

연속간행물 및 단행본 게재기사의 입력형식 표준화에 관한 연구

A Study on the Input Format Standards for the Contributions to Serials and Monographs

한종엽, 한국해양연구소

Jong-Yup HAN, Korea Ocean Research & Development Institute

문헌 데이터베이스를 구축함에 있어서 서지데이터의 입력형식에 관해서는 국내의 표준지침이 마련되어 있지 않다. 국내외의 관련 표준과 선행연구를 검토하고, CSA, OCLC, 국회도서관, 연구개발 정보센터 등의 입력데이터 서지요소를 분석하여 학술지의 기사와 단행본 형태 출판물의 기고문, 특허문헌 등을 대상으로 구축하는 데이터베이스의 실용적인 입력표준안을 설계하여 제시하고자 한다.

1. 서론

각종 출판물을 대상으로 정보자원(1차정보)에 대한 상세한 사항, 즉 제목, 저자명, 발행자, 초록, 주제명 등과 같은 2차정보를 서지데이터라 하며 이러한 서지데이터를 수록한 데이터베이스를 서지데이터베이스라고 통칭하는데, 본고에서는 학술지의 기사와 단행본 형식의 출판물에 수록된 논문, 특허문헌 등을 대상으로 구축하는 데이터베이스라는 관점에서 문헌 데이터베이스라고 명명하였다.

정보화시대에서 각 분야의 표준화 문제는 국제화, 세계화 차원에서 필수적인 선결요소로써 선진각국의 국가적인 주요 관심사가 된지는 이미 오래전이고 많은 분야에서 구체적인 활동이 이루어지고 있다. 특히 문헌정보학 분야에서의 표준화는 정보의 교류와 공동활용, 그리고 이용자의 편의를 고려할 때 필수적인 요소가 된다. 그러나, 문헌 데이터베이스를 구축함에

있어서 서지데이터의 입력형식에 관해서는 체계적인 표준화 작업이 이루어져 있지 않다. 현재 국내외의 문헌 데이터베이스는 그 입력데이터의 구성과 요소가 다양각색인데 그 이유는 통일된 표준이 없기 때문이다. 이에 대해서는 국제적으로도 현재의 디지털환경의 데이터베이스 구축에 부합하는 완전한 표준지침이 미비한 실정이고, 국내의 각 기관들은 70-80년대에 작성된 미국과 일본의 표준과 관련규격에 따라 적당히 이를 참조하여 준용하고 있는 실정이다. 이에 따라, 문헌 데이터베이스의 구축 및 이에 수록된 정보의 상호교류와 유통에는 많은 문제점이 있다. 따라서 본고에서는 국내외의 관련표준을 비교·검토하여 단행본 형식의 출판물에 수록된 기고문이나 학술지 등의 기사 및 특허문헌에 대한 데이터베이스를 구축하는데 필요한 실용적인 입력표준안을 설계하고자 한다. 특히, 학술지는 학회, 협회, 연구기관 등에서 주로 간행되고 있으며, 이들 자료의 데이

CSA는 미국의 Cambridge Scientific Abstracts, Inc.에서 제공하는 데이터베이스이다. CSA는 환경, 해양, 컴퓨터, 기계, 재료, 공학전반, 의학, 생명과학, 농학 등 과학기술 전 분야에 걸친 학술지, Proceedings의 논문단위 서지정보데이터를 포함하고 있는 세계 최대의 과학기술 데이터베이스로서 47종의 주제분야로 구성되어 있다. 검색은 표준검색(저자, 기사제목, 키워드검색)과 확장검색(키워드, 필드코드검색)으로 선택해서 검색을 수행할 수 있고, 데이터항목은 기사명, 저자명, 공저자명, 편자명, 저자성주소, 원문수록처, 회의명, ISSN, ISBN, 초록, 본문언어코드, 초록언어코드, 출판년도, 자료유형, 주제명, 분류번호, 입력일자, 주기사항, 원서명, 발행자, 입력기관번호, 인증자 등으로 구성되어 있다.

3.1.2 OCLC (Online Computer Library Center) FirstSearch

OCLC는 전세계 도서관 소장정보에 대한 접근성을 높이고 최소한의 정보비용으로 효율적으로 정보공유를 하는 것을 목표로 설립되었으며, 미국을 포함한 전세계 70여개국의 3만여 도서관이 참여하고 있는 비영리 연구기관이다. OCLC는 도서관의 편목작업과 원문서비스를 지원하고자 「FirstSearch」를 통하여 「World-Cat」 등 60종이상의 데이터베이스 검색서비스를 시행하고 있다. OCLC의 서지레코드는 대상 데이터 유형에 따라 레코드 작업을 위한 템플릿을 선택하여 해당 내용을 가변장필드 태그로 기술하게 된다. 서지레코드는 미국의회도서관 등 참여기관으로부터 수집된다. 「First-Search」의 검색은 해당주제 데이터베이스를 선별적으로 또는 전체를 선택하여 저자명, 제목, 정보원, 주제명 등으로 제한하여 검색할 수 있는데, 기본검색과 확장검색 및 전문가검색으로 검색양식을 지정하여 검색을 수행할 수 있다. 데이터항목은 기사명, 저자명, 잡지명, 권호, 페이지, 발행일자, 출판사, DDC번호, LC번호, ISSN, 언어코드, 초록, 원문정보 등으로 구

성되어 있다.

3.2 국내

3.2.1 국회도서관 기사색인 DB

국회도서관은 도서 및 기타 도서관자료 및 문헌정보를 수집·정리·보존하고 도서관 봉사 및 입법조사 분석업무를 수행함으로써 국회의 입법활동을 지원하는 기관으로 국가서지를 작성·배포하고 있다. 국회도서관의 「정기간행물기사색인」은 1945년 이후 국내에서 발행된 정기간행물에 수록된 기사중에서 학술적 가치가 있다고 판단되는 기사를 선정·색인하여 구축한 데이터베이스로서 현재 90만건이상의 서지레코드가 수록되어 있다. 데이터항목은 기사명, 저자, 잡지명, 권호, 페이지, 주제코드, 입력일자 등으로 구성되어 있으며, 실제 검색항목은 기사명/잡지명, 저자명, 키워드, 발행자와 보조항목인 발행년을 제한하여 검색할 수 있다.

3.2.2 연구개발정보센터 과학기술문헌 (SATURN) DB

연구개발정보센터는 국가과학기술정보 DB구축 및 유통을 위하여 1999년에 독립법인으로 설립된 정부출연연구기관으로 과학기술문헌정보, 도서정보, 정부연구보고서정보, 과학기술사실정보, 인력정보, 해외동향정보 등을 제공하고 있다. 과학기술 문헌정보데이터베이스는 과학기술계 16개 전문정보센터가 분야별로 수집·소장하고 있는 연속간행물, 회의록(Proceedings), 보고서 등에 포함된 연구문헌에 대한 데이터베이스로서, 키워드 검색, 학술지검색, 논문제목 등으로 검색을 수행하고, 데이터 항목은 입력일자, 자료유형, 언어코드, 원문소장기관, 소장기관등록번호, 보고서번호, 분류번호, 개인저자명, 공동저자명, 기관저자명, 저자성소속기관, 기사명, 대등서명, 출판사항, 형태사항, 규격번호, 규격갱신기호, 주기사항, 초록, 주제명, 잡지

터베이스화 및 유통의 원활화에 대한 사회적 요구가 매우 빠르게 증대되고 있다. 따라서 연속 간행물 및 그 게재 기사의 구성과 요소에 관한 표준화는 이러한 사회적 요구와 관련하여 가장 기본적으로 이루어져야 할 작업이다(고영만 등 1996).

2. 국내외 관련표준 및 선행연구

연속간행물의 구성과 요소에 대한 국외의 관련표준으로는 「ISO 8-1977 : Documentation - Presentation of periodicals」와 「SIST 07-1985 : 學術雜誌の構成とその要素」가 있으며, 연속간행물 게재기사의 구성과 요소에 대한 국외의 관련표준으로는 「ISO 215-1986 : Documentation - Presentation of contributions to periodicals and other serials」와 「SIST 08-1986 : 學術論文の構成とその要素」가 있다. 단행본, 연속간행물, 단행본 및 연속간행물의 게재기사, 특허문헌 등에 대한 서지적 참조의 내용·형식과 관련된 표준으로는 「ISO 690-1987 : Documentation - Bibliographic references - Content, form and structures」가 있고, 전자문헌과 그 구성요소에 대한 서지적 참조 관련표준으로는 「ISO 690-2-1997 Information and documentation - Bibliographic references - Part 2 : Electronic documents or parts thereof」가 있으며, 서지적 참조 기술방식의 표준으로는 「SIST 02-1984 参照文獻の書き方」이 있다. 「ISO 690-1987」은 종래의 인쇄매체에 대한 서지적 참조 표준이며, 「ISO 690-2-1997」은 전자매체와 그 구성요소에 대한 서지적 참조 표준이다. 또한, 「SIST 02-1984」은 문헌을 참조할 때에 記述하여야 할 요소, 선정, 표기법에 대한 원칙과 지침을 주기 위한 기준이다. 「ISO 690-2-1997」의 서문의 주(註)에서는 「ISO 690」은 「ISO 690-1」로 개정될 것이라고 언급하고 있는데, 현재까지 이에 대한 개정판은 발간되지 않고 있다. 개정·발간될 것으로 예상되는 「ISO 690-1」은

보다 현대적 수준에서 디지털환경에 부합하는 다매체의 서지적 참조 표준이 될 것으로 기대되어 상당한 관심의 대상이다.

국내에서는 이에 대응할만한 표준이 아직 확립되지 않고 있는데, 근래 들어서 국가차원의 서지DB구축의 일환으로 그 표준의 필요성이 증대함에 따라, 한국데이터베이스진흥센터와 첨단학술정보센터에서 위탁하여 연구과제를 수행한 두가지의 선행연구가 있다. 이는 고영만 등(1996)의 「연속간행물 및 게재기사의 구성과 표준화 연구」와 정준민(1998)의 「학술지논문 종합목록 입력기준에 관한 연구」이다. 고영만 등(1996)은 연속간행물의 구성과 요소의 분석, 연속간행물 게재기사의 구성과 요소의 분석, 관련 규격 분석 및 실태조사, 표준제안의 내용으로 연구를 수행했다.

또한, 정준민(1998)은 서지데이터의 각 요소별 특성을 분석하여 더블린 코어의 15개 요소 중 10개의 요소를 선정해서 입력기준을 제시하였다. 그러나, 고영만 등(1996)은 공저자가 2인이상인 경우에는 해당 기사에 대하여 기여도가 많은 사람을 필요에 따라 최소한으로 기재하도록 기준을 설정함으로써 이에 누락되는 공저자는 검색되지 않는다는 문제점이 나타난다. 또한, 정준민(1998)의 연구는 더블린 코어가 인터넷과 같은 디지털 환경에서 저자나 정보제공자의 인터넷자원 기술을 위한 최소한의 데이터 요소만을 포함한 메타데이터 기술방안으로써 이의 기본 데이터요소는 제정당시(1995년) 13개의 요소에서 1997년에 두요소를 추가하여 15개의 요소로 구성되었는데, 이 중 10개의 요소를 선정하여 인쇄매체를 포함하는 문헌의 서지 데이터요소에 그대로 적용하기에는 무리한 점이 있다.

3. 대표적인 문헌 데이터베이스

3.1 국외

3.1.1 CSA (Cambridge Scientific Abstracts)

명, 발행일자, 권호, ISSN, CODEN, 발행연도, 원문 이미지 정보 등으로 구성되어 있다.

4. 결론

국내외의 관련 표준과 선행 연구를 검토한 결과 현재의 문헌 데이터베이스 구축에 실용적으로 적용하기에는 미비한 점이 많다. 왜냐하면 고영만 등(1996)이 공저자가 2인이상인 경우에는 해당 기사에 대하여 기여도가 많은 사람을 필요에 따라 최소한으로 기재하도록 기준을 설정함으로써 이에 누락되는 공저자는 검색되지 않는다는 문제점이 나타나며, 정준민(1998)은 전자자료를 기술하기 위하여 만들어진 터블린 코어의 15개 요소 중 10개의 요소를 선정하여 일방 학술지 논문 및 기사에 적용하려고 설계하는데 있어서 무리한 점이 발견되었다.

그리하여 본고를 통하여 도출된 제안점은 다음과 같다. 우리나라의 문헌 DB 구축에는 검색 효율성과 적합성을 위해서 저자명, 단체명, 주제명 등에 대한 전거파일의 구축이 보다 체계적으로 이루어져야 한다. 정준민(1998)은 학술지명 전거와 저자명 전거에 대하여 언급하였고, 정옥경(1999)은 저자주제명일람표를 제시한 바 있다. 또한, 입력데이터의 다양한 서지요소는 완전수준으로 기술하여 DB 구축하고, 이를 통해 정보검색한 결과를 화면에 출력할 때는 이용자에게 적합한 요구수준의 정보를 Screen Display할 필요가 있다. 검색결과정보의 화면 출력방식에 대해서도 표준화의 관점에서 심도 있는 연구가 진행되어야 한다. 향후 다매체 환경에서 존재가능한 다양한 정보원의 DB 구축 및 통합검색을 위해 문헌 데이터베이스를 구축하는 경우에도 서지요소에 대하여 MARC와 준용할 수 있는 표시기호(Tag)를 사용할 필요성이 있다.

참고문헌

고영만 등. 연속간행물 및 계제기사의 구성과

요소의 표준화 연구. 한국데이터베이스진흥센터, 1996.

정옥경. 한국편목규칙의 표목부에 관한 연구. 중앙대학교 대학원, 1999. (박사학위논문)

정준민. 학술지 논문 종합목록 입력기준에 관한 연구. 첨단학술정보센터, 1998.

American Library Association. *The Future of the Descriptive Cataloging Rules*. Papers from the ALCTS Preconference, AACR2000, ALA Annual Conference, Chicago, June 22, 1995.

John J. Riemer, ed. *Cataloging and Classification Standards and Rules*. New York, The Haworth Press, Inc., 1996.

Pat Oddy. *Future Libraries Future Catalogues*. London, Library Association Publishing, 1996.

Ronald Hagler. *The Bibliographic Record and Information Technology*. Chicago, American Library Association, 1997.

P. Stoyanova. *Evaluating the content of bibliographic records for serials : the user's point of view*. Master of Library Science Research Project Report. Faculty of Information Studies. University of Toronto, Toronto, Ontario, 1998.