입한 것이 2006년, 국내 학술지가 PMC에 처음 등재된 것이 2008년이다. 이제 PMC XML을 기반으로 한 JATS XML을 통하여 의학 분야 뿐 아니라 전 과학 분야 나아가서 전 학문 분야까지 포용할 수 있고, 초기에 영어만 다룰 수 있었던 환경에서 나아가 전 세계 모든 문자어를 다룰 수 있는 JATS XML 로 우리나라 과학 학술지와 단행본을 네트워크 세상에서 경쟁력있는 형태로 선보여야 할 것이다. 2015년 현재는 의학 분야가 가장 많이 JATS XML을 다루고 있으나, Open Access Korea나 ScienceCentral 등 의학 이외 다양한 분야에서 작업하고 있음을 알 수 있다.

이렇게 JATS XML 로 제작할 때 큰 장점은 앞에 언급한 PubReader, epub, PDF 전환 뿐 아니라 학술지에 필수인 DOI, CrossMark, Fundref, Text and Data Mining, ORCID 같은 다양한 국제 표준 형식으로 손쉽게 변환이 가능하다는 점이다. 변환 도구가 있어서 클릭 한번 누르면 생성되므로 매우 편리하다. 또한 Directory of Open Access Journal 에 학술지 등록시킨 뒤 DOAJ XML로 메타자료를 기탁할 때 손쉽게 DOAJ XML로 변환 가능하다. 이외도 어떤 XML 이던지 다양한 형태로 변환할 수 있으므로 JATS XML이 학술지 출판에 기본 구조가 되었다. 즉, JATS XML은 이제 선택이 아니라 디지털 출판과 온라인 출판에서 필수이다.

2012년까지 PMC XML을 국제수준으로 다룰 수 있는 회사가 국내에 두 군데에 불과하였으나 2015년에는 다수 회사가 어렵지 않게 다룰 수 있다. 또한 영문학술지이거나 영문 단행본인 경우에는 인도, 파키스탄, 미국 등의 여러 회사에 외주를 주어 작업할 수도 있어서 국제적으로 과학학술 논문 JATS XML 제작 경비는

영문 논문 편당 50 불 수준으로 충분하다.

과편협 특별회원 단체 가운데 JATS XML을 Table XHTML, MathML, ChemML 포함하여 제작할 수 있는 곳은 아래와 같다. 국내 이외도 여러 회사가 제작할 수 있다.

Medang-Inforang Inc. http://www.medrang.co.kr/
M2Community Co.,Ltd. http://publishing.m2comm.kr/
KoreaScholar http://www.koreascholar.com/
SONJO http://www.sonjo.co.kr/

Digital Object Identifier (DOI)

Digital object identifier (DOI)란 온라인상의 디지털 지적 재산에 부여하는 알파벳-숫자 기호 체계이다. 우리나라 국민에게 주민등록번호를 부여하여 어디로 이사 가도 주민등록 전산망을 통하여 어디 거주하는 지 늘 찾을 수 있는 것과 같이 학술논문도 고유의 웹 주소가 있다면 그 주소를 통하여 어디서 볼 수 있는 지 늘 끊임없이 찾아갈 수 있을 것이다. 즉 특정 DOI 주소가 그 학술논문의 웹사이트와 연결되어 있어 DOI는 바뀌지 않으므로 웹사이트 주소가 바꾸어도 언제든지 찾아갈 수 있다. 예를 들어 참고문헌에서 바로 해당 원문으로 찾아갈 수 있다. DOI는 prefix 라는 발행처의 고유 번호와 suffix라는 발행처에서 임의로 부여할 수 있는 번호로 구성한다. 학술문헌의 DOI 는 CrossRef (http://www.doi.org/)에서 다루므로 여기에 가입을 하여야 한다. DOI를 부여하면 cited-by 기능을 구현하여 어는 논문에서 우리 학술지 인용하였는지 알 수 있다. JATS XML 제작하면 CrossRef XML 제작은 배우 쉬우며, sponsoring publisher가 따로 있으면 CrossRef에 편당 1 달러를 기탁한다. sponsoring publisher를 JATS XML 제작 회사가 맡으면 연회비를 CrossRef에 낼 필요가 없다.

CrossMark

1. CrossMark 란 무엇인가?

CrossMark는 CrossRef에서 운영하며 2012년 4월 27일 첫 서비스를 시작하였다. 이 서비스는 온라인상에서 내용을 수정하고, 오류를 고치고, 철회하는 등 출판물의 변경사항을 알리고, 최종본을 파악할 수 있게 하는 것이다. 이 서비스는 다양한 출판 정보를 제공한다. 예를 들면, 출판 이력, 부록(supplement) 위치, 원문 접근 정책(access policy), 연구비 지원기관 정보, 전문가심사 정보 등이다. 운영 경비는 CrossMark를 등록할 때마다 받는 미화 0.2 달러로 충당한다. 이 비용은 DOI를 기탁할 때 내는 1달러와는 별개로 지불한다. 2010년도까지 발행한 과거 분은

편당 0.02 달러이다. CrossMark는 CrossRef DOI를 기탁한 모든 논문에 붙일 수 있다.

2. CrossMark 는 어떻게 구현하는가?

CrossMark XML을 작성하는 방법은 다음에서 소개할 FundRef XML 제작에서 같이 Appendix에서 다룬다. 연구자는 사용할 줄 아는 것으로도 충분하고, 학술지에서 구현하기로 결정하면, CrossMark XML 제작과 기탁 등의 기술적인 내용은 국내 출판사나 학술지 정보 처리 회사에서 추가 경비 없이 쉽게 구현할 수 있다.

FundRef

1. FundRef 정의

연구비지원 기관 고유번호(funder identification) 서비스로 출판물에서 연구비 정보를 기술하는 국제 표준 양식이다. 2013년 5월 28일 정식으로 출범하여 서비스를 시작하였다.

2. 참여하는 기관은?

http://www.e-sciencecentral.org/funder/에서 참여하는 연구비 지원기관을 찾을 수 있다. 이 가운에 우리나라 연구비 지원기관으로 한국대학교육협의회 산하 200 개 대학교와 기타 정부 기관, 제약회사와 100대 기업 가운데 금융업을 제외한 회사 emd 393 곳이 등록되었다. 예를 들어 Hallym University는 FundRef Uniform Resource Identifier (URI)가 http://dx.doi.org/10.13039/501100002632이다.
http://dx.doi.org/10.13039/501100002701은 Ministry of Education, Republic of Korea의 URI이며,
http://dx.doi.org/10.13039/501100002994는 Ministry of Knowledge Economy, South Korea의 URI 이다. 등록되지 않은 연구비 지원기관이 fundref.registry@crossref.org에 메일을 보내어 정보를 제공하거나 기존 정보를 수정할 수 있다.

3. 왜 이런 작업이 필요할까?

표준화된 연구비지원에 대한 이름이나 메타자료가 없어서 text mining이나 분석이 어려운 현황을 개선하고, 연구비 지원 기관도 결과물 확인을 쉽게 할 수 있도록 하는 것이 필요하다. 이런 이유로 미국 National Institute of Health에서는 Medline에 반드시 그 연구가 미국 정부 연구비 지원 받은 것인지 여부를 기록하게 하였고 연구비 번호를 기록한다. 그럼으로써 정부에서 연구비 지원한 논문을 쉽게 추적할 수 있고 또한 국민에게도 연구비가 이렇게 효율적으로 쓰였다는 것을 알릴 수 있다. 발행인도 출판하는 논문의 연구비 지원 기관 추적을 쉽게 할 수 있으며 연구자

소속 연구 기관에서도 구성원의 학문 업적에 대한 주요 지원기관을 쉽게 확인하는 것이 필요하기 때문이다.

4. 어떻게 FundRef가 작동할까?

FundRef Registry에서 표준화한 연구비지원기관 이름을 제공하면, 투고관리시스템 (manuscript management system) 제작사에서 논문 투고 과정에 FundRef Registry를 포함시킨다. 발행인은 저자가 funder_name(연구비지원기관 이름), funder_identifier(고유번호), award_number (연구비 번호)를 원고에 기술하도록 투고 규정에 추가한다. 이후 연구비지원기관 정보가 출판사에 전달된다. 발행인이 CrossRef에 연구비지원기관 정보를 FudnRef XML로 제공하면. 연구비지원기관은 CrossRef를 검색하여 지원한 결과로 나온 논문의 DOI와 메타자료를 받는다.

5. FundRef 서비스로 누가 어떤 혜택을 받는가?

연구비지원기관은 지원 결과를 추적할 수 있으며, 저자도 투고 시스템에서 쉽게 입력 가능하다. 대학이나 연구기관은 구성원의 연구 성과를 파악할 수 있고, 학술지 발행인은 논문마다 연구비 제공 기관을 분석할 수 있다. 정부는 국민에게 연구비 사용 결과물을 더 투명하게 알릴 수 있다.

CrossRef Text and Data Mining (TDM)

1 정의

CrossRef CrossRef Metadata API를 통하여 DOI가 부여된 학술논문 원문에 접근하여 데이터마이닝 연구에 활용할 수 있도록 하는 기능으로 학술지나 단행본 사업 모형과 무관하게 비영리 목적이다. 연구자나 대중이 모두 무료로 활용할 수 있다. http://www.crossref.org/tdm/index.html에 자세한 내용이 있다.

이런 내용은 일부 국가에서는 법 집행으로도 강제하고 있다. 예를 들어, 영국 기업혁신기술부(Department for Business, Innovation & Skills)와 영국 지식재산청(Intellectual property Office)은 2014년 10월 비영리 목적으로 TDM을 이용할 수 있도록 지적재산권법을 개정하였다. 출판사(publishers)나 aggregate는 연구자가 TDM 활용을 불합리하게 방해하지 못하도록 하였다.

2. 참여 방법

별도의 등록을 하는 과정도 없으며 참여에 경비가 드는 것이 아니고 단지 학술지 누리집에서 개별 논문마다 TDM을 설정하여 독자에게 도움을 줄 수 있다.

3. 학회지에서 구현 방법

API를 이용하여 구현하는데 기술적인 자세한 내용은 http://www.crossref.org/tdm/index.html 에서 찾을 수 있다. 국내 회사에서 이런 기능을 어렵지 않게 구현하고 있다. 예를 들어 *Science Editing* (http://www.escienceediting.org/)에서 논문마다 TDM을 활용할 수 있도록 링크를 걸어 CrossRef TDM API 통하여 검색할 수 있도록 구현하였다.

4. 연구자는 어떻게 활용할까?

http://search.crossref.org/에서 원하는 주제로 검색을 하면 DOI가 있는 논문이 결과로 나오고 그 논문을 누르면 원문을 볼 수 있다. 이때 비록 상업출판사 학술지라도 원문을 TDM을 통하여 접금할 수 있도록 허락한 것은 전문을 볼 수 있고, 제한을 둔 곳은 원문 접근이 안된다. 이 접근은 출판사가 정한다. 물론 국내 과학 학술지는 대부분 open access/free access 이므로 DOI 만 부여하면 검색하여 국내 논문이 모두 나타난다.

ORCID

1. ORCID 정의

Open Researcher & Contributor ID (ORCID)는 개별 연구자를 구별하는 영구적인 디지털 고유번호이다. 논문을 투고하거나 연구비를 신청할 때 연구자의 연구업적을 쉽게 알 수 있게 한다. 또한 연구자 고유번호 등록 시스템을 유지 관리하는 기구를 부르는 말이기도 하다. 핵심 기능은 고유번호(unique identifier)를 얻고, 활동 기록을 관리하는 등록시스템과 시스템 간에 연결과 인증을 지원하는 Application programming interface (API) 두 가지이다. API는 응용 프로그램에서 사용할 수 있도록, 운영 체제나 프로그래밍 언어 기능을 제어할 수 있게 만든 인터페이스를 뜻한다.

2. 등록과 취급 정보

등록시스템은 개별 연구자에게 무료로 제공되고 ORCID를 받아서 연구 활동을 관리할 수 있고 다른 연구자의 업적을 검색할 수 있다. 기관도 회원이 될 수 있는데 회원이 되면 기관의 정보를 ORCID 고유번호와 연결하고 ORCID 자료를 갱신하고 ORCID에서 최신 정보를 받을 수 있으며 구성원이나 학생을 ORCID에 등록할 수 있다. 2012년 10월 16일 서비스를 시작하여 2014년 4월 6일까지 643,048 명이 등록하였다. ORCID가 취급하는 정보는 민감하지 않은 것으로 이름, 이메일, 기관명과 연구 활동이다.

3. ORCID에 연구자 등록하기

http://crosstech.crossref.org/2013/01/easily-add-publications-to-your-orcid-profile.html에 설명이 있다. DOI 번호받지 않는 논문 예들 들어, 자신의 석박사 학위 논문 또는 어떤 논문이나 단행본도 수작업으로 등록할 수 있다. 만약 어느 연구 업적도 없는 경우라도 일단 자신의 이름, 교육 배경, 직장 등을 입력하면 등록할 수 있다. 그러므로 학부생이나 대학원생도 미래 연구 업적 등록을 위하여 바로 등록할 수 있다. https://orcid.org/register에서 회원 등록한다. 등록한 연구자인 경우 ID, passwd를 치고 들어가면 입력한 개인 이력 정보가 나오고 업적이 보인다. http://orcid.org/0000-0002-8559-8640 를 누르면 보이는 것과 같다. 연구자는 DOI 에 등록된 자신의 업적을 늘 갱신할 수 있다. CrossRef Metadata Search (http://search.crossref.org/), Europe PubMed Central (europepmc.org), ResearcherID, Scopus to ORCID 등에서 업적을 가져와 등록시킬 수 있다.

4. ORCID를 학술지 누리집에서 제공하는 방법

ORCID에 저자가 등록하여 그 번호를 투고 원고에 제출하고, 학술지에서 ORCID를 제공하기로 결정하였으면 누리집(homepage)에서 ORCID 정보를 제공하여 쉽게 저자의 정보를 확인할 수 있도록 하여야 한다. 학술지 누리집에서 Journal Article Tag Suite (JATS) extensible markup language (XML) 제작하여 제공하는 경우 JATS XML 파일에 ORCID 정보를 어떻게 기술하는 지 또는 hypertext markup language (HTML) 파일에서 어떻게 기술하는 지 Appendix에서 다룬다. 국내에서도 ORCID를 제공하는 학술지가 2013년 2월 처음 등장한 이후(http://jeehp.org/) 많은 학술지가 받아들였다.

투고관리시스템

이 시스템은 연구자가 투고하면 편집인과 심사자가 온라인에서 심사를 하여 출판단계까지 다루는 시스템으로 최근에는 모든 국제 상업 출판사가 이런 시스템을 이용하고 있으며 국내 학회지도 대부분 활용하고 있다. 국제적으로 다양한 제품이 나와 있어 해외 회사 시스템을 이용하는 곳도 있고, 국내회사 시스템을 이용하는 학회도 있다.또한 한국연구재단에서 무료로 제공하는 서비스도 있다. 어느 시스템을 사용하던지 초기 구축에서 편집인과 프로그래머와 긴밀한 의사 소통으로 학회의 실정에 잘 맞는 시스템이면 어려움이 없다. 국내 제품도 국제 상업회사의 제품과 비교하여 전혀 손색없이 잘 만들고 또한 수정 요구에 빠른 의사 소통이 가능하므로 국내 회사의 제품을 사용하는 학회가 훨씬 더 많다. 현재 학회가 이런 시스템을 새로 구축한다면 주위 학술지 편집인의 사용 경험을 듣고 회사와 연락하여 추진하면 충분하다. 국내 회사는 대개 2 개월 내 빠르면 1 개월 내 개통이 가능하다. 앞에

소개한 JATS XML을 제작하는 특별회원사는 이런 제품 제작도 가능하다.

국제 수준 영문 누리집

영문 누리집이 없으면 전 세계에서 우리 학술지를 찾아와서 읽는 것이 거의 불가능하다. 우리가 학술지를 발행하는 것은 우리가 생산한 훌륭한 논문을 많이 읽고 도움 받으라는 것이 큰 목표이므로 당연히 학술지 누리집을 잘 꾸며야 한다.우리가 비록 아직 학술지 내용이 국제 최고 수준은 아니라고 하여도 적어도 누리집만큼은 최고 수준으로 키워야 한다. 이런 작업은 경비만 투자하면 어려움없이 구현할 수 있는 내용이다. 누리집에 다음과 같은 내용이 메뉴나 서브메뉴에서 필수로 들어가야 한다.

Aims and scope
Editorial Board
Current issue
Archiving (Back issue)
Instructions to authors
Research and publication ethics
Copyright policy
Article processing charge or submission fee
Subscription info
Contact info
Manuscript management system

맺는 말

이미 학술지 인쇄본 시장은 점점 축소되고 온라인 학술지 시장이 빠른 속도로 발전하고 있다. 이런 상황에서 JATS XML, DOI, CrossMark, FundRef, ORCID 등의 내용은 이제 더 이상 편집인과 무관한 일이 아니다. 국내에서는 대부분 과학 학술지에서 이런 내용을 적용하고 있다. 우리나라에서도 일부 출판사를 제외하고 여러 출판사는 이런 시장의 변화에 빠르게 따라가지 못하여 별도의 IT 기업이 주도권을 쥐거나 외국 회사의 고객 역할을 한다. 학술지 자료는 대개 다 전자 자료로 나타났고, 다양한 네트워크를 통하여 자료가 모두 연결될 수 있어 이용자에게 빠르게 원하는 정보를 제공할 수 있다. 이런 세상에서 새 편집인은 다양한 디지털 기술을 학술지에 적용하고, 나아가 앞으로는 우리나라에서 학술지 국제 표준을 제안할 수 있어야 한다. 다른 분야는 국제 표준을 쥐락펴락 하기도 하는데 유독 학술지는 늘 따라가기에 바쁘므로 이런 현실을 타개하기 위한 편집인과 관련 인사의 노력이 필요하다.