

국가지식정보자원 분류 체계 표준화 연구*

A Study on the Developing Standard Classification of the National Knowledge and Information Resources

고 영 만(Young-Man Ko)**

서 태 설(Tae-Sul Seo)***

조 순 영(Sunyeong Cho)****

목 차

- | | |
|-------------------------|-----------------------|
| 1. 서 론 | 4. 국가지식정보 분류 표준(안) 모형 |
| 2. 문헌 분류와 지식정보자원 분류 | 5. 결 론 |
| 3. 국가지식정보 표준 분류체계 개발 방안 | |

초 록

본 연구의 목적은 지식정보자원 데이터베이스를 구축하고 관리하고 서비스하는 국내의 모든 공공 기관과 민간 기관에서 공동으로 활용될 수 있으며 다양한 자료유형과 매체유형을 포괄할 수 있는 국가지식정보 분류 표준(안)을 개발하는 것이다. 연구의 기초 작업으로서 문헌분류 체계와 지식정보자원 분류 체계의 차이, 지식정보자원의 분류사례 및 지식정보자원 분류에 대한 문헌분류 체계의 적용 가능성을 분석하였다. 기초 분석에 이어 국가지식정보 표준분류 체계 개발에 필요한 지식의 분야별 분포, 분류의 계층과 유연성 확보 방안, 지식정보자원의 자료유형과 매체유형, 국가지식정보의 포털서비스 방안 등이 분석되었다. 이러한 분석을 토대로 본 연구에서는 주제 분류표, 자료유형 분류표, 매체유형 분류표로 구성된 다원적 국가지식정보 분류 표준(안) 모형을 제시하였다.

ABSTRACT

The purpose of this study is to make out a draft for the standard classification of the National Knowledge and Information Resources. As the result of the Study the standard classification system of the national knowledge and information resources, named "Knowledge Classification : KC" is suggested. KC consists of 3 classification systems : classification by subject, type of resources and type of media. The classification by subject has 12 main classes, and each main class has divisions. Main classes consist each of major discipline or group of related disciplines. The type of resources is classified by 10 types of content, likewise numbered 0-9, and the media of knowledge are classified by 8 types, likewise 0-7. In the practice the notation always consists of 2 characters and 2 digits. The first character designate main class and the second character designate division. The first number designate the type of resources and the second number designate the type of media.

키워드: 표준 분류, 지식정보자원, 주제 분류, 자료유형, 매체유형

Standard Classification, National Knowledge and Information Resources,
Subject Classification, Type of Resources, Type of Media

* 이 연구는 한국정보문화진흥원의 '05년도 지식정보자원관리사업 표준화연구 사업으로 이루어졌음.

** 성균관대학교 문과대학 문헌정보학과 교수(ymko@skku.edu)

*** 한국과학기술정보연구원 NDSL사업단 책임연구원(taeso@kisti.re.kr)

**** 한국교육학술정보원 국가지식정보센터 수석연구위원(chosy@keris.or.kr)

논문접수일자 2006년 8월 14일

게재확정일자 2006년 9월 12일

1. 서론

1.1 연구 목적

2005년부터 국가지식정보자원 DB의 공동이용과 활용 촉진을 목적으로 하는 “지식정보자원관리사업”의 2단계 사업(2005-2009)이 추진되고 있다. 국가지식정보자원을 효율적으로 활용하고 그 가치를 극대화하기 위해서는 범국가적 차원에서 일관적으로 관리할 수 있고, 중장기 전략을 수립할 수 있는 표준화된 지식정보자원의 분류체계가 필요하다. 또한 기관과 분야에 따라 분류 체계와 접근 방식이 상이하게 구성된 지식정보자원을 효율적으로 관리하기 위해서는 생산, 구축, 서비스의 제공 및 이용을 고려한 분류체계가 필요하다. 특히 상업적 민간 포털을 통해 국가지식정보자원 통합검색시스템을 점진적으로 연계하고 있는 상황을 고려할 경우 공공부문과 민간부문에서 공동으로 활용될 수 있으며 입수가능성(Availability), 접근성(Accessibility), 사용성(Usability)을 고려한 지식정보자원 분류체계의 개발과 표준화가 절실한 실정이다.

그러나 지금까지는 사업의 목적에 국한되거나 부분적으로만 유효한 분류체계들이 개발되어 사용되고 있으며 지식정보자원을 국가적 차원에서 관리할 수 있는 표준화된 분류체계가 존재하고 있지 않다. 이는 사업의 주관기관과 분야별 관리기관 간의 효율적인 지식정보자원 관리가 제대로 이루어지지 않는 결과로 이어지고 있으며, 이용자 측면에서의 접근성과 사용성에 있어서도 많은 불편을 초래하고 있다. 본 연구는 지식정보자원 데이터베이스를 구축하고 관리하고

서비스하는 국내의 공공기관과 민간기관에서 공동으로 활용될 수 있으며 다양한 자료유형과 매체유형을 포괄할 수 있는 국가지식정보 분류 표준(안)을 개발하는 것을 목적으로 한다.

1.2 연구대상 및 범위

지식정보자원관리법(일부개정 2004.12.30 법률 7263호)에서는 지식정보자원의 사업 범위를 보존의 가치와 이용가치를 지닌 지식정보로 규정하고 있으며 그 대상영역을 과학기술, 교육학술, 문화예술, 한국학, 정보통신 및 산업 등의 일정한 분야로 설정하고 있다(정보통신부 2000). 그러나 조동성(1993)은 공적자원이 투입되어 민간부문에서 수집, 보유, 관리하고 있는 지식정보자원이라는 광의의 해석을 할 경우 국가지식정보자원의 범위는 국민이 실질적으로 필요로 하는 지식정보자원, 인간의 생존과 유지에 관련된 지식정보자원, 사회적 안전을 유지하기 위한 지식정보자원 등을 포함하는 것이 적절한 것으로 주장한 바 있다. 본 연구에서 개발하고자 하는 국가지식정보자원 분류 표준화의 범위는 국가지식정보자원에 대한 광의의 해석인 “비공개로 분류된 지식정보뿐 아니라 불특정 일반 국민에게 이용될 수 있는 지식정보로서 공공기관 및 민간부문이 일반예산은 물론 기타 기금 등 공적 자원을 투입해 수집·보유·관리하고 있는 모든 지식정보”를 그 대상으로 한다.

1.3 연구방법

본 연구에서는 문헌분류 체계의 분석과 지식

정보분류 사례의 분석을 토대로 국가지식정보 표준분류 체계를 개발하고자 하며, 이를 위한 연구 방법으로 문헌분류 체계 분석, 지식정보 자원 분류 사례 분석 및 표준분류안 개발 연구의 세 가지 방식을 적용하였다.

문헌분류 체계 분석은 듀이십진분류법(Dewey Decimal Classification : DDC), 한국십진분류법(Korean Decimal Classification : KDC), 미 국회의도서관분류(Library of Congress Classification : LCC), 콜론분류(Colon Classification : CC) 등의 문헌분류 체계가 국내외 전자정보원의 분류와 국내의 국가지식포털 분류에 어떻게 반영되고 있는지를 찾아내기 위한 것이다.

지식정보자원 분류 사례 분석은 지식정보를 관리하고 서비스를 제공하는 기관에서 사용하고 있는 분류체계의 특징과 문제점을 찾아내기 위한 것이다. 이를 위해 종합정보센터 8개 기관의 분류체계 및 지식정보 관리기관 분류체계, 포털서비스 시스템 지식정보 분류체계, 공공기관 행정 데이터베이스 분류체계를 분석하였다.

표준 분류안 개발은 문헌분류 체계 및 지식정보 분류 사례의 분석을 토대로 국가지식정보 분류 표준화에 고려할 사항을 도출하기 위한 것이다. 지식의 분야별 분포, 분류체계의 계층과 유연성, 자료유형과 매체유형, 포털서비스 측면에서 고려할 사항들이 분석되었다.

2. 문헌분류와 지식정보자원 분류

분류란 분류의 대상이 되는 것들을 분류의 목적과 기능에 근거하여 유사한 것들은 함께 묶어주고 같지 않은 것을 구분해 주는 것이다.

분류에서는 기준에 따라 형성되는 분류 대상 간의 종속관계에서 상위 개념인 유개념과 하위 개념인 종개념 간의 관계를 체계적으로 조직하는 과정이 이루어진다. 일반적으로 분류는 크게 자연적 분류와 인위적 분류의 둘로 나뉜다. 자연적 분류는 분류의 근거가 자연현상의 객관적 성질인 자연 자체이므로 이미 어느 정도 분류가 이루어져 있다. 인위적 분류에서는 목적과 편의에 따라 일정한 원칙을 설정하여 분류의 대상을 실용적으로 제어하고 질서를 부여한다. 학문분류와 문헌분류는 인위적 분류에 해당하며, 지식정보자원의 분류 역시 인위적 분류에 해당한다.

2.1 학문분류와 문헌분류

일반적으로 지식관련 분류는 학문분류, 문헌분류, 백과사전 분류, 정보서비스기관 분류, 연구활동 분류 및 인터넷 포털서비스 업체의 분류 등으로 구분할 수 있다. 이 중에서 문헌분류는 논리학상의 분류개념과 일치하고 학문 활동에서 이루어지는 학문적 분류이론과 축을 같이 하며 대부분의 현대적 문헌분류표가 학문분류의 바탕 위에서 창안되고 발전되었다는 점에서 넓은 의미의 학문분류에 속한다고 할 수 있다. 학문분류는 인간의 지식을 체계화한 학문을 체계적으로 조직화하려는 시도 혹은 학문의 연구과정에서 얻어진 개념을 체계화하려는 시도라고 할 수 있다. 현대 학문의 뿌리를 형성시킨 아리스토텔레스의 분류법에서 비롯되었으며, 이후 수많은 학자들에 의해 인간의 사유를 조직화하려는 시도가 이어져왔다. 학문의 분류는 문화적 특성과 학파에 따라 다양한 관점에서

분류되었으며, 오늘날에는 학문의 목적, 학문의 대상과 존재양식 외에도 학문의 연구방법, 실용성 등에 따라 그 분류가 이루어지고 있다.

학문분류는 근세까지 철학적 의미와 주요 학문명칭을 부여하는 것에 중점을 두고 발전되었으나 현대에 와서는 각 나라에서 다양한 학문분야의 연구 개발 지원, 측정 및 분석에 실용적으로 활용하기 위해 연구 분야를 바탕으로 고안한 것이 주를 이루고 있다. 따라서 학문분류는 연구분야 분류표에도 많은 영향을 미치고 있으며, 시대와 지역, 분야 및 지식활동 패턴의 변화에 따라 체계화되는 방법과 분류표가 매우 다양해지고 있다. 국내에서도 여러 기관에서 특정 분야별 또는 전 학문분야에 대한 분류표가 다양하게 작성되어 사용되고 있다. 그러나 이는 지식정보의 효율적이고 전체적인 관리 및 범용적인 활용을 어렵게 만드는 결과를 낳고 있다.

문헌분류란 주로 도서관에서 소장하고 있는 자료를 유사성에 따라 논리적으로 체계화하는 것을 말한다. 해당 문헌이 담고 있는 정보의 내용을 토대로 이루어지므로 필연적으로 학문 활동에서 이루어지는 학문분류의 이론과 축을 같이 할 수밖에 없다. 실제로 있어서도 대부분의 현대적 문헌분류표는 학문분류의 바탕 위에서 창안되고 발전되었다. 그러나 문헌분류는 도서관에 소장된 문헌이나 자료의 실제적인 분류라는 실용적인 기능과 목적을 갖는다는 점에서 학문분류와 차이점을 보이고 있다. 학문분류는 학문의 연구과정에서 동류(同類)를 찾아 명확성을 분석하고 체계화하는 것으로서 학문의 논리적 성질이나 성격이 분류기준이 되고 분류 자체가 목적이 되지만, 문헌분류는 학문의 연구

결과에서 얻어진 기록물을 중심으로 내용의 주제와 형식에 따라 이를 체계화하여 이용의 편리성을 추구하기 위한 것이므로 문헌자체가 담고 있는 내용의 주제나 문헌 형식 등이 분류의 기준이 된다.

학문분류는 학문 자체의 논리성을 분류기준으로 삼고 있기 때문에 추상성이 강한 것에 비해 문헌분류는 분류대상이 학문의 결과물이기 때문에 구체적이고 실용성을 갖추어야 한다. 따라서 특정 원리나 규칙에 의해 만들어진 다양한 문헌분류표가 많이 나와 있으며 문헌을 분류할 경우 반드시 특정한 문헌분류표를 사용해야 한다. 문헌분류표는 일반적으로 도서관에서 일차적으로 내용에 따라 주제접근을 한 후 이차적으로 형식구분을 하고 있으며, 출처나 행정 기능, 연대나 지리적 위치에 따라서도 구분을 하고 있다. 오늘날에는 매우 다양한 문헌분류표들이 개발되어 실제에서 사용되고 있으며, 대표적인 것으로 듀이십진분류법(DDC), 한국십진분류법(KDC), 국제십진분류법(UDC), 미의회도서관분류(LCC), 콜론분류(CC) 등을 들 수 있다.

2.2 문헌분류와 전자정보원 분류

연구 분야 분류 또는 지식정보의 분류체계를 작성하는 연구자들은 문헌분류가 장서의 분류와 보관에는 적합할 수 있으나 시대적 변화를 적시에 반영하고 있지 못해 지식활동의 파악과 지식활동 정보의 데이터베이스화에는 거의 유용하지 않다는 비판을 하기도 한다(설성수, 송충한 2000). 이는 문헌분류가 여러 가지 분류원칙, 기준 및 기호법 등의 제약으로 인해 융통성과 확장성이 적은데다가 과거로부터 현재에 이

르는 모든 문헌을 되도록 일관성 있게 분류하고 배가된 장서를 관리하고 이용시키기 위해서 분류표의 수정 폭을 가능한 최소화할 수밖에 없었기 때문이라고 할 수 있다. 또한 문헌분류의 표목으로 채택되는 시점이 연구의 결과물인 문헌의 빈도가 충분히 많아져야만 하므로 시간적인 격차가 발생하는 것도 사실이다.

하지만 문헌분류 역시 현대적 시각에서 주제의 분화와 발전, 통합 그리고 분류표의 구조개선 등과 같은 요구를 적극적으로 수용하고 있다. 문헌분류는 전 학문분야의 주제를 상세하고 망라적으로 다루고 있는 분류표이기 때문에 특정 분야의 분류 체계의 개발에 있어서도 활용가능성이 높다. 또한 어떤 형태의 분류체계라 하더라도 설계원칙과 이론을 바탕으로 이루어져야 하므로 문헌분류에서 오랫동안 개발하고 적용시켜온 분류이론 및 설계원칙을 참고하는 것이 반드시 필요하다.

1990년대 이후 인터넷 정보 및 전자정보원의 양이 급속히 증가하면서 온라인 시스템, 검색엔진 및 인터넷 자원 디렉토리에서 전자정보원의 조직 및 검색에 계층적으로 구조화되어 있는 문헌분류표를 활용하는 예가 늘어나고 있다. DDC, LCC, UDC의 분류표들이 활용되고 있으며 그 중 LCC를 활용한 예는 “CyberStacks (<http://www.public.iastate.edu/~CYBERSTACKS/>)”와 “Scout Report Signpost(<http://signpost.org/index.html>)”를 들 수 있다.

“CyberStacks”에서는 아이오와 주립대학교 도서관의 시스템 중에서 과학, 기술 및 관련 분야의 선별된 인터넷 자원을 조직화하는데 LC 간략분류표를 이용하고 있으며 이 분류표는 G (Geography, Anthropology and Recreation),

H(Social Sciences), J(Political Science), K(Law), Q(Science), R(Medicine), S(Agriculture), T(Technology), U(Military Science), V(Naval Science)의 열개 주류로 이루어져 있다. 구조적 계층은 LCC의 기호법에 따라 주류 다음의 하위류와 세목의 3단계로 점차 세분화되어 표현되고 브라우징 검색을 할 수 있도록 되어 있다. “CyberStacks”를 고안한 맥커맨(Gerry McKiernan)은 LC 분류표를 선택한 이유로서 도서관들이 여러 세대를 거쳐 다양한 출판물과 매체를 조직하는데 사용되어온 잘 구축된 분류표이며 주제의 범위와 내용뿐 아니라 정보의 포맷과 개념적 관계를 표시해주기 때문에 인터넷 자원을 조직하는데도 적절한 내용과 구조를 제공해줄 수 있다는 점을 들고 있다.

“Scout Report Signpost”는 미국의 고등교육 커뮤니티에 양질의 전자정보원을 제공하는 데 있어서 LC 분류표를 수정하여 22개의 대분류와 각각의 하위분류 2단계로 정보를 범주화해서 사용하고 있다. “Internet Scout Project”에서의 LC 분류표 수정 내용의 특징은 크게 세 가지를 들 수 있다. 하나는 숫자로 된 기호의 확장 없이 문자로 된 분류코드만을 사용하는 점이며 다른 하나는 접근을 용이하게 하기 위하여 계층구조에서 주제가 한 가지 이상의 분야에 속하는 경우 많은 항목들을 두 개의 다른 분류코드에 배정한 점이다. 세 번째 특징은 이용자에게 도움이 되도록 선택된 LC 분류표를 설명해주는 주기들이 부가되었다는 점을 들 수 있다.

DDC를 활용하는 경우는 고려대학교(<http://library.korea.ac.kr>)에서 전자정보원의 분류와 단계별 브라우징 검색을 위해 활용하는 예를 들

수 있다. 이 전자정보원 검색 시스템에서는 모든 학문분야가 골고루 분포되어 있지 않은 전자저널 매체의 특성을 고려하여 해당 대학에서 구독하는 모든 전자저널을 DDC로 분류하여 주제의 분포도를 조사한 다음, 조사 결과에 따라 10개로 이루어진 DDC의 주류에 의학을 대분류로 이동시켜 11개의 주류를 작성하였다. 또한 브라우징 검색을 위한 수직적 단계를 류, 강, 목의 3단계로 제한하여 설정하였으며, 세목(소수점 이하의 분류기호)으로 분류되는 저널의 빈도가 많은 항목은 요목이상으로 이동하여 배정하였다. 해당 저널이 없는 분류기호는 검색결과가 없게 되는 경우를 방지하기 위해 분류표에서 제외시켰으며, 전자저널은 서가의 배열을 염두에 둘 필요가 없는 것이므로 학제적 성격의 저널은 중복으로 분류하여 여러 분야에서 검색이 될 수

있도록 하였다. 특히 항목 이동에 따른 DDC의 매칭표를 제시하고 있으며(표 1 참조), 향후 사용되지 않는 분류기호가 나타나는 경우 이를 수용할 수 있도록 설계되었다. 이는 분류기호를 일관성 있게 유지하면서 이를 활용한 주제별 브라우징 검색을 위한 별도의 테이블 작성이 가능하도록 하기 위한 것이라 할 수 있다. DDC를 활용한 장점으로는 첫째, 별도의 복잡한 분류표를 작성하지 않아도 된다는 점 둘째, DDC로 분류되는 전자정보원뿐 아니라 도서관 온라인 목록을 모두 포함하여 분류기호에 의해 같은 주제의 자료를 통합검색 할 수 있다는 점 셋째, 출판사로부터 제공되는 DDC 기호를 그대로 수용함으로써 저널의 메타데이터를 구축하기 위해 분류작업에 소요되는 시간과 비용을 줄일 수 있다는 점 등을 들 수 있다.

〈표 1〉 DDC를 활용한 고려대학교 도서관 전자저널 분류표 사례

대분류	중분류	소분류	DDC
Social Science	Management	General	650
		Office services	651
		Processes of written communication	652
		Shorthand	653
		Accounting	657
		General management	658
		Personnel management	658.3
		Executive management	658.4
		Management of production	658.5
		Management of materials	658.7
		Management of distribution	658.8
		Advertising and public relations	659
Technology	Computer Science	Computer science	004
		Inferfacing and communication	004.6
		Systems	003
		Computer programming, programs	005
		Special computer methods	006
		Artificial intelligence	006.3
		Computer pattern recognition	006.4
		Computer graphics	006.6

2.3 문헌분류와 국가지식포털 분류

국가지식포털에서 제공하는 디렉토리 검색은 주제별 검색과 연계기관별 검색의 두 가지 유형이 있다. 기관별 디렉토리 검색은 문화, 역사, 산업, 정보통신, 해양수산, 과학, 교육학술, 건설기술의 8개 대분류로 구분하고 있으며 하위 주제구분은 각 분야별로 지정된 정보센터에서 제공하는 분류체계를 따르고 있다. 대부분 공공기관에 의해 정보가 제공되고 있기 때문에 각 기관의 특성을 반영할 수 있는 장점은 있으나 분류표 전체의 통일성과 일관성이 결여되어 있을 뿐아니라 항목간의 수평적, 수직적 균형이 이루어지지 않았다. 주제별 디렉토리 검색의 분류체계는 인문, 철학, 종교, 사회, 과학, 기술, 예술, 언어문학, 정보통신, 역사의 10개 항목으로 대분류가 구성되어 있으며, 계층은 소분류까지 3단계로 이루어져 있다. 중분류까지의 분류체계는 <표 2>와 같다.

국가지식포털의 주제 분류는 DDC나 KDC

와 같은 문헌분류표를 상당부분 참조하고 있다. 어학과 문학을 언어문학으로 통합시킨 점과 신설 항목인 인문학과 정보통신을 제외한 모든 항목과 배열순서가 매우 유사하며, 문헌분류와 같이 특정 분야에 한정되지 않는 지식 전 분야를 다루고 있다는 점이 이를 뒷받침한다. 주요한 특징으로는 한국학, 경학, 천도교, 서예, 사부, 자부, 집부, 한국사 등과 같이 한국적 상황을 반영한 지식정보 분야가 많다는 점과 문헌분류에서는 총류의 컴퓨터학, 공학의 컴퓨터공학, 통신 등 여러 분야에 걸쳐 있는 정보통신이 국가지식포털에서는 대분류로 배정되어 있어 특정 기관의 지식정보에 비중을 두고 있다는 점을 들 수 있다.

국가지식포털 주제 분류의 문제점은 분류표 전체의 통일성과 일관성 결여 및 항목간 수평적, 수직적 균형 결여를 들 수 있다. 크게 언어, 문학, 역사, 철학으로 구성되는 인문학의 경우 다른 항목과의 수평적 균형이 맞지 않고 있으며, 인문학의 중분류로 설정되어 있는 항목들

<표 2> 국가지식포털 주제별 디렉토리

대분류	중분류
인문	도서학, 문헌정보학, 고고학, 인류학, 한국학, 문화학, 기호학, 심리학, 윤리학
철학	동양철학, 서양철학, 논리학, 경학(經學), 형이상학, 철학체계
종교	종교학, 불교, 기독교, 도교, 천도교, 신도, 인도교, 회교, 비교종교학
사회	사회학, 정치외교학, 법학, 통계학, 경제학, 경영학, 행정학, 교육학, 민속학, 군사학, 무역학, 관광학, 회계학, 지리학
과학	수학, 물리학, 화학, 천문학, 지학, 광물학, 생물과학, 식물학, 동물학
기술	토목공학, 건설교통, 기계공학, 전기전자, 화학공학, 의학, 제조업, 가정학, 해양수산, 환경공학, 생명공학, 자원에너지, 농업공학
예술	건축예술, 음악, 연극, 조각, 공예, 서예, 회화, 사진술, 운동
언어문학	한국문학, 중국문학, 일본문학, 영문학, 독일문학, 프랑스문학, 세계문학, 세계언어, 사부(史部), 자부(子部), 집부(集部)
정보통신	인터넷, 하드웨어, 소프트웨어, 통신네트워크, 컴퓨터과학, 전자상거래, CALS, 전자문서
역사	역사학, 한국사, 중국사, 동양사, 서양사, 문화제, 박물관, 주제별역사, 역사관

은 대부분 사회과학, 철학, 역사 등 이미 설정되어 있는 항목에 배정이 가능하다. 또한 문헌분류표에서는 학제적 또는 복합주제를 가진 저작들이 주기에 의해 특정기호에 분류하도록 되어 있고 이밖에도 총류와 같은 일반적인 지식을 분류할 수 있는 항목이 없다. 중분류 및 소분류 항목의 전개가 인위적으로 이루어진 점 역시 문제점 중의 하나이다.

2.4 지식정보자원 분류 사례

지식정보를 관리하고 서비스를 제공하는 기관들의 분류 사례를 분석하기 위해 지식정보자원관리사업에 의해 2005년 8월 현재 지정되어 있는 종합정보센터 8개 기관(정보통신부 2001), 국내 지식정보 관리기관의 분류체계, 해외 지식

정보 서비스 기관 분류체계, 포털서비스 시스템 지식정보 분류체계, 공공기관 행정 데이터베이스 분류체계를 분석하였다(표 3 참조).

분석 결과 종합정보센터의 분류는 기본적으로 보유 데이터베이스를 중심으로 분류가 이루어져 있으나 분류 체계의 통일성과 논리성이 결여되어 있는 것으로 나타났다(서태설 2003). 문화포털과 역사 분야의 분류항목들은 시대, 지역, 특정 연구분야의 특수성을 포괄하고 있으며, 정보통신 분야의 경우 동향정보를 제공하기 위하여 자원의 유형과 유통 형식을 기준으로 분류하여 나열하고 있는 특징을 지니고 있다.

국내의 지식정보 관리기관에서 자체적으로 구성하여 사용하고 있는 지식 분류체계는 매우 다양하다. 분석 대상 기관들의 분류체계 분석 결과 데이터베이스 주제 분류는 인터넷이 활성화

〈표 3〉 분석 대상 지식정보자원 분류 사례

(1) 지식정보자원관리사업 종합정보센터 분류체계	과학기술종합정보시스템(science.knowledge.go.kr)
	학술연구정보서비스 RISS4U(www.riss4u.net)
	문화포털(www.culture.go.kr)
	한국역사정보통합시스템(kh2.koreanhistory.or.kr)
	정보통신통합검색시스템 ITFIND(www.itfind.or.kr)
	정부출연 연구기관 지식정보 검색시스템 IKIS(www.ikis.re.kr)
	건설교통전자정보관(www.codil.or.kr)
	해양수산학술연구정보검색서비스(www.ocean.go.kr)
(2) 국내 지식정보 관리기관 분류체계	한국데이터베이스진흥센터(1997) : 데이터베이스 분류
	한국학술진흥재단(1999) : 연구분야(학술활동) 분류
	한국문화콘텐츠진흥원 : 문화콘텐츠 분류
	저작권심의조정위원회 : 저작물 분류
	한국소프트웨어진흥원(2004) : 디지털콘텐츠 분류
(3) 해외 지식정보 서비스 기관 분류체계	한국전산원(1998) : 전자상거래 표준 분류
	OCLC WorldCat, OCLC FirstSearch, BL, CDL, USGPO
(4) 포털시스템 지식정보 분류체계	KTH 컨소시엄(엠펙스, 야후, 파란)
	네이버 지식시장, 네이버 지식레퍼런스
(5) 공공기관 행정 데이터베이스 분류체계	전자정부, 국가기록원, 통계청 등 46개 기관

화되기 이전의 데이터베이스를 대상으로 하고 있어서 국가지식정보자원 분류에 많은 참고가 될 수 있는 것으로 나타났으며, 연구분야(학술 활동) 분류 역시 국가지식정보 표준 분류 체계의 기반 구조가 될 수 있는 대안 중의 하나로 평가되었다. 문화콘텐츠 분류는 문화예술 분야의 분류에 참고할 가치가 있는 것으로 나타났으며 디지털콘텐츠와 전자상거래 분류는 지나치게 전문적이어서 세부분야 분류에 한정적으로만 참고할 수 있는 것으로 평가되었다.

학문과 지식의 변화가 급속하고 분류의 목적이 각기 상이하기 때문에 해외의 경우에도 지식 정보 관리 또는 서비스 제공 기관에서 사용하는 분류체계가 표준으로 채택된 사례는 없다. 디지털 지식정보는 통일된 분류체계가 필요할 정도로 양이 많지 않고 모두 데이터베이스화되어 온라인 검색이 가능해서 자원의 분류체계에 대해 자유롭기 때문에 자원을 보유하고 있는 기관과 자원의 특성에 따라 각 기관에서 원자료의 보유 관, 원 자료의 물리적 특성, 편집 형태, 데이터 유형에 따라 임의로 분류하고 있는 것으로 나타났다. 디지털 지식정보의 분류체계는 온라인 검색의 장점을 반영하여 도서 등의 유형 자원에 비해 상위 개념 수준에서 소수의 구분체제를 이루고 있으나, 종합적인 지식정보를 다루는 기관에서는 공통적으로 기존 도서관에서 채택하는 LCC, DDC 등의 주제별 분류체계를 채택하고 있는 것으로 나타났으며, USGPO 와 같이 원자료가 생산되는 기관의 주제별 영역이 뚜렷한 경우에는 생산기관 중심으로 분류하고 있는 것으로 나타났다.

검색포털의 국가지식서비스는 일부 주제는 제공기관의 분류체계를 그대로 사용하고 있으

며, 네이버의 지식 레퍼런스 서비스의 경우 한국학술진흥재단의 학술분류체계를 기본 틀로 삼고 있다. 네이버의 경우 이용자들의 검색 요구(검색 쿼리를 통한 특성) 및 데이터 양을 반영하여 중 카테고리를 대 카테고리로 배치하는 형태를 취하고 있으며, 특히 중 카테고리 이하의 카테고리는 유동적으로 구성되어 있어서 주제별 전개와 동시에 유형별 전개를 하고 있으며 링크카테고리 및 두개 이상의 카테고리를 부여하고 있다.

공공기관의 경우 조사대상 40개 기관 중 전 범위의 주제 분류체계를 사용한 기관은 8 개 기관이었으며, 대부분 업무 중심, 이용자 중심의 분류로 행정조직구조를 반영하거나 민간 포털 서비스의 분류와 유사한 분류체계를 적용하고 있는 것으로 나타났다.

3. 국가지식정보 표준 분류 체계 개발 방안

3.1 지식의 분야별 분포와 국가지식정보 표준 분류

대한출판문화협회 주제별 발행부수, GPO Subject, Ulrich's Universe 저널, BL 저널, OCLC FirstSearch DB, 대학의 학과 분포의 분석 결과에 따르면 주제별 분류의 수준 차이로 일관된 순위를 보이지는 않으나 대체로 사회과학분야의 지식정보의 양이 가장 많은 것으로 나타나고 있다(고영만 2004). 이는 사회과학 분야가 너무 포괄적이며, 연구 결과물이 대부분 서적으로 출판되기 때문에 실제 출

판량이 상대적으로 과다한 것을 가장 큰 원인으로 들 수 있다. 한편 한국데이터베이스진흥센터(2004)에서 조사한 국내의 주제분야별 지식정보자원 분포에 따르면 대체적으로 과학기술 분야의 정보량이 가장 높았으며 교육학술, 문화, 역사 등의 순서대로 특정분야의 정보량이 많은 것으로 나타나고 있다.

지식정보의 양이 중요한 기준으로 작용할 경우 교양(문학, 어학)과 상식 분야의 분포가 높게 나타나게 되며 교양, 상식의 이용자는 다른 전문 분야 이용자와도 중복되는 문제점을 안고 있다. 따라서 물리적 자원 중심이 아니라 디지털 자원이 중심이 되는 국가지식정보의 분류체계에서는 지식정보자원의 물량 분포보다는 이용자 분포에 따른 분류가 보다 합리적인 대안이 될 수도 있다.

3.2 국가지식정보 표준분류의 계층과 유연성

지식정보자원은 문헌분류를 기반으로 하지만, 내용적 측면에서는 지식의 체계화인 학문과 밀접한 연관성을 지니고 있으므로 필연적으로 전통적인 학문분류를 고려하는 것이 바람직하다. 그리고 각 분야마다 정해진 항목 이외의 주제를 배정할 수 있는 기타 항목 및 새롭게 나타나는 지식 분야를 용이하게 추가할 수 있는 구조가 반드시 고려되어야 한다. 일반적으로 문헌분류의 유효한 계층은 4단계 정도이므로 문헌분류 만큼의 세목이 나열될 필요는 없는 국가지식정보 분류체계의 수직구조는 대분류, 중분류, 소분류 정도의 현행 유지가 바람직하다.

같은 계층에 분포되는 항목 간의 수평구조는 분류 차원에서 이공계와 인문사회계의 배치문

제가 있고, 이어서 중분류, 소분류 배치문제가 뒤따른다. 이들 항목은 각 분류표마다 차이가 있기 때문에 국가지식정보 분류에 가장 적합한 분류체계를 구축할 필요가 있으며, 국내외 분류체제, 대학 학과 등 지식분류체제의 근간이 되는 분류체제에 공통적으로 포함되어 있는 분야를 중심으로 하는 것이 바람직하다. DDC의 “류”를 중심으로 주제별 분류 체제를 채택하고, 이용측면을 고려하여 DDC의 형식 구분과 같이 데이터의 유형을 조합하여 사용하는 방식이 고려될 수 있을 것이다. 또한 계층 구조에서는 “강” 수준이지만 학문의 범위가 넓고, 이용자 그룹이 명확하여 사회적으로도 독립성이 인정되는 분야(예 : 의학, 법학, 신학, 경영학)는 별도로 구분할 필요가 있다. “목” 이하의 수준일 경우에도 정보통신 분야와 같이 사회적인 변화에 의해 지속적인 학문의 세분화가 이루어지고 자원의 생산량이 급속히 증가되는 분야는 “류” 수준으로 구분할 필요가 있다.

학제간 주제와 복합 주제를 분류표에 반영하는 것과 관련하여 지금까지의 열거형 분류표는 다면적 주제를 가진 정보를 표현하기가 어려운 문제점을 지니고 있다. 하지만 다면적 이원배치형식의 분류표는 이용자가 사용하는데 어려움을 느끼는 단점을 가지고 있다. 그러나 이용 편의성과 실용성이 지나치게 강조될 경우 범용적인 표준분류 체계가 될 수 없는 문제점이 있다. 따라서 열거형 분류표를 활용하되 주제부여를 복수로 하여 검색할 수 있도록 하는 것이 효율적인 방법이 될 수 있다.

분류표의 기호사용법에 있어서는 숫자나 문자로 된 코드를 사용할 것인지 또는 표목과 같은 용어를 그대로 사용할 것인지를 결정해야

하며, 정보의 관리 측면에서는 코드사용이 편리하다는 점도 고려되어야 한다.

3.3 국가지식정보의 자료유형과 매체유형

대부분의 대학도서관은 기존의 전통적인 도서관에서 다루고 있는 자료들과 새롭게 구축되는 디지털 콘텐츠의 통합검색과 관리를 위해 자료의 유형과 매체의 유형을 구분한다. 통계, 사전, 고서, 교과서와 같이 즉 이용자의 접근 동기가 일반 자원과는 달리 주제 보다 형식에 가까운 자원을 별도로 구분함으로써 주제 분류가 모호하고 출판 목적과 이용행태가 뚜렷이 구분되는 자원은 편집형식으로 별도 구분하기 위한 것이다. 매체유형의 구분은 정보기술의 발전에 따라 종전의 인쇄 매체뿐 아니라 새로운 유형의 매체가 지속적으로 개발되어 유통 됨에 따라 물리적으로 표현되는 자료의 특성을 기준으로 매체를 구분할 필요성이 생겼으며 실제로 있어서도 일반인쇄물, 원문(PDF), URL, VOD 등으로 구분하고 있다. 따라서 국가지식정보자원의 분류체계는 주제별 분류항목, 자료유형별 분류항목, 매체유형별 분류항목이 혼합되어 있는 상태이므로 이를 일정한 원칙에 따라 통일성 있는 분류표로 통합할 수 있어야 한다.

3.3.1 자료유형

(1) 문헌분류표의 표준세구분표

DDC나 KDC와 같은 문헌분류표에서는 우선 주제와 학문에 따라 분류하고, 다음으로 필요에 따라 제 1보조표인 표준세구분표에 의해 주제를 표현하는 형식으로 세분한다. 표준세구분표는 형식 가운데 반복적으로 나타나는 사전

이나 정기간행물, 서지 등과 같은 물리적 형식 또는 이론이나 역사, 연구와 같은 접근법 내지는 표현 형식에 대해 공통적인 기호를 부여하기 위해 마련된 조기표로서 많이 사용되는 부호는 다음과 같다 : 01 철학 및 이론, 021 도표, 025 인명록/기관주소록, 027 특허/상표, 03 사전/용어집, 05 연속간행물, 072 연구방법/통계 방법, 076 연습문제/문제집, 08 특정 부류의 사회집단, 09 역사적, 지리적, 인물의 취급

표준세구분표의 일반적인 특성은 크게 두 가지를 들 수 있다. 하나는 해당 문헌이나 자료의 일차적인 주제보다는 자료와 관련된 비주제적 특성을 나타내주는 점이며, 다른 하나는 주제의 범위와 상세도에 관계없이 해당 분류기호 전체에 상당하는 모든 주제의 분류기호에 적용할 수 있다는 점이다. 이 외에 표준세구분표에 의한 패싯(facet)을 적용하기 위한 기호의 특성과 부가지시사항 등의 특성을 지니고 있으나 이를 국가지식정보 표준분류에 적용시키는 것은 무리이다.

(2) MARC 21의 008필드 내용형식

기계가독형 자동화목록형식인 MARC 21의 부호화정보필드에서 표현할 수 있는 단행본의 자료유형은 문헌분류표의 표준세구분에 비해 다양한 유형으로 구분하고 있다 : a - 초록/요약(Abstracts/summaries), b - 서지(Bibliographies), c - 목록(Catalogs), d - 사전, 용어집(Dictionaries), e - 백과사전(Encyclopedias), f - 핸드북(Handbooks), g - 법률기사(Legal articles), i - 색인(Indexes), j - 특허(Patent document), k - 음반해설서(Discographies), l - 법령집(Legislation), m - 학위논문(Theses),

n - 주제별 문헌분석서(Surveys of literature in a subject area), o - 평론(Reviews), p - 프로그래밍 텍스트(Programmed texts), q - 영화해설서(Filmographies), r - 인명록(Directories), s - 통계(Statistics), t - 기술보고서(Technical reports), u - 표준/명세서(Standards/specifications), v - 판례 및 판례주석(Legal cases and case notes), w - 법률 보고서 및 요약(Law reports and digests), z - 조약집(Treaties).

MARC 21에서는 단행본 뿐만 아니라 다양한 형태의 정보자원에 적용할 수 있는 장르(Genre) 코드를 제공하고 있으며 비도서 정보자료에 해당되는 80여종의 항목이 제시되어 있다: 데이터베이스, 웹사이트, 게임, 플래시 카드, 전기(biography), 회의록, 영화, 리허설, 연설, 정부간행물, 그래픽, 녹화자료, 음향자료 등 (<http://www.loc.gov/marc/sourcecode/genre/genrelist.html> 참조).

(3) 종합정보센터의 자료유형

국가지식정보자원관리사업에 의해 지정된 종합정보센터에서 사용하고 있는 분류체계 상의 자료유형은 각 기관마다 상이한 것으로 나타나고 있다. 자료유형으로 분류하는데 있어서 기준이 모호한 첨단 DB, 해외전자정보 등을 제외할 경우 대략 14개의 유형으로 구분할 수 있다: 1 특허/표준/규격, 2 초록/해제, 3 서지/목록, 4 학위논문, 5 연속간행물, 6 사전, 7 문서, 8 정부간행물, 9 통계, 10 세미나/회의간행물, 11 공연/전시, 12 고서, 13 지도, 14 연표.

3.3.2 매체유형

매체유형은 전자정보원의 메타데이터를 기

술하는 요소 중에 포함되어 있다. MARC 21에서는 리더 006에서 기술되며, MARC와 호환성이 높은 MODS(Metadata Object Description Scheme)에서는 정보원의 유형(Type of Resource)을 다음과 같이 정의하고 있다: 텍스트(text), 지도(cartographic), 악보(notated music), 녹음자료(sound recording), 이미지(still image), 영상이미지(moving image), 입체자료(three dimensional object), 소프트웨어/멀티미디어(software/multimedia), 혼합자료(mixed material).

MODS나 DCMI에서는 전자정보원의 물리적 형식과 디지털 표현형식을 보다 구체적으로 설명하기 위해 MIME형식(internetMediaType)을 적용시키고 있다. ASCII, text/html, Postscript file, executable application, JPEG image 등을 예로 들 수 있으며 이들 형식은 자원을 보여주거나 실행시키는데 필요한 소프트웨어나 하드웨어 또는 기타 장치를 결정하는데 사용된다. MODS와 DCMI(Dublin Core Metadata Initiative)에서 전거로 사용하는 IANA(Internet Assigned Numbers Authority)의 MIME type(<http://www.iana.org/assignments/media-types/index.html>)은 자료의 유형에 따라 다음과 같은 디렉토리 리스트가 제공되고 있다: 텍스트(Text), 멀티파트(Multipart), 어플리케이션(Application), 메시지(Message), 이미지(Image), 오디오(Audio), 비디오(Video)

3.4 국가지식정보의 포털 서비스

국가지식정보의 포털 서비스를 위해서는 검색포털의 특성상 다양한 이용자 층이 존재하

로 이들의 요구를 수용하기 위한 분류체계가 마련되어야 한다. 검색포털의 상위개념은 대체로 이용자들의 요구를 대부분 반영하기 때문에 전통적인 분류체계와는 상당한 차이가 있다. 따라서 기본적인 분류의 틀은 학술분류 혹은 문헌분류 체계를 채택하면서 유동적인 구성으로 설계 하는 것을 고려하는 것이 바람직하다.

검색포털의 디렉토리는 주제분류에 유형분류가 다소 혼합된 형태로 구성되어 있는 것이 일반적이며 이용자들 또한 이러한 체계에 익숙해져 있다는 점이 고려되어야 한다. 포털의 이용자들은 디렉토리를 통한 검색보다 통합검색에 익숙해져 있으므로 상세한 분류 전개 시 오히려 이용률을 떨어뜨릴 수도 있다는 점이 감안되어야 한다. 카테고리 명을 부여할 때 일반 이용자들에게 친숙한 용어를 사용하고 카테고리에 키워드를 부여하여 카테고리의 검색결과 또한 유용하게 활용할 수 있도록 하는 것이 필요하다. 각 상위 주제의 분류 시에 제공기관의 분류체계를 부분적으로 이용하는 것은 좋은 방법일 수 있으나, 상이한 전개 방식에 의해 이용자에게 혼란을 주는 것은 피해야 한다.

4. 국가지식정보 분류 표준(안) 모형

4.1 국가지식정보 분류 표준(안) 기반체계와 구조

8개 분야 종합정보센터의 분류 체계는 지식정보자원의 주제만을 분류한 경우와 자료유형과 매체유형 등을 혼합하여 분류한 경우로 구

분된다. 주제와 자료유형 등의 혼합 분류는 지식정보자원의 주제, 자료의 유형 및 매체의 유형을 동시에 표현해야 하는 실용적 측면이 고려된 것이라 할 수 있다. 본 연구에서 제시하는 ‘국가지식정보 표준분류’는 종합정보센터 분류체계와의 상호운영성을 고려하여 주제 분류와 자료유형 및 매체유형을 동시에 표현할 수 있는 구조를 갖추도록 하였으며, 영문 명칭으로는 “Knowledge Classification : KC”를 사용하였다. 분류 방식으로는 일반적으로 문헌분류에서 복합주제나 합성 주제를 조합하기 위한 방법으로 사용되고 있는 패싯(facet)식 분류기법을 적용하였으며, 사용의 용이성을 위해 분류체계 자체를 단순하고 간결하게 구성함으로써 분류의 복잡성을 최소화하였다. 분류체계의 틀은 기본이 되는 주제 패싯과 자료유형 패싯 및 매체 유형 패싯으로 구성되며, 이들 세 패싯의 조합은 기호와 열거순서에 의해 이루어진다(표 4 참조). 자료유형이나 매체유형 분류 시 복수의 유형이 적용될 경우에는 상위 기호에 우선순위를 두어 분류한다.

4.2 국가지식정보 분류 표준(안) 기호법과 열거 순서

4.2.1 기호법

분류기호는 분류표에 있어서 특정하게 배열된 주제의 항목을 기호화한 것으로 숫자, 문자, 부호 등을 사용할 수 있다. 한가지의 분류기호만으로 구성되는 기호를 순수기호라고 하고 두 가지 이상으로 구성되는 기호를 혼합기호라고 한다. 계층화 방식에 따라 구분할 경우 DDC와 같이 0에서 9까지의 숫자를 사용하여 주제 영

〈표 4〉 국가지식정보 분류 표주(안) 기본 구조

주제 분류	자료 유형(Type of Content)	
A. 총류 B. 철학/심리학/종교 C. 사회과학 D. 경영/경제학 E. 법학 F. 자연과학 G. 기술과학 H. 전자/정보통신공학/컴퓨터학 J. 의학 K. 문화/예술 L. 어문학 M. 역사/지리	1. 백과사전, 사전, 용어집 2. 연속간행물(연감) 3. 학위논문 4. 통계 5. 세미나/회의록/보고서	6. 법령, 특허, 표준, 규격 7. 전집, 총서 8. 고서 9. 문서 0. 기타(해당사항 없음)
	매체 유형(Type of Media)	
	1. 원문(Full-text) 2. 메타데이터(서지/목록/초록/요약/해제 등의 Text) 3. 이미지(지도 등의 Image) 4. 녹음자료(Audio) 5. 동영상(Video) 6. 멀티미디어(Multimedia) 7. 컴퓨터프로그램(소프트웨어 등의 Application) 0. 기타(해당사항 없음)	

역을 하위 영역으로 유별화하도록 고안된 십진식 분류표와 LCC, CC와 같이 문자에 숫자나 기호를 결합하여 사용하는 비십진식 분류표로 구분할 수 있다.

국가지식정보 표준분류의 구조는 기본적으로 DDC의 주류(main classes)와 강목(division)의 2단계 분류표를 적용하였다. 그러나 분류기호에 있어서는 정형성과 세분류의 용이성이라는 장점을 가진 십진식에 비해 기호에 구애받지 않고 분류체계를 합리적으로 구성할 수 있으며 전개능력이 뛰어난 비십진식의 장점이 실용적 측면에서 우선적으로 고려되었다. 따라서 열개로 한정되는 십진식을 탈피하여 영문 대문자를 사용하여 주제를 분류하였으며, 보조 패시의 식별이 용이하도록 숫자를 혼합하여 사용하였다. 주제분류표의 주류(대분류)와 강목(중분류)에 각각 영문 대문자 한자리를 부여하여 두개의 영문 대문자로 주제를 표현할 수 있도록 하였으며, 식별상 다른 문자나 숫자와 혼동될 수 있는 I와 O는 제외하였다. 신규항목의 삽

입을 용이하게 하기 위하여 하나의 주류에 두개 이상의 학문분야를 배정한 경우 학문분야 간에 미정의 기호를 남겨두었으며, 각 주제분야의 기타항목은 조기성을 살려서 가능한 Z문자를 사용하였다. 보조 표의 기능을 하는 자료 유형 패시에는 0에서 9까지의 숫자를, 매체 유형 패시에는 0에서 7까지의 숫자를 적용하였다.

4.2.2 열거순서

열거순서(facet order)는 어떤 유(class)에 나타나는 둘 이상의 특성 즉 패시들의 기호를 합성하기 위해 결합할 때 적용하는 순서를 말한다. 예를 들어 문학이라는 피구분체는 문학이라는 '학문' 패시와 '언어' 패시, '문학형식' 패시, '시대' 패시 등으로 분석할 수 있으며, 이 때 발생하는 문제 중의 하나가 이 패시들을 어떤 순서로 열거할 것인지에 관한 것이다. 열거순서의 적용 방식에 따라 합성주제나 복합주제를 다루는 문헌들이 함께 모이거나 분산될 수 있다는 점에서 열거순서는 분류에서 매우 중요한

의미를 갖는다.

열거순서는 이용자의 요구와 필요성에 근거하여 결정해야 하므로 이용자에게 도움이 되는 유용성을 순위 결정의 기준으로 삼는 것이 합리적이다. 그러나 모든 이용자에게 유용한 순서, 그리고 모든 주제에 공통적으로 적용될 수 있는 유용한 순서는 존재하지 않으므로 어떤 순서가 이용자에게 유용한가를 결정하는 것은 쉬운 문제가 아니다.

국가지식정보 표준분류의 열거순서는 보편적으로 자료에 접근하는 방법에 따라 '주제-자료유형-매체유형'의 차례를 따르도록 하였다. 그리고 패시의 숫자가 많지 않고 분류기호를 문자와 숫자를 혼합하여 사용하였기 때문에 각 패시의 구별이 용이할 것이므로 패시 지시기호(연결부호)는 사용하지 않는 것을 원칙으로 하였다. '주제-자료유형-매체유형'의 열거순서는 모든 지식정보에 일관성 있게 적용하며 주제분류표는 반드시 중분류에 해당하는 두 자리 영

문 코드를 부여하도록 하였다.

자료유형은 1에서 9까지의 숫자를 부여하되 기타 유형이나 해당사항이 없는 경우는 0을 적용한다. 매체유형 역시 1에서 7까지를 부여하고 기타 유형은 0을 부여한다. 세 가지 패시 기호를 사용해서 반드시 4자리 분류기호를 유지하도록 해야 한다. 이는 패시의 순서와 기호를 용이하게 식별할 수 있고 일관성 있는 규칙을 검색조건에 적용하기 위한 것이다. 패시 순차와 기호체계 및 분류 사례는 <표 5>와 같다 :

4.3 국가지식정보 분류 표준(안) 주제 분류표

4.3.1 주제 분류의 기본구조

주제 분류의 기본 구조는 DDC의 주류를 토대로 구성되었다. 주류는 분류표에 나타나는 제일 첫 번째의 구분단계로서 분류표의 총체적인 구조에 직접적인 영향을 미치게 되므로 그 구분 정도뿐만 아니라 순서 역시 분류표에서

<표 5> 국가지식정보 표준분류 열거순서 및 기호체계

분류기호(4자리) :

- 주제 대분류(영문1) + 주제 중분류(영문1) + 자료유형(숫자1) + 매체유형(숫자1)

분류(패시)단위		기호체계	비고
주제분류표	대분류(주류)	(영문 1자리) 총 12개 A ~ M	식별상 I 제외
	중분류(강목)	(영문 1자리) 총 190여개 A ~ Z	* 기타항목은 Z 부여 * 식별상 I, O 제외
자료유형 보조표		(숫자 1자리) 총 10개 0 ~ 9	
매체유형 보조표		(숫자 1자리) 총 8개 0 ~ 7	
분류예시		* 통신 표준(원문) HD61 * 영어사전(Web-DB) LU11 * 대동여지도 MH03 * 단곡문집 LC01 * 지리정보시스템 MQ21 (정기간행물 기사)	

중요한 역할을 하게 된다. 일반분류표의 경우 주류는 지식의 모든 영역을 망라하게 되며 이론적으로는 각 주류는 서로 대등한 수준의 것이어야 한다. 그러나 현실적으로 현행 분류표들은 분류표마다 주류의 수와 구분정도 및 대등관계에 있어서 서로 차이를 보이고 있다.

베이컨(F. Bacon)의 학문분류로부터 영향을 받은 DDC의 주류 역시 인문과학에 편중하고 있는 점과 언어와 문학의 분리, 역사와 사회과학의 분리, 그리고 주류의 배열 순서에 많은 문제점이 있는 것으로 지적되고 있다. 이는 한편으로는 DDC를 개발한 듀이(M. Dewey)가 주류의 배열에 관심이 부족했던 것에 기인한다. 그렇지만 다른 한편으로는 DDC 서문에서 듀이 스스로가 '실용주의자'임을 밝혔듯이 총체적인 배열이 반드시 이론적이거나 논리적인 필요는 없다는 입장에서 나온 것이라고 할 수 있으며, 문헌분류표가 철학적 이론 및 정확성보다 실제적 유용성을 우선시 한다는 것을 의미하는 것이기도 하다.

국가지식정보 표준분류의 주제 분류 역시 지식정보자원의 이용과 관리의 실용성을 고려하였다. 지식의 전체 영역을 포괄하고 있는 DDC를 기본 모델로 하되 주류와 강목을 10개의 항목으로 한정하지 않고 강목 이하의 수준일 경우에도 분야의 범위가 넓고 문헌의 출판량과 이용도가 높아 독립성이 인정되는 분야는 주류나 강목으로 상향 이동시켰다.

주제분류표의 기호는 영문 대문자를 채택함으로써 주류의 숫자에 제한을 두지는 않았으나 주제분류 체계에 따라 브라우징 검색항목을 설정하게 된다는 점을 고려하여 12-13개의 항목 이하로 제한하였다. 각 주류는 다시 2단계의 강

목으로 세분되며 역시 분야에 따라 전개되는 항목의 수에 유연성을 부여하였다.

주제 항목의 설정에 있어서 DDC 외에 한국적 상황에서의 문헌 특성을 반영하기 위해 KDC를 참고하였으며, 주류항목과 순서를 결정하는데 있어서는 LCC나 CC와 같은 분류표, 한국학술진흥재단의 연구분야분류표, 과학기술정보표준화위원회의 과학기술정보분류 등의 비교 결과를 반영하였다. 또한 분야별로 구성되어 있는 여덟 개의 종합정보센터에서 사용되고 있는 분류표(또는 주제별 디렉토리 검색항목)를 참조하였다. 이들 분류표의 분석 결과는 종합정보센터마다 분류 기준과 분류 수준 및 계층구조가 상이한 것으로 나타났다. 주제가 매우 세분화되어 있는 경우와 자료의 유형이나 특성에 의해 간략하게 구성한 경우를 고려하여 국가지식정보 표준분류는 가능한 각 기관에서 사용되고 있는 주제와 형식을 포괄할 수 있도록 구성하였다. 특히 2004년 한국전산원에서 수행한 국가지식정보자원 현황조사에서 주제별 정보자원을 파악하여 정보량이 많은 분야가 분류체계의 계층에 반영될 수 있게 하였다. 그러나 정보량이 가장 많은 것으로 나타난 교육학술 분야의 경우 교육학에 해당되는 자원보다 학위논문이나 대학의 연구논문집과 같은 다양한 주제 분야를 포괄하는 것이어서 교육학을 주류에 배정하지는 않았다.

4.3.2 주제 분류표

국가지식정보 표준분류의 주제분류표는 주제의 범주를 크게 총류, 인문과학, 사회과학, 자연과학, 기술과학, 어문학으로 구분하였으며, 2단계 분류체계를 유지하기 위해 강목과 세목

이하의 주제 분야를 상위 항목으로 이동하였다. 이에 따라 사회과학의 경영/경제와 법학은 각각 주류 D, E로 독립되었고 전자/정보통신공학/컴퓨터학과 의약학 역시 H와 J로 기술과학으로부터 독립되어 주류로 배정되었다. 또한 어문학 아래의 중분류 중 한국문학의 문학형식, 역사/지리 아래의 한국역사의 시대구분과 지리구분은 동위류의 다른 항목에 비해 많은 기호를 배정함으로써 한국적 특수성을 반영하였다. 그러나 주류 B의 철학/심리학/종교의 경우 DDC에서는 2개의 주류로 구분되어 있는 것을 정보의 양과 이용률을 고려하여 하나의 주류로 병합하였다. 주류의 처음에 나오는 총류 및 지식일반은 포괄적이고 학제적인 정보자료를 분류하는데 적용되는 것으로 문헌분류에서 필요한 도구적 장치에 해당한다. 타 교육기관이나 해외 데이터베이스와의 교환가능성을 위해 모든 분류항목은 DDC기호와 매칭되도록 테이블을 함께 제시하였으며, 종합정보센터에서 이미 DDC로 분류되어 있는 국가지식정보자원은 그대로 적용될 수 있도록 하였다. 다만 기타 학문 분야, 한국의 역사구분, 지리구분과 같이 새롭게 전개한 항목에는 DDC기호를 부여하지 않았

다(부록 참조).

4.4 국가지식정보 분류 표준(안) 자료유형 분류표

도서관에서 사용하는 분류표상의 자료유형은 자료의 비주제적 특성을 보여주기 위한 것으로 일반적으로 주제 분류와 구분하여 정보내용의 특성을 표현한다. 기계가독형 자동화목록형식(MARC)에서도 자료유형은 주제로 구분하지 않고 부호화정보필드(008) 중 내용형식 필드에서 기술하도록 되어 있다. 국가지식정보 표준분류의 자료유형을 구분하기 위해 문헌분류표의 표준세구분표, MARC 21의 008 필드 내용형식, 종합정보센터 자료유형을 분석하여 참조하였다.

국가지식정보 표준 분류에서 주제분류에 부가되는 첫 번째 패킷인 자료유형은 각 종합정보센터에서 사용하고 있는 유형을 수용하는 범위로 한정하였으며 기호는 아라비아 숫자 코드 0-9를 사용하였다. 1-9에 해당되지 않는 기타 항목은 0에 배정하였으므로 반드시 자료유형 코드를 사용해야 한다(표 6 참조).

〈표 6〉 국가지식정보 표준분류 자료유형 분류표

- | |
|--|
| <ol style="list-style-type: none"> 1. 백과사전, 사전, 용어집 2. 연속간행물(연감) 3. 학위논문 4. 통계 5. 세미나/회의록/보고서 6. 법령, 특허, 표준, 규격 7. 전집, 총서 8. 고서 9. 문서 0. 기타(해당사항 없음) |
|--|

4.5 국가지식정보 분류 표준(안) 매체유형 분류표

정보기술의 발전에 따라 인쇄물 중심이었던 정보자원들이 다양한 전자정보 매체로 제공되고 있으며, 이에 따라 동일한 자료유형에서도 매체의 유형이 다른 경우가 발생한다. 연속간행물이나 학위논문의 경우 인쇄물과 전자형식의 원문으로 각각 구축되어 이용자에게 서비스되고 있는 것이 그 예이다. 국가지식정보 표준분류의 제2 보조표에 해당하는 매체유형의 분류에서는 메타데이터 기입과의 호환성을 유지하기 위하여 IANA의 MIME 매체유형을 수용하였으며, 다만 메시지의 경우 본 연구의 적용대상이 되는 지식정보자원의 범위에 포함되지 않아 적용에서 제외하였다. 이렇게 함으로써 정보원의 유형(Type of Resource)에 가까운 형태로 구성하였으며 아라비아 숫자 0에서 7을 기호로 채택하였다. 기호 8과 9는 확장이 가능하도록 미정의 기호로 남겨 놓았다(표 7 참조).

5. 결 론

본 연구에서는 다양한 주제 분야와 다양한

자료유형 및 매체유형으로 이루어져 있는 지식정보자원을 포괄할 수 있는 분류 체계의 표준모형을 제시하였다. 표준 모형으로 제시된 국가지식정보 분류 표준(안)의 개발을 위하여 먼저 문헌분류 체계와 지식정보자원 분류 체계의 차이, 지식정보자원의 분류사례 및 지식정보자원 분류에 대한 문헌분류 체계의 적용 가능성에 관한 기초 분석을 실시하였다. 이러한 기초작업을 통해 국가지식정보 표준분류 체계 개발에 필요한 지식의 분야별 분포, 분류의 계층과 유연성 확보 방안, 지식정보자원의 자료유형과 매체유형, 국가지식정보의 포털서비스 방안 등을 도출하였다. 여러 측면에서의 분석을 토대로 본 연구에서는 DDC를 수정 적용한 주제분류를 기본으로 하면서 다양한 자료유형과 매체유형을 수용할 수 있는 다원적 국가지식정보 분류 표준(안)을 개발하였다.

본 연구에서 제시된 국가지식정보 분류 표준(안)은 그 자체로서 향후 지식정보자원관리 중장기 계획을 수립하는데 활용될 수 있으며, 국가지식정보자원의 개발 전략과 방향제시의 기초도구로 활용될 수 있다. 특히 본 표준(안)이 향후 관련 사업의 표준분류체계로 확정되어 실제에 적용될 경우에는 국가 차원에서의 지식정보자원에 대한 현황 파악과 관리 기능이 체계

〈표 7〉 국가지식정보 표준분류 매체유형 분류표

- | |
|--|
| <ol style="list-style-type: none"> 1. 원문(Full-text) 2. 메타데이터(서지/목록/초록/요약/해제 등의 Text) 3. 이미지(지도 등의 Image) 4. 녹음자료(Audio) 5. 동영상(Video) 6. 멀티미디어(Multimedia) 7. 컴퓨터프로그램(소프트웨어 등의 Application) 8. 기타(해당사항 없음) |
|--|

적으로 이루어 질 수 있게 될 것이다. 이는 지식 정보자원의 관리환경을 개선하는데 많은 기여를 할 것이며, 공공과 민간 분야에서 지식정보

자원을 관리하는데 있어서도 시스템의 중복구축이나 디지털화 업무의 중복이 방지될 수 있을 것이다.

참 고 문 헌

- 고영만. 2004. 국가지식정보자원 분류체계 및 현황평가 방법론. 한국전산원.
- 서태설. 2003. 과학기술정보 분류 표준 개발. 제4차 과학기술정보표준화위원회 총회 및 세미나. 서울. 2003.11.19.
- 설성수, 송충한. 2000. 『지식활동분류의 이론과 실제』. 대전: 한남대학교 출판부.
- 정보통신부. 한국. 2000. 지식정보자원관리 기본계획(안).
- 정보통신부. 한국. 2001. 종합정보센터지정 등 분야별 지식정보유통체계 확립방안(안).
- 조동성. 1993. 공익 DB의 효과적 개발과 민간업체의 제도적 참여 방안. DB 산업활성화를 위한 세미나, KOEX.
- 한국데이터베이스진흥센터. 1997. 데이터베이스 분류(DPCS 001), 1997.12.4 제정.
- 한국데이터베이스진흥센터. 2003. 국내데이터베이스 산업 현황 조사.
- 한국소프트웨어진흥원. 2004. 디지털 콘텐츠 산업분류체계에 관한 연구. 2004. 2.
- 한국전산원. 1998. EC 기술체계 분류 및 Life Cycle 분석. NCA IV-RER-97128 / 1998.6.23
- 한국전산원. 2001. 국가지식정보자원 디지털화 로드맵 수립. NCA III-RER-01081 / 2001. 11.
- 한국전산원. 2002. 국가정보화 백서.
- 한국전산원. 2002. 가치있는 지식 창출 및 공유활성화를 위한 지식커뮤니티에 관한 연구.
- 한국전산원. 2003. 공공지식정보자원의 부가가치 창출방안 연구. NCA III-RER-03088 / 2003. 12.
- 한국전산원. 2004. 국가지식정보자원 현황조사 및 중장기 개발계획 연구. 2004.10.31.
- 한국학술진흥재단. 1999. 학문분류표 재설정에 관한 연구. 1999.11
- OCLC/ALISE. 2001. "Global Collective Resources: A Study of Monographic Bibliographic Records in WorldCat", OCLC/ALISE 2001 Research Grant Report
- U.K. DTI. 1996. Government-held Tradeable Information Guidelines for government departments in dealing with the private sector. London: Department of Trade and Industry.
- CyberStacks.
 <<http://www.public.iastate.edu/~CYBERSTACKS/>>
- Scout Report Signpost.
 <<http://signpost.org/index.html>>

〈부 록〉 국가지식정보 표준분류 주제 분류표

대분류	중분류	DDC기호
A. 총류 지식일반	AA 총류	000
	AB 지식, 학문일반	001
B. 철학/심리학/종교	BA 철학일반	100
	BB 심리학	150
	BC 논리학	160
	BD 윤리학	170
	BE 고대철학, 중세철학	180
	BF 동양철학	181
	BG 한국철학	181.119
	BH 현대서양철학	190
	BJ 미학/예술학	111.85
	BK 종교일반	200
	BL 비교종교	200
	BM 기독교	230
	BN 불교	294.3
	BP 바라문교/인도교	294.5
	BQ 유대교	296
	BR 회교(이슬람교)	297
	BS 유교	299.512
	BT 도교	299.514
	BU 한국종교	299.57
	BV 기타제종교	299.9
C. 사회과학	CA 사회학	301
	CB 사회복지학	361
	CC 인류학	301
	CD 정치외교학	320
	CE 행정학	351
	CF 국방/군사학	355
	CG 교육학	370
	CH 신문방송학	070
	CJ 문헌정보학	020
	CK 지역개발	307
	CL 풍속/민속학	390
	CZ 기타 사회과학	300
D. 경영/경제학	DA 경영학일반	658
	DB 재무관리	658.1
	DC 인사관리	658.3
	DD 조직관리	658.4
	DE 생산관리	658.5
	DF 물류관리	658.7
	DG 시장관리	658.8
	DH 회계학	657
	DJ 광고학	659
	DL 경제학일반	330
	DM 수리/미시경제	338.5
	DN 노동경제	331
	DP 금융경제	332

	DQ 토지/에너지경제	333
	DR 재정/공공경제	336
	DS 국제경제	337
	DT 거시경제	339
	DU 각국의 경제	330.9
	DV 농업경제학	338.1
	DW 무역학	380
	DY 보험학	368
E. 법학	EA 법학일반	340
	EB 국제법	341
	EC 헌법	342
	ED 행정법	342.06
	EE 형법	345
	EF 민법	346
	EG 조세/세법	343.04
	EH 경제법/산업법	343.07
	EJ 노동법	344.01
	EK 사회보장/사회법	344.02
	EL 교육법/문화법	344.07
	EM 환경/교통법	344.046
	EN 의료/보건법	344.041
	EP 각국의 법	349
	EZ 기타 제법	
F. 자연과학	FA 자연과학 일반	500
	FB 수학	510
	FC 통계학	519.5
	FD 천문학	520
	FE 물리학	530
	FF 화학	540
	FG 지구과학/지질학	550
	FH 고생물학(화석학)	560
	FJ 생명과학/생물학	570
	FK 식물학	580
	FL 동물학	590
	FZ 기타 자연과학	
G. 기술과학	GA 기술과학일반	600
	GB 전기공학	621.3
	GC 수력공학	621.2
	GD 기계공학	621.8
	GE 자원공학	622
	GF 해양공학	623
	GG 토목공학	624
	GH 교통공학	625
	GJ 위생공학/환경공학	628
	GK 항공/우주공학	629.1
	GL 자동차공학	629.2
	GM 농업/농학	630
	GN 축산학	636

	GP 수산학	639
	GQ 가정학(생활과학)	640
	GR 화학공학	660
	GS 금속공학	669
	GT 제조업	670
	GU 건축공학	690
	GZ 기타 기술과학	
H. 전자/정보통신공학/컴퓨터학	HA 반도체	537,622
	HB 신경망	006.32
	HC 제어계측시스템	621.37
	HD 통신시스템	384
	HE 마이크로파	537.534
	HF 회로시스템	621.395
	HG 신호처리	006.4
	HK 기타 전자/정보통신공학	621.38
	HP 컴퓨터시스템	004.3
	HQ 시스템소프트웨어	005.43
	HR 컴퓨터구조	004.7
	HS 데이터베이스	005.74
	HT 소프트웨어공학	005.1
	HU 인공지능	006.3
	HV 컴퓨터그래픽스	006.6
	HW 프로그래밍언어	005.13
	HX 정보보호	005.8
	HZ 기타 컴퓨터학	
J. 의학학	JA 의학일반	610
	JB 해부학	611
	JC 생리학	612
	JD 공중보건/예방의학	614.4
	JE 내과학	616
	JF 외과학	617
	JG 치의학	617.6
	JH 안과학	617.7
	JJ 이비인후과학	617.8
	JK 산부인과/소아과	618
	JP 약리학	615
	JQ 간호학	610.73
	JR 한의학	610.9
	JZ 기타 제의학	
K. 문화/예술	KA 문화·예술일반	700
	KB 도시계획 및 조경	710
	KC 건축	720
	KD 조형/조각	730
	KE 서예	745.61
	KF 공예/장식미술	740
	KG 회화	750
	KH 사진	770
	KJ 음악	780
	KK 레크레이션/오락	790

	KL 연극/영화	791
	KM 무용	792.8
	KN 체육/스포츠	796
	KZ 기타 예술	
L. 어문학	LA 문학일반	800
	LB 한국문학일반	895.7
	LC 한국시	895.71
	LD 한국희곡	895.72
	LE 한국소설	895.73
	LF 한국수필	895.74
	LG 한국연설	895.75
	LH 한국서간/일기/회고록	895.76
	LL 중국문학	895.1
	LM 일본문학	895.6
	LN 영미문학	810
	LP 기타 제문학	890
	LQ 언어학일반	400
	LR 한국어	495.7
	LS 중국어	495.1
	LT 일본어	495.1
	LU 영어	420
	LZ 기타 제어	490
M. 역사/지리	MA 전기(biography)	920
	MB 족보	929
	MC 고고학	930.1
	MD 역사일반, 세계사	900
	ME 한국역사(통사)	951.9
	MF 고대(조조선, 삼국시대) ~ 936 A.D.	
	MG 고려시대(高麗時代) 중세(中世) 918~1392	
	MH 조선시대(근세) 1392~1910	
	MJ 항일시대(抗日時代) 최근세(最近世) 1910~1945	
	MK 1948~ 대한민국	
	MQ 지리, 여행일반	910
	MR 한국지리/여행(관광)	915.19
	MS 서울특별시	
	MT 경기도	
	MU 강원도	
	MV 충청도	
	MW 경상도	
	MX 전라도	
	MY 제주도	
	MZ 북한	