

XML을 활용한 고문헌의 원문디지털화 방안에 대한 연구

- 고문헌을 위한 DTD개발을 중심으로 -

A Study on Methods of Digitalization of Older Oriental Books by Using XML

리 상 용(Sang-Yong Lee)

초 록

본 논문은 1996년 W3C에서 웹문서의 표준으로 제안된 차세대 웹언어인 XML(eXtensible Markup Language)을 활용하여 고문헌의 원문을 디지털화하고 웹상에서 이용자들이 고문헌의 원문을 손쉽게 검색할 수 있도록 하기 위한 방안에 대해 연구한 것이다. 고문헌을 디지털화하기 위하여 일반고문헌을 위한 XML DTD(Document Type Definition)를 개발하고, 이 DTD를 토대로 실제로 XML 상에서 고문헌의 원문을 입력해 보고, 그 과정에서 발생하는 문제점과 기술적인 한계를 살펴보고, 향후 본 고문헌 XML파일의 국가적인 활용방안에 대한 시스템 모형을 구축하였다.

ABSTRACTS

This article is a study on methods of digitalization for older oriental books using XML(eXtensible Markup Language) proposed as a standard for Web documents by W3C in 1996. In this paper, to digitalize older books, I designed the XML DTD(Document Type Definition) for general older oriental books and input the texts of the older books on the basis of the XML DTD, examined the problems and limitations in light of the technical level these days, and established the model of the system for using XML files of older oriental books at a national level in the future.

키워드: XML, DTD, 고문헌, 고서, 디지털도서관, Older books, Digital library

이 논문은 2000년도 한국학술진흥재단의 지원에 의하여 연구되었음. (KRF-2000-003-C00476)

건국대학교 충주캠퍼스 인문학부 문헌정보학전공 부교수(sylee@kku.ac.kr)

논문접수일자 2003년 2월 25일

게재확정일자 2003년 3월 19일

참고문헌

- 강순애. 1999. 한국고문헌정보시스템의 구축 및 전망. 『한국문헌정보학회지』, 31(4) : 83-112.
- 권영하. 1998. PDF를 이용한 디지털 도서관 구축 방안. 『디지털도서관』, 5: 86-95.
- 김녹환. 2002. 『XML을 활용한 전자박물관시스템 구현에 관한 연구』, 연세대학교 대학원 석사학위논문임.
- 김영숙. 조성호. 2001. 『XML BIBLE』. 서울: 삼양출판사.
- 김 용, 남궁황. 1998. 전자문헌개발도구에 관한 고찰: SGML, HTML과 PDF를 중심으로. 『정보관리연구』, 29(1): 1-19.
- 김용규. 1998. 『XML을 이용한 원격강의용 강의 공간의 설계 및 구현』. 석사학위논문. 목포대학교 대학원.
- 김은중, 홍웅기, 이현주 외. 2001. 『전문가로 가는 지름길 XML』. 서울: 대림.
- 김천식, 임도빈 편역. 2000. 『XML』, 서울: 대림.
- 김현희, 장혜원. 1999. 디지털도서관 문서양식으로서의 XML과 HTML의 특성 및 검색기능 비교연구. 『한국정보관리학회지』, 15(2): 105-134.
- 김흥규. 2002. 한국문화의 국제적교류와 옛한글, 『유니코드정보환경과 옛한글』, 옛한글국제표준화심포지엄, 고려대학교민족문화연구원 비표준문자등록센터.
- 남영준. 1998. 전자도서관 구축에 따른 장서의 디지털 우선 순위에 관한 연구. 『도서관학논집』, 한국도서관정보학회, 29: 343-380.
- 리상용. 2000. PDF를 활용한 고문헌의 원문 디지털화 방안에 대한 고찰, 『한국문헌정보학회지』, 34(1): 132-153.
- 박병진. 1999. 『XML기반의 통합에디터 구현』. 석사학위논문. 서강대학교.

- 박준서. 1999. 『XML을 위한 웹브라우저 설계 및 구현에 대한 연구』. 석사학위논문. 송실대학교 대학원.
- 박형욱. 1998. 『쉽게 배우는 Acrobat3』. 서울: 비비컴.
- 선정일. 1999. 『Internet EDI 구현을 위한 XML 활용에 대한 연구』. 석사학위논문. 국민대학교 대학원.
- 손지혜. 1999. 『시문헌의 XML DTD 개발에 관한 연구』. 석사학위논문. 한양대학교 교육대학원.
- 심대익. 1999. 『IR을 지원하는 RDBMS를 이용한 XML 문헌 검색시스템』. 석사학위논문. 서울대학교 대학원.
- 옥영정. 1999. 고서검색시스템의 설계에 관한 연구, 『한국도서관정보학회지』, 30(4): 255-278.
- 유사라. 1997. 『하이퍼미디어 도서관 정보시스템』. 서울: 한국도서관협회.
- 유사라. 1996. 멀티미디어 정보관련기술과 표준안에 대한 고찰. 『정보관리학회지』, 13(2): 55-68.
- 윤희중. 1999. 행자부, 『행정 전자문서유통 표준안 마련』.
- <http://www.etnews.co.kr/etnews/new_jetnews_content?J199905070110104>.
- 이건식. 유니코드 기반 대규모DB개발에서 옛한글 구현방안, 『유니코드정보환경과 옛한글』, 옛한글국제표준화심포지엄, 고려대학교민족문화연구원 비표준문자등록센터.
- 이윤석. 1997. 『홍길동전 연구』, 대구: 계명대학교.
- 이용석. 1999. 『XML 문서 저장시스템의 설계 및 구현』. 석사학위논문. 한국외국어대학교 대학원.
- 이지영. 1998. 고서의 디지털화와 색인에 관한 연구, 『서지학연구』, 15: 359-384.
- 장수진, 손광덕. 1999. 『Adobe Acrobat 4 workshop』. 서울: 영진출판사.
- 장혜원. 1999. 『XML과 HTML로 제작한 웹문헌의 특성 및 검색 기능 비교 연구』, 석사학위논문, 명지대학교 대학원 문헌정보학과.
- 정지범. 1997. 『차세대 웹 문서 포맷 XML』.
- <<http://www.knti.co.kr/KntiWorld/PcLine/199709/709ef00.htm>>.
- 정희경. 1999. 『문서기술언어-SGML, HTML, XML』.
- <<http://mie.paichai.ac.kr/bbs/data/sxh.htm>>.
- 정희경. 1998. 『XML 가이드』. 서울: 도서출판 그린.
- 조성원. 1999. 『WWW의 현재와 미래: DHTML & XML』.
- <<http://members.iword.net/aster/web/dhtml-xml.html>>.
- 채진석. 1999. “인터넷 악보 표현을 위한 XML DTD 설계”. 『한국정보관리학회지』, 15(3): 37-52.
- 한국전산원. 2002. 『지식정보자원관리를 위한 XML DTD 표준』, 서울: 한국전산원.
- 한국전자책컨소시엄. 2001. 『제2차 한국 전자책(eBook) 문서표준화 포럼』.
- 홍윤표. 2002. 옛한글의 개념에 대하여, 『유니코드정보환경과 옛한글』, 옛한글국제표준화심포지엄, 고려대학교민족문화연구원 비표준문자등록센터.
- 정영미, 안현수. 1998. 『전자도서관구축론』. 서울: 구미무역.
- 장수진, 손광덕. 1999. 『Adobe Acrobat 4 workshop』. 서울: 영진출판사.
- Adobe. 1999. Acrobat developer conference.
- Jon, Bosak, Jon. 1997. “Overview: XML, HTML, and all that.”
- <<http://www.oasis-open.org/cover/bosakWWW6-over.html>>.
- Bray, Tim, Hollander, Dave and Layman, Andrew. 1999. “Namespaces in XML.”
- <<http://www.w3.org/TR/REC-xml-names/>>.
- Bray, Tim, Paoli, Jean and Sperberg-McQueen, C. M. 1998. “XML 1.0.”
- <<http://www.w3.org/TR/REC-xml>>.
- Chinenyanga, T. T., Kushmerick. 2002. And Expressive and efficient language for XML information retrieval. 『Jasist』, 53(6): 438-453.
- Frentzen, Jeff. 1997. “Will you be ready when/if XML replaces HTML?”
- <<http://www.zdnet.com/zdnn/content/pwkw/1501/266546.html>>.
- Funke, Susan. 1998. XML Content Finally Arrives on the Web!. 『Searcher』. 6(9): 63-68.
- ____. “XML: How It Will Affect Content, Knowledge Management, and Information Professionals.” Proceedings of the 20th National Online Meeting. Medford, NJ Information Today, Inc., 123-131.
- Garshol, Lars Marius. 1999. “Free XML software.”
- <<http://www.stud.ifi.uio.no/~larsga/linker/XMLtools.html>>
- Levitt, Jason. 1997. “XML is the future of HTML.”
- <<http://www.global.ebscohost.com/cgi-bin...axrecs=10/reccount=250/startrec=21/ft=1>>

Light, Richard. 1998. Presenting XML Indianapolis, IN : Sams.net Publishing.

_____. 1998. "XML Pointer Language (XPointer)."
 <<http://www.w3.org/TR/WD-xptr>>

Luk, Robert W.P., H.V Leong, Tharam S. Dillion, Alvin T.S. Chan, W. B. Croft and J. Allen. 2002. A survey in indexing and searching XML documents *Jasist*, 53(6): 415-437.

Maruyama, Hiroshi, Uramoto, Naohiko and Tramura, Kent. 1998. "The Power of XML: XMLs Purpose and Use in Web Applications."
 <<http://www.ibm.com/XML/>>

Meggison, David. 1998. 『Structuring XML Documents』. Upper Saddle River, NJ: Prentice Hall PTR.

Rath, Hans Holger. 1998. "XML: Chance and Challenge for Online Information Providers." *Proceedings of 22nd International Online Information Meeting*, Oxford: Learned Information Europe Ltd., 339-345.

Sperberg-McQueen, C. 1998. XML and the future of Digital Libraries. *Journal of Academic Librarianship*, 24(4): 314-317.

Lovegrove, W. S. and Brailsford, D. F. 1995. "Document analysis of pdf file: methods, results and implications", *Electronic Publishing*, 8(2/3) : 207-220.

Smith, P. N. and Brailsford, D. F. 1995. "Towards structured, block-based PDF:", *Electronic Publishing*, 8(2/3) : 153-165.

<http://www.adobe.com>

<http://www.adobepdf.com>

<http://www.dlibrary.or.kr>

<http://www.dyenc.co.kr>

<http://www.infoconamerica.com>

<http://koreanhistory.or.kr/>

<http://www.unicode.org/>

<http://xml.webmania.co.kr/>

<http://www.xmlportal.co.kr/>

<http://www.xmlgo.net/>