

서지 유틸리티와 CIP 데이터 활용에 관한 연구

- OCLC와 미국국회도서관 CIP 프로그램을 중심으로 -

A Study on the Bibliographic Utility and Its Use of CIP Data : With a Focus on the LC CIP Program and OCLC

윤 정 옥(Yoon, Cheong-Ok)*

〈 목 차 〉

1. 서론	3. 1 서지 유틸리티의 목표와 발전
2. 미국국회도서관의 CIP 프로그램	3. 2 서지 유틸리티와 CIP 데이터의 중요성
2. 1 CIP 프로그램의 개관	3. 3 서지 유틸리티와 CIP 데이터 업로딩
2. 2 CIP 프로그램의 대상과 절차	3. 4 서지 유틸리티와 CIP 데이터 개선
2. 3 CIP 데이터의 특성과 배포	4. 결론과 제언
3. 서지 유틸리티와 CIP 데이터의 활용	

초 록

이 연구에서는 미국국회도서관의 CIP (Cataloging in Publication) 프로그램을 중심으로 서지 유틸리티를 통한 CIP 데이터 접근, 배포, 개선 및 활용의 관계를 고찰하였다. LC CIP 프로그램은 많은 출판사의 참여와 지속적 성장, CIP 서지정보의 품질에 대한 신뢰와 즉각적 접근을 위한 서지 유틸리티 종합목록 데이터베이스의 활용, 출판 후 서지정보의 완전성과 정확성 개선을 위한 품질제어 노력 등의 특성을 갖는다. LC와 OCLC와 같은 서지 유틸리티 사이의 협력은 국립중앙도서관의 '출판시 도서목록' 프로그램의 제한점에 대해 시사하는 바가 크다. 국립중앙도서관이 국가의 대표 서지기관으로서 출판예정 단행본에 제한된 CIP 데이터 뿐 아니라 모든 유형의 국내간행 신구간 자료에 대한 정확하고 완전한 서지 및 전거 레코드에 대한 국내 도서관의 요구를 충족시켜야 할 필요성이 지적되었다.

주제어: 미국국회도서관, CIP 데이터, 종합목록, 공동편목, 서지 레코드 품질제어, 국립중앙도서관

ABSTRACT

This study examines how CIP program of the Library of Congress distributes, provides an access to, and helps upgrade CIP data through bibliographic utilities. LC CIP program can be characterized by the participation of many U.S. mainstream publishers, a continuing increase in the number of CIP data assigned annually, and an effort to immediately upgrade the completeness and accuracy of CIP data in cooperation with bibliographic utilities. Cooperation between LC and bibliographic utilities in performing quality control is worth noting, when compared with the limitation of CIP program of the National Library of Korea. It is suggested that the National Library of Korea should provide complete and accurate bibliographic and authority records for other types of new and old library materials published in Korea as well as CIP data for new monographic materials.

Key Words: Library of Congress, CIP data, OCLC WorldCat, Shared cataloging, Quality Control of Bibliographic Records, National Library of Korea

* 청주대학교 인문대학 문헌정보학 전공 교수 (jade@cju.ac.kr)
• 접수일: 2005년 5월 1일 • 최초심사일: 2005년 5월 30일 • 최종심사일: 2005년 6월 10일

I. 서론

국내에서는 ‘출판시 도서목록’이라고 불리는 CIP (Cataloging in Publication) 데이터는 일반적으로 한 나라의 대표 도서관이 CIP 프로그램을 시행하면서 그 나라에서 출판예정인 도서를 위해 작성하여 제공하는 기본적 서지 레코드를 말한다. CIP 데이터는 도서관, 출판사, 서적중개인 등 여러 이해당사자들이 도서의 신속한 처리를 위해 활용할 수 있을 뿐 아니라, 궁극적으로는 이용자로 하여금 빠르게 신간 및 출판예정 도서의 서지정보에 접근하게 하는 역할을 할 것으로 기대된다. 세계에서 처음으로 1971년에 CIP 프로그램을 시작한 미국국회도서관 (이하 LC라고 함)이 천명한 것처럼, CIP 프로그램의 목표는 출판사로 하여금 도서를 출판할 때 CIP 데이터를 저작권면에 삽입하여 도서관과 서적중개인들이 빨리 도서를 처리할 수 있게 하는 것이기 때문이다¹⁾.

우리나라의 국립중앙도서관이 실제로 CIP 프로그램을 시작한 것은 2002년 7월로 2005년 현재 이년 남짓한 역사를 지니고 있다. CIP 프로그램을 국내에 도입해야 될 필요성에 대해서는 일찍이 팔십년 대 초반부터 논의되어왔지만²⁾, 실제로 국립중앙도서관과 대한출판문화협회를 중심으로 한 한국형 CIP 제도의 도입방안에 관한 논의가 본격화된 것은 이십여 년이 지나서였다³⁾. 미국을 비롯하여 오스트레일리아, 브라질, 캐나다, 독일, 영국 등이 이미 칠십년 대에 CIP 프로그램을 시작하였던 것에 비하면 다소 늦게 시작되었다고 할 수 있다.

현재 국립중앙도서관의 CIP 프로그램은 그 연륜이 짧을 뿐 아니라, 지금까지 CIP 데이터를 부여한 실적이 그리 두드러지지 못한 것으로 나타나고 있다⁴⁾. 또한 CIP 데이터가 도서관이나 출판사들에 의하여 실제로 어떻게 활용되는지에 대해서는, 국립중앙도서관이 도서관의 홈페이지와 KOLIS-NET 연계 서비스를 통하여 “전국도서관에서 신속하게 공동 활용” 하도록 제공하고 있다고 주장하는 것⁵⁾ 외에는, 아직은 아무도 구체적으로 언급한 적이 없다. 그보다는 도서관 현장에서 정리업무를 담당하는 사서들은 아직 CIP 데이터가 활용가치가 없다고 단적으로 지적하기도 하는 형편이다⁶⁾.

국립중앙도서관에서 CIP 프로그램을 시작하고자 할 당시, 오동근과 여지숙은 국내에서 CIP 프

1) Library of Congress, "Purpose of CIP," 2004, <<http://cip.loc.gov/purpose.html>> [인용 2005.3.7]

2) 홍순영, "Cataloging in Publication (CIP) 및 그 도입에 관한 소고," 한국문화연구원논총 (이화여자대학교), 42 (1983), pp.97-110.

3) 오동근, 여지숙, "한국형 CIP 제도의 도입방안에 관한 연구," 한국문헌정보학회지, 제 35권, 제 2호 (2001), pp. 59-74.

4) 조왕근, "국립중앙도서관 e-CIP 와 연계한 도서관 업무개선 방안" 출판과 도서관 그리고 서비스(2005년 국립중앙도서관 '열린 정책 세미나' 자료집 1), pp.85-107. (주최: 국립중앙도서관 자료기획과 일사 2005년 3월 23일, 장소: 국립중앙도서관 대강당)

5) 국립중앙도서관, 출판시도서목록(CIP) 편람, 개정2판(서울: 동도서관 출판시목록센터 2004).

6) 김포옥, 노옥림, "목록 업무의 아웃소싱에 대한 인식도 연구" 情報管理學會誌 제21권 제 호(2004), pp.145-771; 윤정옥, "대학도서관 편목 업무 외주의 동향과 논쟁," 한국문헌정보학회지, 제38권, 제2호(2004), pp.119-136.

로그랩을 어떻게 시행할 것인지에 대하여 포괄적 연구를 수행하였다⁷⁾. 이들은 우선 LC를 중심으로 한 CIP의 등장배경과 역사를 고찰하였고, 미국, 캐나다, 영국 등 주요국가의 CIP 프로그램 현황에 대하여 CIP 대상 자료의 범위, 운영주체와 방식, 데이터 근거자료와 송부절차 등을 분석하였다. 이들은 우리나라에서 채택해야 할 모형으로 중앙집중식 국가 CIP 센터를 건립하여, 그 안에 CIP 부서, 수서부서 및 정리부서를 운영할 것을 제안하였다. 이들이 강조한 것은 국가문헌정보자원을 총괄하는 기관에서 국가서지센터를 운영하고, CIP 프로그램을 운영함으로써 국가문헌과 서지정보의 유기적 수집, 정리와 관리체제를 확립하자는 것이었는데, 이들이 제안한 이상적 모형의 체제는 아직 구현되지 못하고 있다.

최근 국립중앙도서관이 주최한 ‘열린 정책 세미나’에서 조왕근⁸⁾은 2002년부터 2004년 국내에서 출판된 도서통계와 국립중앙도서관의 CIP 부여실적을 비교하여, 2003년 한해 전체 CIP 부여대상 신간도서 33,440건 가운데 단 5.5%인 1,835건에만 CIP를 부여하였고, 2004년에는 대상 신간도서 33,909건 가운데 6.9%를 차지하는 2,348건에만 CIP를 부여하였음을 보고하였다. 이 같은 수치를 고려하면, 국립중앙도서관이 2002년을 시범사업의 해로 삼았고, CIP 프로그램을 단지 두 해 동안 진행했을 뿐이라는 점을 감안할 때, 프로그램의 성과를 아직 단적으로 평가하기는 어려우나 현재 출판사들의 참여와 CIP 부여율이 매우 저조한 것은 사실이라고 할 수 있다. 조왕근은 현행 CIP 제도를 시행하는 데 있어서 국가 CIP 센터의 필요성 및 운영정책에 대한 공감대 부족, 인력 및 예산 부족, 출판유통정보 시스템과 연계한 CIP 데이터 수집방안 부재, e-CIP 시스템과 타 시스템 간의 데이터 연동 방안 부재, 도서관 업무 자동화 시스템에서 사용할 표준 API의 개발 보급 필요, CIP 구축을 위한 서지도구의 정비 및 표준화 부족, CIP와 실제출판정보의 불일치 해소방안 부재 등과 같이 많은 문제점이 있음을 지적하였다.

불과 이년 남짓 진행된 국내의 CIP 프로그램이 안정되고, 활발하게 진행되며, 그 결과물인 CIP 데이터가 도서관, 출판사, 서적중개인에 이르는 다양한 관심집단에게 광범하게 인정되고 활용될 것을 기대하기엔 다소 이르다고 할 수는 있다. 그러나 미국을 비롯한 다른 나라에서 생산된 CIP 데이터가 다양한 방법으로 활용되고 있는 사례를 참조하면, 현행 국내의 CIP 제도가 간과한 측면을 새롭게 인식할 수 있을 지도 모른다. 이런 전제 하에 본 연구에서는 LC의 CIP 프로그램의 특성을 살펴보고, 국제적 서지정보의 공유조직인 서지 유틸리티를 통한 CIP 데이터의 활용 현황을 고찰하도록 한다. 특히 LC가 국가의 대표적 서지기관으로서 CIP 데이터에 근거한 서지정보의 생성, 공유 및 활용 체계를 주관하는 역할 및 신간 및 출판예정 도서의 서지정보원으로서 서지 유틸리티가 수행하는 역할에 주목하여 논하도록 한다.

7) 오동근, 여지숙, 전계논문.

8) 조왕근, 전계논문.

II. 미국국회도서관의 CIP 프로그램

1. CIP 프로그램의 개관

LC에서 CIP 프로그램을 1971년 처음 시작하면서 천명한 목표는 “출판사에게 전문적 편목 데이터를 제공하여, 데이터가 책에 인쇄될 수 있게 함”으로써 도서관들로 하여금 편목 비용을 절감하게 하고, 책이 독자들에게 신속하게 배포될 수 있게 한다는 것이었다⁹⁾. CIP 프로그램이 가진 목표는 전적으로 새로운 것은 아니었고, 이미 1950년대 말에 CIS (Cataloging in Source) 프로그램을 통해서 시도되었던 것이다. 당시 CIS 프로그램은 미국 내 출판사가 보낸 교정쇄를 이용하여 해당 자료가 출판되기 전에 목록을 작성하여 배포하려고 하는 것이었고, 그 목표나 처리과정 등은 지금의 CIP 프로그램과 크게 다르지 않았다. 불과 이년 가량 시행된 CIS 프로그램 자체는 LC에서 편목에 투자해야 하는 시간과 비용부담, 출판 전과 후의 서지 데이터 불일치 등과 같은 문제 때문에 부정적으로 평가되긴 하였지만, 도서관과 출판사들은 전반적으로 이 같은 프로그램이 필요하다는 점을 인정하였고, 이 같은 호의적 반응은 이후 CIP 프로그램이 새롭게 시행되는 기반이 되었다고 할 수 있다¹⁰⁾.

LC의 CIP 프로그램은 처음으로 시행된 1971년 7월 단 27개의 출판사만 참여했던 것에서 그해 말에 157개 출판사로 늘어났고, 1973년까지는 매년 3만 이상의 CIP 데이터를 생산할 것을 목표로 하였다¹¹⁾. CIP 프로그램 보고서는 1973년 7월에서 1974년 7월 사이에는 모두 640여 출판사가 참여하도록 성장하였음을 보여주고 있다¹²⁾. 2001년에는 모두 3,500 이상의 출판사가 여기 참여하여 사실상 거의 모든 주류의 미국 출판사가 참여하였고, 그 밖에 많은 소규모 출판사 및 상당수의 다국적 회사들도 참여하였다고 한다¹³⁾.

LC는 CIP 프로그램이 시행된 지 십 년 동안 어떻게 이용되었는가를 살펴보기 위하여 2,366개의 무작위로 선정된 미국내 학술, 공공, 전문 및 학교도서관을 서베이 하였다. 이 연구에서는 이 도서관들에 대하여 도서관 운영의 3개 영역, 즉, 수서, 편목과 공공 서비스 영역에 CIP 프로그램이 미치는 영향을 조사하고, 또한 별도의 대규모 도서관 집단의 CIP 데이터 활용 현황을 검토하였다고 한다. 이 연구에 따르면, 대부분의 도서관들이 CIP 데이터를 하나 이상의 영역에서 사용하고 있

9) Library of Congress, Processing Dept., *Cataloging in Publication: Progress Report, July-December 1971*, (Library of Congress: Washington, D.C., 1972) (PR-1, ED 060905) [초록].

10) 오동근, 여지숙, 전개논문.

11) Library of Congress, *Ibid*.

12) Library of Congress, Processing Dept., *Progress Report - Cataloging in Publication, July 1973 to June 1974*, (Library of Congress: Washington, D.C., 1974) (LC-CIP-5, ED 095933) [초록].

13) Marilyn McCroskey and Michelle R. Turvey, "What Is CIP and How Does It Benefit the School Library Media Specialist?" *Knowledge Quest*, vol.32, no.2(2003), pp.45-46.

며, 편목을 위해 가장 많이 사용한다고 하였다. 수서에서도 비교적 많이 사용되고 있지만, 공공 서비스 영역에서는 그리 많이 사용하지 않았다. 대부분의 도서관들은 CIP 데이터에 대하여 대체로 ‘호의적’이거나 ‘매우 호의적’인 의견을 가짐을 발견하였다¹⁴⁾. 이처럼 CIP 프로그램에 대한 긍정적인 반응은 1999년 말, 사천 이상의 출판사 및 관련기관들이 교정쇄를 제출하게 되고, 1998년 한해에만 51,792 CIP 데이터를 생산하였던 것과 같이 지속적으로 이 CIP 프로그램의 규모가 커지게 되는 기반이 되었다고 할 수 있다¹⁵⁾.

LC의 CIP 프로그램은 1999년 말, 2000년 1월에 출판 예정인 도서 *Write It Down, Make It Happen*의 CIP 데이터를 작성함으로써 백만 번째 CIP 기록을 달성하였는데, 당시의 CIP 부서장 존 셸리가 다음과 같이 말한 것은 고품질의 서지정보를 신속하게 제공할 수 있는 CIP 프로그램의 역할에 대한 자부심을 반영하는 것으로 보인다 :

“CIP 레코드는 -- 미국의 도서업계는 말할 것도 없고 -- 미국의 115,000여 도서관뿐만 아니라 전 세계의 도서관들을 지원한다... 수요가 많은 책들을 위해 최고 수준의 목록을 생성함으로써, 그리고 책이 간행되기 전에 그렇게 함으로써, 세계의 도서관계에 매년 엄청난 액수의 돈을 절감한다. 돈 보다 더 중요한 것은 CIP 목록이 제공하는 국가의 지식과 창의성에 대한 접근이다.”¹⁶⁾

2. CIP 프로그램의 대상과 절차

LC의 CIP 프로그램은 그 대상이 단행본 도서임을 명백히 하고 있다 즉, 출판 예정인 모든 도서가 CIP 데이터를 받을 수 있는 것은 아니고, “표제면에 미국의 지명을 출판지로 포함하고 있는 출판예정 (forthcoming) 단행본”만이 자격을 갖게 된다. 또한 CIP 프로그램에 참여할 자격이 있는 출판사도 제한되어 있다. 그 자격은 “미국의 도서관들이 널리 입수할 가능성이 가장 많은 서명을 출판하는 미국의 출판사”로 한정되어 있고, 서적 벤더, 배급자, 인쇄자, 프로덕션 회사 및 기타 중개자들은 참여할 수 없도록 되어 있다¹⁷⁾.

CIP 프로그램의 절차를 간략히 살펴보면, 우선 자격이 있는 도서를 위해, 참여하는 출판사들은 완성된 CIP 데이터 신청서와 더불어 도서의 전문 (畵文) 혹은 적어도 ‘핵심적인 요구 자료’를 LC에 제출해야 한다. 도서관의 CIP 담당부서는 신청서와 본문의 완전성과 자격 여부를 검토하고, 제

14) Kathryn Mendenhall, *A Survey of the Cataloging in Publication Program. Final Report* (Library of Congress: Washington, D.C., 1982) (ED 221215) [초록

15) Library of Congress, “Cataloging Milestone: Library Resources 1 Millionth CIP Record,” *Library of Congress Information Bulletin*, vol.58, no.11(1999), p.259.

16) Library of Congress, *Ibid.*

17) Library of Congress, 2004, *Ibid.*

어번호 (LCCN, Library of Congress Control Number)를 부여하여, 적절한 주제분과의 편목부서로 전달한다.

목록사서는 기술편목을 수행하고, 주제표목과 LC 분류번호 및 DDC 번호를 부여하며, 필요하면 LC 이름전거 파일에 새로운 이름전거 레코드를 작성하여 올리거나, 주제전거 파일에 새로운 주제표목을 작성하여 올리는 등의 업무를 수행한다. 이렇게 완성된 CIP 데이터는 출판사에게 송부되며, 출판사는 이 정보를 표제지 뒷면에 인쇄하여 포함시키게 된다. 이전에 CIS 프로젝트가 실패하였던 원인 가운데 하나가 CIS를 책에 인쇄하는 데 소요되는 추가의 비용이었다고 하였으나¹⁸⁾, 지금은 그것이 문제가 되지 않는 것으로 알려지고 있다. LC는 또한 CIP의 기계가독형 버전인 MARC CIP를 세계의 대형도서관, 서지 유틸리티, 및 서적 벤더에게 배포한다. 이 기관들 가운데 많은 수가 이 데이터를 사용하여 도서관계에 출판예정 도서를 공지하고, 도서주문을 가속화시킨다고 한다¹⁹⁾.

한편 CIP 프로그램에서는 출판사가 출판된 도서를 CIP 부서에 보내야만 CIP 처리가 비로소 완료된 것으로 여기는데, 도서관 직원은 출판된 도서를 받고, 면수와 크기 같이 CIP 데이터에서 결여된 데이터 요소를 목록 레코드에 추가하고, 레코드에 이미 있는 요소들이 출판물을 제대로 기술하고 있는지 확인한다. 이때 출판된 도서에 나타난 정보에 기초하여, 본서명, 부서명, 시리즈, 저자나 주제 등을 변경할 수도 있고, 확인을 마친 후에 레코드의 기계가독 버전을 다시 배포하게 된다. 이 마지막 확인단계는 서적 벤더와 자동화된 시스템을 가진 도서관들에게 특히 중요한 것으로 여겨진다고 한다. 특히 개별 도서관들은 LC가 도서가 출판된 즉시 CIP 레코드를 확인해 줄 것을 기대한다고 한다. 그렇지 않으면, 각 도서관이 개별적으로 이 레코드를 갱신하고 확인해야 하기 때문이다²⁰⁾. 또한, 많은 도서관들이 벤더로부터 도서를 구입하면서 최종적으로 확인된 CIP 레코드를 반영하는 목록 카드 세트를 함께 보내 줄 것을 원하기 때문에, 서적 벤더들도 LC가 반드시 확인한 서지 레코드를 요구하는 추세라고 한다²¹⁾.

LC의 보고서에 따르면, 출판사가 CIP 프로그램에 교정쇄를 제출하고 CIP 데이터를 받기까지 평균 12.7일이 걸리고, 모든 CIP 레코드의 78%가 14일 이내에 완성된다고 한다²²⁾. 이들 가운데 요약문을 추가해야 되는 아동 및 청소년용 도서와 국립의학도서관에서 의학전문주제표목 (MeSH)이 부여되어야 하는 의학학 분야의 도서들의 경우에는 CIP 레코드 생성이 다소 지연됨을 보고하고 있다.

CIP 프로그램에 참여하기 위하여 출판사들은 아무런 비용도 지불하지 않아도 되지만, 도서가

18) 오동근, 여지숙, 전개논문.

19) Library of Congress, 2004, *Ibid.*

20) McCroskey and Turvey, 2003, *Ibid.*

21) Library of Congress, 2004, *Ibid.*

22) Library of Congress, Bibliographic Division, *Annual Report: Fiscal Year Ending September, 30, 2004*, <<http://www.loc.gov/catdir/bad04.pdf>> [인용 2005.3.7]

출판된 후에 반드시 한 부의 기증본을 제출하도록 되어 있다. 물론 그것이 주된 목적이 아니긴 하지만, LC는 CIP 프로그램을 통하여 입수하는 도서로 인하여 막대한 재정적 이익을 얻고 있다. 2004 회계연도의 LC 서지접근부 (Bibliographic Access Division)의 연차보고서는 1971년 미국이 세계 최초로 CIP 프로그램을 시작한 이래 삼십여 년 사이에 모두 1,280,995책에 대한 편목 데이터를 생산했고, 그 가운데 2004년 회계연도에만 53,349책에 대한 CIP 편목 데이터를 생산했음을 보고하고 있다. 이와 더불어 Electronic Preassigned Card Number Program (EPCN)을 통하여 CIP 프로그램에 자격이 되지 못하거나, 자격이 되지만 너무 늦게 교정쇄를 제출하여 출판 전에 완전한 CIP 편목 데이터를 얻을 수 없었던 28,290 책에도 LCCN을 부여하였다고 한다. 이렇게 CIP 프로그램과 EPCN을 통하여 LCCN을 부여받은 도서는 모두 출판된 후에 한 부를 반드시 CIP 담당부서에 제출하게 되는데, 2004년 한해에 이러한 경로로 모두 91,747책이 입수되었고, 그 가치는 5백여만 달러 이상인 것으로 추정되고 있다²³⁾.

결국 LC의 CIP 프로그램은 출판사, 서적 벤더, 도서관들에게 도움이 될 뿐 아니라, LC 자체에도 막대한 이익이 된다는 점에서 주목할 만하다. 그리고 CIP 데이터를 통해 궁극적으로는 이용자들이 국내에서 출판되는 신간 및 출판예정 도서의 서지정보에 신속하게 접근할 수 있게 된다는 점이 중요하다고 할 수 있다.

3. CIP 데이터의 특성과 배포

CIP 데이터는 기본적으로 서지 레코드 (bibliographic record)이다. <그림 1>에서 Adrienne

<p>Library of Congress Cataloging-in-Publication Data Tannenbaum, Adrienne Metadata solutions : using metamodels, repositories, XML, and enterprise portals to generate information on demand / Adrienne Tannenbaum. p. cm. Includes bibliographical references and index. ISBN 0-201-71976-2 1. Information technology. 2. Metadata. I. Title</p> <p>T58.5 .T37 2001 005.7-dc21</p> <p style="text-align: right;">01031646</p>

<그림 1> LC CIP 데이터 블록의 예

23) Library of Congress, Bibliographic Division, 2004, *Ibid.*

Tannenbaum의 *Metadata Solutions* (Boston: Addison-Wesley, 2002)의 표제지 뒷면에 수록된 LC CIP 데이터의 예에서 보는 것처럼, 기본적으로 서명, 저자명, 참고문헌과 색인 등 주기사항, ISBN, LC 주제표목 (LCSH), LC 분류번호, DDC 번호 및 LC 제어번호 (LCCN)를 제공하고 있다. CIP 데이터에는 발행지, 발행자, 발행연도 등의 정보, 면수나 삽도 등의 정확한 형태사항 및 내용 정보의 일부가 결여되어 있지만, 그럼에도 불구하고 CIP 데이터가 도움이 되는 이유는 도서관들로 하여금 도서를 입수하는 즉시 얻을 수 있는 기본적 서지정보를 바탕으로 하고, 실물 (items in hand)을 검토함으로써 얻을 수 있는 정보를 추가하여 완전한 서지 레코드를 생성할 수 있게 하고, 따라서 개별적으로 편목에 소요하는 비용과 시간을 절감할 수 있게 하기 때문이다.

사우스웨스트 미주리 주립대학의 교수이자 사서인 McCroskey와 Turvey는 다소 오래 된 통계이고, 데이터를 미국의 학교도서관 미디어센터에 한하고 있긴 하지만, 1991년 LC의 서베이를 인용하여 CIP 데이터를 목록의 기초로 사용할 때, 미디어센터 당 연 평균 \$3,865를 절감할 수 있다고 하였다²⁴⁾. 이 같은 비용의 절감은 모리스 등²⁵⁾이 아이오와 주립대학 도서관에서 제시한 것과 같이 레코드 당 오리지널 편목과 카피 편목 수행 시 소요되는 비용으로 산출된 것은 아니지만, 편목에 투입되는 시간이 줄어들면 줄어들수록, 목록에 완성된 서지 레코드를 올리고, 도서를 서가에 배치하기까지 걸리는 시간이 단축되어, 결국은 이용자가 보다 빠르게 자료에 접근할 수 있는 가능성이 높아지게 된다는 점을 고려할 때, 이들이 CIP 데이터를 활용하도록 강조한 것은 마찬가지로 설득력 있는 주장이라고 할 수 있다.

앞에서 언급한 것처럼, CIP 데이터는 두 가지 형태로 배포되는데, 책의 저작권 면 (copyright page)에 인쇄되는 CIP 데이터 블록과 기계가독형 버전인 MARC CIP가 있다. CIP 데이터 블록은 LC가 출판사에게 보낸 CIP 데이터를 도서의 저작권면에 인쇄한 것이므로, 교정쇄 제출 당시의 서지정보가 고정된 채로 배포된다. 그러나 MARC CIP는 우선 도서의 출판 전에는 CIP 데이터 블록과 같은 서지정보를 수록한 MARC 포맷으로 배포되지만, 도서가 출판된 후 LC에 제출되면 수정 및 추가된 레코드로 갱신되어 다시 목록에 올려지게 되므로, 보다 최신 버전을 제공하게 된다는 특성이 있다²⁶⁾.

24) McCroskey and Turvey, 2003, *Ibid*.

25) Dilys E. Morris, Collin B. Hobert, Lori Osmus, and Gregory Wool, "Cataloging Staff Costs Revisited," *Library Resources & Technical Services*, vol.44, no.4(2000), pp.70-83.

26) McCroskey and Turvey, 2003, *Ibid*.

Ⅲ. 서지 유틸리티와 CIP 데이터의 활용

1. 서지 유틸리티의 목표와 발전

서지 유틸리티 (Bibliographic utility)의 기능을 CIP 데이터 활용과 관련지어 주목할 중요한 측면은 미국의 OCLC (Online Computing Library Center)와 RLIN (Research Libraries Information Network) 및 1998년 이후 OCLC에 통합되어 운영되고 있는 WLN (Western Library Network) 등의 서지 유틸리티들이 기본적으로 “협동편목을 조정하기 위하여 (to coordinate cooperative cataloging)” 생겨났다는 점이다²⁷⁾. 즉 서지 유틸리티가 구축된 동기는 협동편목을 통해 서지정보를 공유함으로써 정보와 자원에 대한 접근을 증대시키고 비용을 절감하는 것이었다. 도서관들은 국가적 혹은 국제적 서지 유틸리티를 통해 협동편목을 수행하고 종합목록 데이터베이스를 유지하여, 여기에서 한 목록자가 생산한 오리지널 서지 레코드를 비용으로 수많은 다른 목록자들이 카피로서 사용할 수 있게 함으로써 ‘규모의 경제’라는 효과를 달성할 것을 기대하였다²⁸⁾.

대표적 서지 유틸리티 가운데 1974년 미국의 하버드 대학, 예일 대학, 콜롬비아 대학 및 뉴욕 공공도서관 등이 구축한 RLIN이 현재에도 북미주의 연구도서관 그룹 (Research Libraries Group)을 중심으로 종합목록 데이터베이스를 유지하고, 협동편목, 상호대차, 문헌제공 (document delivery) 등을 통해 학술정보와 자원의 공유에 주안점을 두고 있다 그러나 1967년에 RLIN보다 먼저 생겨난 OCLC는 처음에는 오하이오주의 대학도서관들이 자원을 공유하고, 비용을 절감하기 위한 목적으로 결성되었으나, 지금은 전 세계 모든 유형의 도서관을 포괄하는 국제적 서지 유틸리티로 성장하였다. 처음에는 학술도서관만이 OCLC 회원이 될 수 있었으나 1972년에 모든 유형의 도서관에 개방되었고, 2005년 현재 84개국의 오만여 개 기관이 OCLC의 서비스를 이용하고 있는 정도로 성장하였다²⁹⁾.

초기의 OCLC는 MARC 데이터베이스에 대한 접근을 제공하고, 회원도서관들은 서지 레코드를 생성하여 데이터베이스에 올렸다. 회원도서관들은 필요한 경우에 서지 레코드를 찾고 검토하고, 필요한 대로 개별 기관의 명세에 맞추어 수정하였으며, OCLC에서 목록카드를 전자적으로 주문하여 사용하였다. 이것이 바로 온라인 공동편목 네트워크가 되었다³⁰⁾. 현재 OCLC는 편목과 메타데이

27) Deborah A. Fritz and Richard J. Fritz, *MARC21 for Everyone: A Practical Guide*(Chicago: American Library Association, 2003).

28) Clare B. Dunkle, “Outsourcing the Catalog Department: A Meditation Inspired by the Business and Library Literature,” *Journal of Academic Librarianship*, vol.22(1996), pp.33-43.

29) OCLC, “About OCLC,” <<http://www.oclc.org/asiapacific/en/about/default.htm>> [인용 2005.3.2]

30) Richard E. Rubin, *Foundations of Library and Information Science* (New York: Neal-Schuman Publishers, Inc., 2002).

터, 장서관리, 디지털화와 보존, e-Content, 참고정보봉사 및 자원공유와 ILL 서비스를 제공하는 한편, OCLC Research Institute를 통하여 정보공유와 접근에 대한 연구 및 개발을 수행하고 있다. 그러나 처음부터 OCLC가 발전하는 기반이 되었고, 지금도 가장 핵심적 역할을 하는 것은 온라인 공동편목을 통한 서지정보 공유라고 할 수 있다³¹⁾. OCLC는 서지 유틸리티가 지역이나 국내 도서관들 사이의 협력 뿐 아니라, 국제적인 서지정보와 자원의 공유 및 협력에 대한 요구를 충족시킬 수 있음을 보여주었고, Rubin이 OCLC를 “역사상 가장 두드러지는” 서지 유틸리티라고 표현한 것이 적절하다고 할 수 있다³²⁾. 다음에서는 서지 유틸리티의 보다 광범하고, 보편적인 기능을 이해하기 위하여 OCLC의 발전과 서비스를 중심으로 살펴보도록 한다.

2. 서지 유틸리티와 CIP 데이터의 중요성

서지 유틸리티가 출판정보의 유통에 기여하는 가장 큰 이유는 이들이 유지하는 온라인 종합목록 데이터베이스가 도서관정의 도구역할을 하며, 신간 및 출판예정 도서의 서지정보를 신속히 제공하는 기능을 한다는 것을 들 수 있다.

각종 도서관 및 정보센터의 장서관리 정책과 전략에 꾸준한 연구를 수행해 온 미국의 로올라 메리마운트 대학도서관장 Evans는 서지 유틸리티는 어떤 면에서 “국가서지의 대체 (replacements for national bibliographies)” 역할을 하며 도서관정의 도구로 사용된다는 점을 지적하였다³³⁾. Evans는 OCLC나 RLIN의 데이터베이스는 전 세계 수천 도서관이 소장한 수많은 레코드를 수록하고 있으므로 대부분의 도서관들이 장서관리를 위해 개별 국가의 국가서지를 접근할 필요가 없음을 강조하였다. 물론 소급적 작업을 위해서나, 특히 온라인 시스템에서 잘 반영되지 않고 있는 국가들을 위해 아직도 국가서지가 중요하긴 하지만, 대부분의 미국도서관에서는 이들만으로도 충분하다고 하였다. 그는 또한 자료주문을 위한 서지정보의 확인 작업을 위해서, 이들 온라인 데이터베이스가 종종 “가장 뛰어난 단일 정보원 (the best single source)”이 되기도 함을 지적하였다. 그 이유는 특정한 서명의 존재 여부를 확인할 수 있고, 적절한 도구가 있다면 그 정보를 도서관의 수서 시스템으로 다운로드 할 수도 있고, 차후에 해당 도서가 입수되고 편목되면 이 정보가 온라인 목록에서 서지 레코드 역할을 할 수 있기 때문이라고 하였다³⁴⁾.

OCLC는 종합목록 데이터베이스인 WorldCat이 매우 다양한 목적으로 이용되고 있지만, 주로 서지 레코드의 편목, 상호대차 자료의 검색, 및 도서관 장서에 추가할 목적으로 주문할 새로운 자료

31) Audrey N. Grosch, *Library Information Technology and Networks* (New York: Marcel Dekker, 1995).

32) Rubin, *Ibid.*

33) G. Edward. Evans, *Developing Library and Information Center Collections*, 4th ed.(Greenwood Village, Colorado: Libraries Unlimited, 2000), p.117.

34) Evans, *Ibid.*

를 식별하는 데 사용된다고 하였다³⁵⁾. 이 같은 주장은 Evans가 지적한 바와 같은 맥락에 있지만, OCLC가 특히 새로운 자료의 식별을 강조한 것은 이들의 목록에 CIP 데이터를 수록하고, CIP 데이터의 개선을 수행함으로써 도서관들로 하여금 신간 및 출판예정 도서의 서지정보를 신속하게 얻을 수 있게 한다는 사실에 바탕을 두고 있다. OCLC의 WorldCat을 통하여 얻을 수 있는 CIP 데이터를 실제로 생성하고 입력하는 것은 LC를 비롯한 여러 나라의 국가대표도서관들이다. 따라서, 서지 유틸리티와 CIP 데이터에 기반한 신간 및 출판예정 도서의 서지정보에 대한 접근은 궁극적으로는 한 국가의 CIP 프로그램이 어떻게 운영되는가, 그리고 그 CIP 데이터를 어떻게 활용하는가에 크게 의존한다고도 할 수 있다.

LC의 CIP 프로그램의 주요한 목표는 순수히 단행본 도서의 서지적 특성을 기술하고, 그럼으로써 도서관 목록에서 그 저작에 대한 접근을 가속화하겠다는 것이므로, 연속간행물, 전자도서, 컴퓨터 파일, 악보, 시청각자료, 각종 마이크로폼 등 도서 이외의 매체가 해당되고, 도서 형태로 되어 있어도 이미 간행된 도서이거나 선물용 도서, 수요에 따라 (on demand) 출판한 도서, LCCN이 미리 부여된 도서, 특정한 교과서 등 다양한 부류의 도서가 CIP 데이터 부여 대상에서 제외되고 있다³⁶⁾. 이처럼 CIP 프로그램이 특정한 부류의 도서만을 대상으로 한다는 것이 다소 제한적이라고 여겨질 수 있으나, 지금까지 대부분의 도서관 장서에서 도서가 차지하는 비중 혹은 중요성이 절대적이었다는 사실을 고려한다면 그 제한성이 크게 문제가 되지는 않는 것으로 보인다. 예를 들어, OCLC의 WorldCat의 서지 레코드 통계를 보면 도서의 비중이 압도적임을 알 수 있다. 물론 전 세계 자료유형별 출판 비율의 실제 수치를 인용 및 분석하면 도서 출판 자체에 대해서 보다 더 정확한 데이터를 얻을 수 있을 것이다. 그러나 도서관에 소장되어 있는 도서는 적어도 일정한 선정기준에 따라 선별된 자료라는 사실을 고려할 때, 전 세계 수많은 도서관의 소장 상황까지를 반영하는 종합목록의 도서 서지 레코드의 통계만으로도 충분한 정보를 얻을 수 있다.

2005년 1월 현재 OCLC의 WorldCat에 수록된 54,826,796 건의 고유한 서지 레코드 가운데 그 가운데 거의 4,600만 레코드 (83.79%)가 도서의 서지 레코드이다. 도서만이 아니라 연속간행물, 시각자료, 혼합자료, 지도, 음성녹음, 악보, 컴퓨터 파일까지를 망라하는 이 서지 레코드들은 모두 95,340만 건의 소장사항 ("Location of Items Cataloged")을 반영하고 있고, 그 가운데 도서의 소장사항은 87,082만 건에 달하고 있다³⁷⁾. 이 통계에 따르면 도서의 소장사항이 전체 소장사항의 91.34%를 차지하고, 각각의 고유한 서지 레코드가 대표하는 고유한 도서 한 권을 평균 19개의 도서관이 실제로 소장하고 있다는 사실을 보여준다. 다른 유형의 자료들은 모두 합하면 전체 서지

35) OCLC, *WorldCat Facts and Statistics*, 2005, <<http://www.oclc.org/asiapacific/en/worldcat/statistics/default.htm>> [인용 2005.3.2]

36) Library of Congress, *The Scope of CIP*, 2004, <<http://cip.loc.gov/scope.html>> [인용 2005.3.7]

37) OCLC, *WorldCat Statistics by Format*, 2005, <<http://www.oclc.org/worldcat/statistics/format.htm>> [인용 2005.3.11]

레코드의 16% 정도를 차지하고, 각 레코드가 대표하는 도서 이외의 자료를 소장한 도서관은 평균 9개 정도이다. 다시 말하면 한 도서의 오리지널 목록 레코드가 서지 유틸리티의 종합목록 데이터베이스에 올려지면 적어도 그 생성기관을 제외한 18개 도서관이 사용하고 있다는 것이다.

한편 이들 고유한 서지 레코드 가운데 OCLC 회원도서관들이 기여한 오리지널 레코드는 45,756,840건이며, 모두 6,898,147건이 LC가 생성한 오리지널 레코드이다(전체의 12.6%). 또한 LC가 입력한 서지 데이터를 기반으로 하여 회원도서관들이 생성한 오리지널 레코드는 2,171,809건에 이른다(전체의 3.9%)³⁸⁾. 즉, 전체 레코드의 16.5% 정도가 LC가 전체 혹은 부분적으로 서지정보 생성에 기여한 서지 레코드라는 것이다. 따라서 LC는 단일기관으로서 서지 유틸리티에 서지 레코드를 가장 많이 기여하는 기관이다. LC는 이처럼 선별한 도서의 CIP 데이터를 생성하고 이를 서지 유틸리티에 제공하여 회원도서관과 이용자들로 하여금 출판예정 도서의 서지정보에 접근할 수 있게 함으로써, 또한 출판된 도서를 입수하면 바로 해당 도서의 CIP 데이터를 개선하여 완전한 서지정보를 도서관들이 이용할 수 있게 함으로써 서지정보 공유와 활용을 촉진하는 역할을 한다. 결국 서지 유틸리티의 종합목록 데이터베이스는 국내 및 국제적 서지정보 공유와 활용을 위한 LC의 실질적 노력이 반영되는 도구라고도 할 수 있다.

3. 서지 유틸리티와 CIP 데이터 업로딩

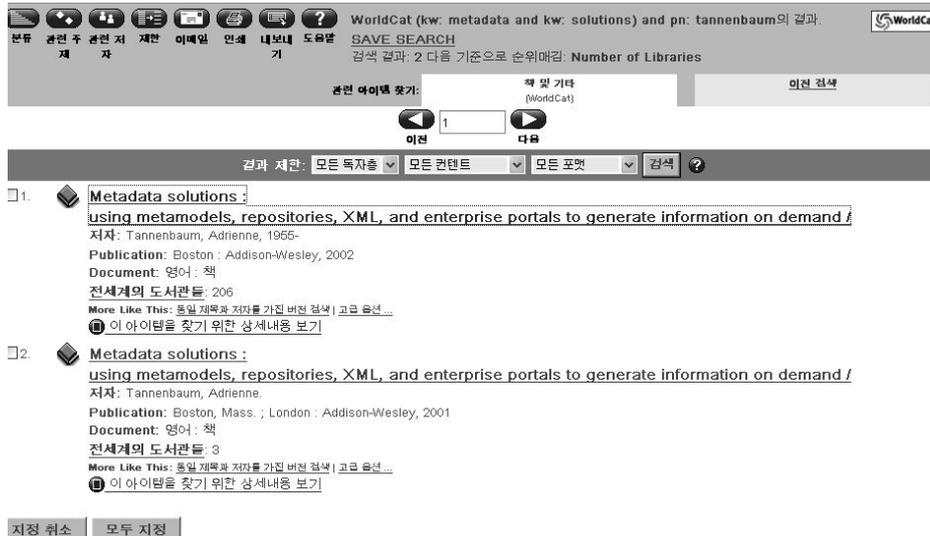
LC는 주요한 대형도서관 서지 유틸리티 및 서적 벤더에게 배포된 MARC CIP 데이터가 도서관계에 출판예정 도서를 공지하고, 도서주문을 가속화시키는데 사용된다는 점을 강조하고 있다³⁹⁾. 서지 유틸리티에 제공된 CIP 데이터가 출판정보 유통에 기여하는 것은 앞에서 지적한 바대로 OCLC의 WorldCat과 같은 종합목록 데이터베이스가 도서선정의 도구로 사용될 때라고 할 수 있다. 물론 LC에게서 직접 MARC CIP를 제공받는 대형도서관들은 서지 유틸리티를 거치지 않아도 주문을 위해 그 정보를 사용할 수 있을 것이다. 그러나 다른 도서관들은 일부의 정보가 결여되어 있기는 하지만 CIP 데이터의 형태로라도 목록 데이터베이스에 올려진 출판예정 도서의 서지 레코드를 빨리 접할수록, 그 정보를 주문을 위해 사용할 수 있는 가능성 또한 높아질 것이라고 추측할 수 있다.

다음에서는 실제 서지 레코드의 예를 들어, 출판예정 도서의 정보가 얼마나 빨리 서지 유틸리티에 올려지는가를 살펴보도록 한다. 앞서 예시한 Tannenbaum의 *Metadata Solutions*는 OCLC의 WorldCat에 <그림 2>에서 보는 바와 같이 두 개의 서지 레코드가 존재하고 있다. 그 가운데 한

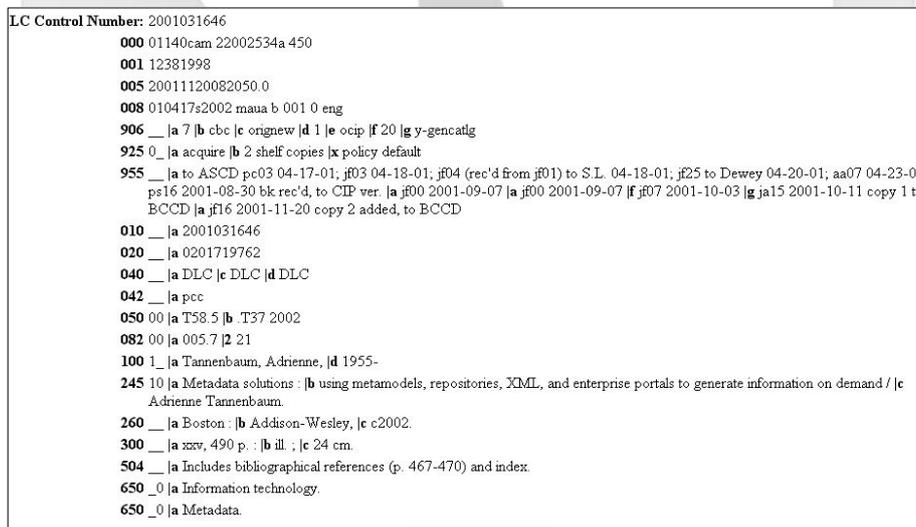
38) OCLC, *WorldCat Facts and Statistics*, 2005, <<http://www.oclc.org/asiapacific/en/worldcat/statistics/default.htm>> [인용 2005.3.2]

39) Library of Congress, 2004, *Ibid.*

서지 레코드 (OCLC Accession Number 46809645)의 최초 입력 일자는 2001년 4월 17일로 실제로 이 책의 표제지 뒷면에 기록된 최초 인쇄시점인 2001년 8월 (First printing, August 2001)보



<그림 2> OCLC WorldCat 검색 결과: Metadata Solutions 의 서지 레코드



<그림 3> LC의 OPAC 검색 결과: <그림 2>의 서지 레코드 1의 MARC 정보

다 앞서 있음을 알 수 있다. 즉, 책이 출판되기 이전, 다시 말하면 어느 도서관에서도 이 책의 실물을 입수하지 않았던 시점에서 이 책의 존재는 WorldCat의 서지 레코드를 통해서 알려지게 되었던

것이다. 언제 어느 도서관에서 이 서지 레코드를 최초로 생성하였는가를 보기 위하여 LC 목록에서 찾아본 MARC 레코드 (LCCN 2001031646 <그림 3> 참조)의 040 필드 (편목 정보원, Cataloging source) 서브필드 \$a (최초 편목기관, Original cataloging agency)는 최초의 레코드 생성기관이 LC (기관부호 DLC)이며, 이후 두 차례에 걸친 레코드 수정도 동 기관에 의하여 이루어졌음을 알 수 있다⁴⁰⁾. 이 레코드에 드러난 이 책의 이력은 LC가 출판 전 카피에 따른 CIP 데이터로 불완전수준 레코드를 입력하고, 이 책이 출판되고 실물이 제출된 후에 서지 레코드를 완전수준으로 개선했음을 보여준다. 또한 955 필드를 보면 LC 내부에서 이 책이 언제 어떤 부서에서 처리되었는지도 알 수 있는데, 여기에서는 2001년 4월 17일에 처음 이 책의 교정쇄가 입수되었고, DDC 부여를 맡은 부서를 비롯하여 몇 개의 부서를 거쳐 2001년 8월 30일에 출판된 책이 입수되어 CIP가 확인되었음을 알 수 있다.

이 서지 레코드 1은 <그림 1>에서 본 CIP 데이터의 상태에서 몇 개의 필드 혹은 서브필드가 수정 혹은 추가됨으로써 완전수준 서지 레코드로 개선되었음을 볼 수 있다. 우선 100 필드 (기본기입)에서 저자의 출생연도 (서브필드 \$d)가 추가되었고, 이는 전거 데이터베이스에서 확인된 형태를 의미한다. 또한 260 필드 (발행사항)에 발행지, 발행자 및 발행연도, 그리고 300 필드 (형태사항)가 추가 되었으며, 504 필드 (서지참조)에 참고문헌의 본문 정보가 추가되었다.

서지 레코드 1에 비해, 서지 레코드 2 (OCLC Accession Numer 48066773)의 최초 생성 일자 는 2002년 10월 21일로 서지 레코드 1보다 일년 반 정도 후이다. 이 레코드에 근거한 실물자료를 소장한 3개 도서관 가운데 두 군데는 영국의 국가대표도서관 (British Library)이고, LCCN도 없는 것으로 보아 이 서지 레코드는 영국에서 생성되었다고 추측할 수 있다. 결국 유추해서 보면 영국의 국가대표도서관과 같은 기관이 이 도서의 실물을 입수하여 편목하기 훨씬 전에 CIP 데이터는 이미 WorldCat에 올려져 있었고, 다른 도서관들이 도서선정을 위해 이 목록 데이터베이스를 참조했을 경우엔 이미 주문에 필요한 충분한 서지정보를 접할 수 있었을 것이다.

CIP 데이터가 서지 유틸리티에 실려 있는 예로는 현재 서적상에 출판예정 도서로 광고되어 있고, 출판 전 주문을 받고 있는 Harry Potter 시리즈의 여섯 번째 책인 *Harry Potter and the Half-Blood Prince*를 들 수 있다 (<그림 5> 참조). OCLC WorldCat에서 찾은 이 책의 서지 레코드는 그 입력일자에서 보는 바와 같이 2005년 1월 3일에 처음 입력되었고, 형태사항이 결여되어 있는 CIP 데이터 상태인데, LCCN이 없는 것으로 보아 LC가 아닌 다른 도서관에서 입력한 것임을 알 수 있다 (<그림 4> 참조). OCLC의 WorldCat에 여러 다른 나라의 국가대표도서관들도 서지 레코드를 입력하고 있으므로, 미국이 아닌 다른 나라에서 출판예정인 도서들의 CIP 레코드 또한 이 목록을 통해서 볼 수 있고, 이 또한 주문을 위해 사용할 수 있을 것이다.

40) MARC Field는 Library of Congress, *MARC 21 Format for Bibliographic Data: Field List*, 1999, English Ed. Update No.1 (Oct 2000) (Washington, D.C.: Library of Congress, 1999)를 참조.

title: Harry Potter and the half-blood prince /	
Author(s): Rowling, J. K.	
Publication: New York : Scholastic, Inc.,	
Year: 2005	
Description: p. cm.	
Language: English	
Standard No: ISBN: 0439784549	
SUBJECT(S)	
Descriptor: Potter, Harry (Fictitious character) -- Juvenile fiction.	
Wizards -- Juvenile fiction.	
Magic -- Juvenile fiction.	
Schools -- Juvenile fiction.	
Hogwarts School of Witchcraft and Wizardry (Imaginary place) -- Juvenile fiction.	
Other Titles: Half-blood prince; Harry Potter and the half blood prince	
Responsibility: by J. K. Rowling.	
Material Type: Fiction (fic)	
Document Type: Book	
Entry: 20050103	
Update: 20050118	
Accession No: OCLC: 57358293	
Database: WorldCat	

<그림 4> OCLC WorldCat: Harry Potter and the Half-Blood Prince의 CIP 레코드

The screenshot shows the Amazon.com product page for 'Harry Potter and the Half-Blood Prince (Book 6)'. The page includes a search bar, navigation links, and promotional banners. The main product information shows the book cover, a list price of \$29.99, a current price of \$17.99, and a 40% discount. It also features a 'Pre-order this item today' button and an 'Add to Wish List' button. The 'ITEM INFORMATION' section provides links for buying info, editorial reviews, and rating the book, along with the ASIN: 0439784549.

<그림 5> Amazon.Com에서 *Harry Potter and the Half-Blood Prince*의 출판예정 광고

4. 서지 유틸리티와 CIP 데이터 개선

서지 유틸리티의 역할은 CIP 데이터를 개선하는 것에서도 볼 수 있다. LC의 CIP 프로그램 절차에서 보는 것처럼 CIP 데이터는 책이 출판되기 전에 WorldCat에 올려지기 때문에, 완전한 목록 레코드에서 필요한 서지정보를 갖추지 못하고 있다. 특히 기본적으로 형태사항과 기타 요소들 가운데 출판 전에는 얻을 수 없는 것들이 있고, 간혹 서명이나 기타 요소들이 실제로 책이 출판되었을 때에는 다르게 나타날 수도 있기 때문에, 모든 CIP 데이터는 정확성과 완전성을 위해서 반드시 확인되고, 개선되어야 할 필요성이 인식되고 있다⁴¹⁾.

LC가 CIP 프로그램에서 제출된 도서의 실물을 최종적으로 확인한 후에야 종료된 것으로 여긴다고 한 것처럼, MARC CIP 데이터는 누가 개선하든 반드시 출판된 도서의 실물을 확인한 후에야 완전수준 (full level) 혹은 핵심수준 (core level) 서지 레코드로 개선되어야 할 것이 강조되고 있다. 1995년에 OCLC에서 CIP 업그레이드 프로그램 (CIP Upgrade Program)이 시작되기 이전에는 LC를 제외하고는 WorldCat 내에서 정규적으로 CIP 데이터를 개선한 기관이 없었고, 개별 도서관이 CIP 레코드를 입수할 때마다 개선해야 했다⁴²⁾.

OCLC CIP 업그레이드 부서와 그에 연계된 도서관 자료 밴더의 작업을 통하여, 1995년 11월부터 2003년 8월 사이에 242,000 레코드 이상이 개선되었고, 연 평균 거의 34,000 레코드가 개선되고 있다고 한다⁴³⁾. LC가 매년 대략 오륙 만 CIP 레코드를 생성한다고 할 때, OCLC의 업그레이드 프로그램이 전체 CIP 레코드의 절반 이상을 개선하고 있음을 알 수 있다.

OCLC가 공개하고 있는 CIP 개선을 위한 명세⁴⁴⁾를 참조하면, 서지정보의 완전성과 정확성을 기하기 위해, 특히 어떤 사항을 검토해야 하는지 알 수 있다. 개선을 위해서는 “반드시 출판된 도서의 실물을 손에 갖고” 편목하도록 정하고 있으며, 우선 고정 필드 (Fixed field)에서 인코딩 수준 (Encoding level), 레코드 유형, 발행연도 1과 발행연도 2, 발행국가 및 기타 실물에 근거한 구체적인 특성 (예를 들어, 내용의 특성, 삽도, 색인 등) 정보를 수정하도록 하고 있다. 또한 가변 필드에서는 LCCN은 어떤 상황에서도 변동되지 않는다는 전제 하에, 020 (ISBN), 050 (LC 청구기호: 정확한 형태와 연도에 주목), 1XX (기본기입), 245 (표제 및 저작자 사항), 250 (판사항), 260

41) Marilyn McCroskey and Michelle R. Turvey, “Why CIP Verification Is Important,” *Knowledge Quest*, vol. 32, no.5(2004), pp.36-37.

42) 실제로 본 연구자가 1980년대 말 UCLA의 과학기술도서관(Physical Sciences and Technology Libraries)의 편목부서에서 일하던 당시에 도서관에 입수된 과학기술 분야 단행본의 실물을 갖고, OCLC 데이터베이스에 있는 CIP 서지 레코드를 반입하여 완전수준 편목을 위하여 기술정보를 보완하는 업무를 수행하였던 경험이 있다.

43) OCLC, *Upgrading of Cataloging-in-Publication(CIP) Records*, 2003, <<http://www.oclc.org/worldcat/about/quality/cip/default.htm>> [인용 2005.3.11]

44) OCLC, *OCLC's CIP Record Upgrade Project*, 2001, <<http://www.oclc.org/worldcat/about/quality/cip/CIPspecs01.pdf>> [인용 2005.3.11]

(발행사항), 300 (형태사항), 4XX과 8XX (시리즈), 504 (서지참조 주기) 필드 등을 검토하도록 정하였다. 실제로 이 필드들의 내용이 추가되거나 변화된 것은 위에서 든 <그림 3>의 예로도 알 수 있다.

그 밖에 기존의 추가적 5XX 필드들의 정확성을 검토하고, 부출 기입을 정당화시키거나, 리프린트에 원래 출판정보를 추가하기 위한 목적으로만 추가적 주기 필드를 만들도록 하였다. 또한 주제, 이름 및 시리즈 부출 기입이 저작에 적절한지 확인하고, 필요하면 해당 전거 파일에 대해 확인하고 수정하는 작업까지 수행하고 있다. 필요하면 7XX 필드에 두 번째, 세 번째 저자를 위한 부출 기입을 만들어 주거나, 아틀라스 자료를 위해 255 필드 (Scale)와 034 필드 (Coded scale)를 확인하고, 동반 자료가 있는 도서에 적용되면 538 (시스템 요구사항) 등도 기입해 주도록 하고 있다⁴⁵⁾.

CIP 데이터를 개선하기 위하여 이 같이 상세한 명세를 제시하는 이유는 첫째, CIP 데이터의 최초 생성기관이 고급의 목록 전문사서들로 구성된 LC라고 할지라도 작성 시 서지정보 오류의 가능성이 있고, 불완전한 출판 전 서지정보의 속성 상 모든 필드를 검토하고 수정 및 보완할 필요성이 있기 때문이다. 예를 들어, 표제에서 library가 literacy로 입력되었다든지, CIP로 제출된 표제에 있던 단어들 실제 출판 시에는 빠져 있다든지 하는 오류나 변동이 발견되기도 하는데, 이러한 오류들은 종종 LC에서 발생한 것이고, 거기서 제공한 MARC 레코드 테이프에 근거한 서지 제품들에는 계속 남아있게 된다는 문제점이 있기 때문이다.⁴⁶⁾

둘째, OCLC와 같은 서지 유틸리티는 수많은 도서관들이 협력하는 가운데, 어떤 도서관이라도 이와 같은 명세를 참조하여 일정 수준 이상의 레코드를 만들도록 독려할 필요가 있기 때문이다. 사실상 OCLC 편목 서비스 이용자 가운데 완전 이상 인증 (Full or higher authorization)을 갖는 누구든지 300 필드를 추가할 수 있고, Enhance 프로그램의 정규적 참여자는 인코딩 수준을 제외한 전체 레코드를 편집할 수 있으므로, 모든 개별 도서관이 공통적으로 적용시켜야 하는 기준이 필요한 것이다. 참고로, 인코딩 수준은 국가적 수준의 Enhance 참여자들만이 수정할 수 있는데, 이 인증을 갖는 OCLC 회원 도서관들은 어떤 마스터 서지 레코드이든 완전수준 레코드로 만들도록 변경할 수 있는 권한을 갖는다⁴⁷⁾. 이 같이 CIP 레코드 (Encoding level: 8 'Prepublication level')를 완전수준 (Encoding level: 공백)으로 개선 할 권한을 일부 도서관에게만 허용하는 것은 종합목록 서지 레코드의 품질을 제어하기 위한 노력의 일부이다. 또한, WorldCat에 수록된 다른 서지 레코드를 수정할 때에도 도서관의 인증범위 규정에 따른 동일한 기준의 품질 제어가 적용되기 때문에, 서지 유틸리티가 출판된 도서, 혹은 출판예정인 도서에 대한 정확하고 완전한 정보를 확인하는 소스로 사용될 수 있는 것이다. 결국 서지 유틸리티의 종합목록 데이터베이스는 아직은

45) OCLC, 2001, *Ibid*.

46) Norman Desmarais, "Validator Subjects and Name Authority," *Electronic Resources Review*, vol.3, no.2 (1999), pp.26-27.

47) OCLC, 2003, *Ibid*.

불완전하지만 잠재적 유용성을 갖는 CIP 데이터의 정보원으로서, 또한 CIP 데이터를 기반으로 하였으나 정보의 정확성과 완전성이 보완되고 개선된 서지 레코드의 정보원의 역할을 함으로써 CIP 프로그램이 목표로 한 양질의 서지정보의 신속한 확산에 기여한다고 할 수 있다.

IV. 결론과 제언

지금까지 LC가 국가의 대표적 서지기관으로서 시행하는 CIP 프로그램의 전반적 특성을 간략히 살펴보고, CIP 데이터에 근거한 서지정보의 생성, 공유 및 활용 체제를 주관하는 현황을 살펴보았다. 또한 서지 유틸리티가 CIP 데이터에 대한 접근을 제공함으로써 신간 및 출판예정 도서의 서지 정보원 역할을 할 뿐 아니라, CIP 데이터에 기반한 서지 레코드의 완전성과 정확성을 개선하기 위하여 품질을 제어하는 역할을 살펴보았다. 특히 이 같이 CIP 데이터의 생성, 배포, 개선 및 활용까지의 과정에서 OCLC와 LC가 협력하고 공조함으로써 미국 국내 뿐 아니라 국제적 서지공유 및 협력체제가 유지, 발전된다는 사실을 볼 수 있었다.

도서관들이 서지 유틸리티의 목록 데이터베이스를 도서선정의 정보원으로 사용할 때, CIP 데이터를 이용함으로써 출판예정 도서의 정보를 얻을 수 있고, 또한 출판된 도서를 입수했을 때 개선된 CIP 레코드를 이용함으로써 정확하고 완전한 서지정보에 접근할 수 있다는 것은 비용과 시간을 절감하고, 궁극적으로는 이용자에게 더 나은 서비스를 제공할 수 있다는 것을 의미한다. 따라서 국내에서도 신뢰할 수 있는 포괄적이며 광범한 서지 유틸리티나 국가 대표도서관의 목록 데이터베이스가 있다면, 이를 통해 CIP 데이터를 보다 널리 활용할 수 있는 방안을 모색할 수 있을 터이지만, 아직은 그 같은 여건이 조성되어 있지 않다고 할 수 있다.

우선 OCLC나 RLIN과 같이 미국을 중심으로 발전해 온 서지 유틸리티라는 도서관들의 협력 네트워크 혹은 서비스는 국내에서는 아직 유사한 형태의 조직을 찾을 수 없다. 물론 국내에도 대학도서관 종합목록과 상호대차를 중심으로 한 전국대학학술정보 공유사업을 주관하는 KERIS와 같은 기관이 있고, 국립대학도서관협의회나 한국사립대학교도서관협의회와 같은 여러 가지 형태의 협력 네트워크가 존재하고 있으나, 다른 나라의 서지 유틸리티가 유지하는 수준과 범위에 준하는 종합목록 데이터베이스를 갖지 못하고 있다.

또한 국립중앙도서관이 유지하는 KOLIS.NET 역시 국가의 대표도서관이 모든 유형의 도서관과 정보센터를 대상으로 제공해야 할 다양한 정보와 서비스를 제공하지 못하는 한편, OCLC의 WorldCat이나 LC의 목록 및 전거 데이터베이스가 유지하는 수준으로 품질 제어된 서지와 전거 데이터베이스를 제공하지 못하고 있는 형편이다. 현재 국내의 도서관들이 국내자료나 동양자료의 편목을 위해서 비교적 많이 참조하고 있는 국립중앙도서관의 목록이 그 완전성이나 정확성 면에서

개선되어야 할 점이 많고, 특히 국내도서관의 CIP 데이터를 아직은 별로 활용할 수 없다고 하는 발 견⁴⁸⁾은 바꾸어 말하면 국립중앙도서관이 적어도 국내자료에 관한 한 가장 “정확하고 완전한” 서 지정보원의 역할을 할 수 있도록 보다 적극적 노력을 기울여야 한다는 기대를 입증하는 것이다.

국립중앙도서관에서 CIP 프로그램을 활성화시킨다고 해도, 전체적으로 목록 데이터베이스의 서 지 레코드를 고품질화 하고, 수록 자료의 범위를 확대시키지 않는다면, 서지정보원으로서의 역할에 많은 제한점이 있을 것임에 틀림없다. 미국의 CIP 데이터가 출판사, 서적 벤더, 및 도서관들에게 매우 중요한 정보원의 역할을 한다고 인정되는 반면에⁴⁹⁾, 현재 우리나라에서 시행하고 있는 CIP 데이터의 유용성은 아직은 확실하게 입증된 바가 없다. 오히려 국내도서관의 목록 전문사서들은 온라인 레코드를 생성하기도 쉽고, 서지 레코드를 얻기도 쉬운 신간 단행본의 CIP 데이터보다는 학위논문, 구간 도서, 비도서자료, 전자자료 등과 같이 타기관의 서지 레코드 입수가 어렵고, 자체 적 온라인 레코드 생성이 어려운 자료를 위해 포괄적이고 품질이 높은 서지 레코드를 제공할 수 있는 제도적 장치가 마련되어야 할 것을 요구하고 있다⁵⁰⁾.

도서관간 서지정보와 자원의 공유 및 협력이라는 기본적 목표를 갖고 있는 서지 유틸리티의 모 형이 지금까지 국내에서 발전되지 못했던 만큼, 유일한 서지정보원이 될 수 있는 국립중앙도서관의 목록 데이터베이스의 수준과 범위에 대한 기대는 어찌면 당연한 것이라고도 할 수 있다. 국가서지 센터로서 국립중앙도서관이 광범하게, 망라적으로 수집한 모든 유형의 구간 및 신간 국내 정보자료 를 대상으로 우수한 품질의 목록과 전거 레코드를 제공함으로써, 국내 도서관들로 하여금 시간, 노 력 및 비용을 절감할 수 있게 한다면, 최근 캐나다의 국가대표도서관이 제공하는 양질의 서지 레코 드를 활용하면서 그 나라의 도서관들이 최근에 달성한 시간과 비용의 절감, 그리고 그에 따른 이용 자 서비스의 확대 사례⁵¹⁾에서 목격하는 바와 같이, 도서관계 전반에 보다 높은 생산성과 경쟁력을 가져오게 될 것이 기대된다.

〈참고문헌은 각주로 대신함〉

48) 김포옥, 노옥림, 전계논문.

49) McCroskey and Turvey, 2004, *Ibid*.

50) 윤정옥, 전계논문.

51) Jamshid Beheshti, Andrew Large, and Pat Riva, “Cost savings to Canadian university and large urban public libraries from their use of National Library of Canada MARC records,” *Library Resources & Technical Services*, vol.47, no.2(2002), pp.44-57.

к с і