

상호운용적 분류체계 관리를 위한 반자동 분류체계 관리방안

Semi-Automatic Management of Classification Scheme with Interoperability

이원구, 신성호, 김광영, 정도현, 윤화목, 성원경, 이민호
한국과학기술정보연구원 정보기술연구실

Won-Goo Lee(wglee@kisti.re.kr), Sung-Ho Shin(maximus74@kisti.re.kr),
Kwang-Young Kim(kykim@kisti.re.kr), Do-Heon Jeon(heon@kisti.re.kr),
Hwa-Mook Yoon(hmyoon@kisti.re.kr), Won-Kyung Sung(wksung@kisti.re.kr),
Min-Ho Lee(Cokeman@kisti.re.kr)

요약

과학기술의 융·복합현상은 21세기 지식 기반 경제하에서 더욱 활발하게 진행됨에 따라 과학기술 분야를 적절히 분류해내고, 미래의 신성장 분야까지 포용할 수 있는 체계를 만드는 것이 결코 쉽지 않다. 이에, 본 연구에서는 각 콘텐츠 관리·서비스 기관이 분류체계 간 상호운용성을 갖을 수 있도록 반자동적인 입수/관리 분류체계 이력관리 및 입수-관리 분류체계 간 매핑 방안을 시스템적 측면에서 제시하였으며, 이를 통해, 기존의 수작업 방식에서 발생할 수 있는 관리적 어려움과 비용적 발생을 최소화할 수 있을 것이다.

■ 중심어 : 주제분류 | 분류체계 | 주제어 자동분류 |

Abstract

Under the knowledge-based economy in 21C, the convergence and complexity in science and technology are being more active. Therefore, we have science and technology are classified properly, make not easy to construct the system to new next generation area. Thus we suggest the systematic solution method to flexibly extend classification scheme in order for content management and service organizations. In this way, we expect that the difficult of classification scheme management is minimized and the expense of it is spared.

■ keyword : | Topic Classification | Classification Scheme | Automatic Keyword Classification |

1. 서론

과학기술 콘텐츠의 의미적 상호운용성 문제는 지속적으로 대두되어 왔으며, 더불어 콘텐츠의 관리 및 서비스를 위한 콘텐츠 분류에 대한 필요성 또한 꾸준히 제기되어 왔다. 이에 국가적 차원에서, 국제적 표준제정 기관 차원에서 또는 과학기술 분야 콘텐츠 관리 및 서비스 기관 자체적으로 분류체계를 제정하여 사용하고

있다. 이러한 여러 기관으로부터의 다양한 분류체계의 사용은 상호운용성 문제를 야기하였고, 체계적 관리 및 효율적인 서비스를 위해 여러 가지의 분류체계를 어떻게 조화시켜 줄 수 있는가에 대한 문제 또한 지속적으로 대두되고 있다[1].

더욱이, 과학기술의 융·복합현상은 21세기 지식기반경제 하에서 더욱 활발하게 진행되어져 왔고, 국가별·연구주체별로 다양한 과학기술의 융·복합 관련

연구들이 진행되면서 매우 복잡한 양상으로 다변화되어 오늘에 이르고 있으며, 향후에도 더욱 다양한 분야로 분화되고 융합될 것으로 쉽게 예측할 수 있다. 따라서 이러한 과학기술 분야를 적절히 분류해내고, 미래의 신성장 분야까지 포용할 수 있는 체계를 만드는 것이 결코 쉽지 않은 일임은 자명하다[2]. 즉, 기존의 과학기술 콘텐츠의 분류체계는 매우 빠르게 증가하고, 변화하는 융·복합기술, 기술융합, 융합기술을 전체적으로 내포(內包) 및 신기술을 모두 담아내는데 그 한계를 가지고 있다[3].

이에, 많은 연구에서는 과학기술의 융·복합 현상을 동적인 흐름으로 파악하고, 기술 융합형태에 따른 특성을 분석하여 종합적인 관점에서 새로운 분류체계를 제안하는 것에서부터, 다양한 과학기술분류체계를 아우를 수 있는 기본 개념체계를 만들고, 이를 통해서 분류체계 간의 상호운용성을 점검할 수 있는 방법론을 개발하는 것, 특정 분류체계에 대한 분류규정 및 분류기법의 개선방향을 모색하는 것 등 분류체계 문제를 해결하기 위한 다양하고도, 많은 노력을 기울여 왔다[9-11]. 이것은 과학기술계 뿐만 아니라 정보서비스 분야의 전반적인 요구이기도 하고, 정보유통 개발 및 서비스 업무를 보다 체계적으로 수행하기 위해서는 분류체계가 일관된 체계로 재편성하거나 새롭게 개발할 필요가 있다는 것을 충분히 인식하고 있다. 하지만 이러한 필요가 발생할 때마다 새로운 분류체계를 만든다면 그것은 또 하나의 분류체계가 추가되고, 관리·적용할 대상이 하나 더 늘어나는 결과를 초래하여 문제를 더 어렵게 만들 것이다.

이에 본 연구에서는 분류체계 문제의 해결을 분류체계에 대한 새로운 체계 또는 개선된 체계를 제안하는 것이 아니라, 다른 각도에서 분류체계 관리방식을 조명하고자 한다. 즉, 위와 같은 노력에도 불구하고 과학기술 콘텐츠 분류체계는 지속적으로 개선되고, 새로운 분류체계가 공표될 것임은 자명하다. 다만, 이는 수용하는 기관이 이를 어떻게 적용할 것인가의 문제로 봉착될 것이며, 본 연구는 각 기관이 분류체계 간 상호운용성을 갖을 수 있도록 분류체계를 유연적으로 수용·확장하기 위한 시스템적 적용방안 내지는 해결방안을 제시하

고자 한다.

이를 위해, 본 논문의 2장에서는 현재의 분류체계 현황에 대해 기술하고, 3장에서는 기존의 분류체계에 대한 관리의 문제점을 살펴보고, 4장에서는 이를 해결하기 위해 제안된 반자동 분류체계 관리방안에 대해 파싱 및 매핑 방안에 대해 설명하고자 한다.

II. 분류체계

2.1 개요

분류체계(classification scheme)라 함은 “유사 특성(형식, 주제, 품목, 관계 등)을 갖는 개체들을 그룹으로 배열(clustering) 또는 구분(distinguishing)하기 위한 체계”라고 할 수 있다[1]. 이러한 분류는 구조적인 면에서 몇 가지 형태로 구분되는데, ① List 형: 무차원 list(ordered/unordered), 차원 list ② Array 형: table(unordered array), matrix(ordered array) ③ Tree 형: single root(taxonomy), multiple roots ④ Network 형: directed graph, map 등으로 구분할 수 있으며, 이 중에서 과학기술분류에 가장 많이 사용되는 구조는 Tree형이다. 이러한 Tree 형에는 포함관계(is a)를 기반으로 하여 상속이 이루어지는 계층형(hierarchy)과 포함관계와 상속이 없고 하나의 차원에 따라 지식을 표현하는 비계층(non-hierarchy)형이 있다[1].

2.2 분류체계 현황

[표 1]과 같이 우리가 과학기술 콘텐츠를 관리·서비스할 때 참조하는 분류체계로는 한국연구재단의 “연구분야 분류”, 지식경제부의 “산업기술 표준 분류”, 과학기술정책연구원의 “국가 과학기술 표준 분류”, 소위 6T(IT, BT, NT, ST, ET, CT) 분류하고 하는 “미래 유망 신기술 분류”, 비전 로드맵과 함께 제시된 “국가 기술지도 분류”, 일반적 도서분류법인 “DDC(Dewey Decimal Classification)”와 국제십진분류법인 “UDC(Universal Decimal Classification)” 등이 있다[9]. 이외에도 KDC(Korean Decimal Classification), CIP(Cataloging In Publication), IPC(International Patent

Classification) 등과 각 기관별로 제정된 UCAT, DIMD, DIGS, KEST, OCIS, 과총 분류, KISTI 표준분류와 같은 다양한 코드 가 존재하고 있다[5][6]. 본 절에서는 우리가 흔히 적용하고 있는 과학기술 분류체계에 대해서만 살펴보고자 한다.

표 1. 기존의 분류체계

연구분야 분류				
분야코드	대분류	중분류	소분류	세분류
산업기술 표준 분류				
대분류	중분류	소분류	Code No.	
국가 과학기술 표준 분류				
대분류	중분류	소분류		
미래 유망 신기술 분류				
구분	코드값	기술명	분류기준	
국가 기술지도 분류				
코드값	핵심기술	분류기준		
KISTI 표준 분류				
대분류	중분류	소분류		
Dewey 십진 분류				
000(총류), 100(철학), 200(종교), 300(사회과학), 400(언어학), 500(순수과학), 600(응용과학), 700(예술), 800(문학), 900(역사)				

본 연구의 특성(모든 분류체계에 대한 상호운용성 연구)상, 기존의 분류체계에 대한 특징이나, 문제점 또는 개선점에 대해서 구체적으로 논의하지는 않지만, 문제점에 대해서만 간략하게나마 기술하기로 한다.

“연구분야 분류”는 세분류까지 분류함으로써 관련 분야 검색 시 적절한 연구분야 검색결과를 얻기 어려울 것으로 보인다. “산업기술 표준 분류”는 향후 융·복합에 의하여 생성될 다양한 분야들을 수용할 수 있는 체계로는 부적합하다. “국가 과학기술 표준 분류” 및 “KISTI 표준 분류”는 앞의 두 분류체계에 비해 좀 더 포괄적인 분야를 중분류로 수용하는 듯하나, 그 범위가 매우 넓어 그 안에 속한 과학기술 분야들을 적절히 수용할 수 없을 것으로 보인다. “미래 유망 신기술 분류”는 중분류 및 소분류를 두지 않고 곧바로 ‘기술명’과 ‘분류기준’을 제시함으로써 융통성 있는 수용 체계를 갖추

었지만, 다양한 분야에서 융·복합되고 있는 콘텐츠 기술들을 적절히 수용하는 데는 한계가 있어 보인다. “국가 기술지도 분류”는 다소 융통성 있는 분류체계를 형성하고 있지만, 국가가 정한 로드맵 비전을 중심으로 정리한 것이므로 미래지향적으로 수용하는 데는 한계를 드러낼 수밖에 없다[2]. 그 외에 DDC 체계 또한 여전히 학제적인 주제영역의 분류에 있어 문제점이 지적되고 있고 나아가 전통적인 지식의 분류체계에 대한 융통성 문제가 제기되고 있다[4].

2.3 기존 분류체계 관리의 문제점

2.3.1 기존의 분류체계 관리

일반적으로 분류체계에 대한 관리는 각 문헌의 메타데이터에 각 분류체계의 항목을 기술하는 정도의 수준이다. 이러한 관리 형태는 분류체계의 개선 시, 버전 이력 내지는 버전별 증적자료 관리를 어렵게 하며, 이는 기존의 분류체계 형태로 구분지어진 문헌에 대한 서비스를 더욱더 어렵게 하고 있다. 더욱이 새로운 분류체계의 도입은 기존 분류체계 적용 문헌에 대한 관리 및 서비스에 치명적인 악영향을 미치게 된다.

우선, 기존의 분류체계 관리, 즉 메타데이터 상의 분류체계 표기에 대해 살펴보고자 한다.

아래의 [표 2]는 한국과학기술정보연구원에서 최근에 개발 중인 관리시스템에 포함된 학술논문/학위논문/특허/연구보고서/동향·분석/산업·표준 콘텐츠의 메타데이터 상에 나타난 분류체계 관련 항목을 나타낸 것이다. 학술논문은 분류체계가 너무나 많아 기술하지는 않지만, 단순히 분류코드로만 관리하고 있는 상황은 다른 콘텐츠와 마찬가지로였다[7].

특히(이미 국제적으로 버전관리 활동을 수행 중임) 콘텐츠에 한해 <분류명>과 <분류버전명>을 통해 버전(이력)관리를 하고 있는 것으로 나타났지만, 대부분의 문헌의 분류체계는 <분류명>과 <분류코드> 형태로만 관리되고 있었으며, 이러한 형태의 관리는 본 조사기관뿐만 아니라 대부분의 콘텐츠 관리 및 서비스 기관에서도 상기와 같은 형태로 관리되고 있다.

표 2. 기존의 분류체계

<p>학위논문 DB : 주제분류코드, CC (ex, C64, S56, S12 등) XML : article-meta/article-categories/subj-group@subj-group-type/subject@content-type, (classification, dimd)</p>
<p>특허 DB : 대표IPC분류코드, RPRST_IPC_CLSFC_CD, IPC분류버전명, IPC_CLSFC_VRSN_NM XML : bibliographic-data/classification-ipc/main-classification, further-classification, bibliographic-data/classification-national@country/main-classification, bibliographic-data/classification-ipc/main-classification, bibliographic-data/classification-ipc/edition</p>
<p>국가 과학기술 표준 분류 DB : 과학기술표준분류명, SCTEC_STND_CLSFC_CD XML : article-meta/article-categories/subj-group@subj-group-type/subject@content-type, (classification, stcd)</p>
<p>산업표준 DB : 규격분야코드, STND_FLD_CD XML : article-meta/article-categories/subj-group@subj-group-type/subject@content-type</p>
<p>동향분석 DB : 기술산업분류코드, TREND_ANLS_SUBJ_CLSFC_CD, 규격분야코드, 차세대분류코드, NXGEN_TECH_CD XML : article-meta/article-categories/subj-group@subj-group-type/subject@content-type, (classification, trend_theme tech ngtch ...)</p>

② 다른 분류체계/같은 분류체계의 새로운 버전의 입수에 따른 기관 표준분류체계로의 적용이 어려우며, 이를 해결하기 위한 기관 표준분류체계의 변경/확장 또한 어렵고, 마찬가지로 임의 분류된 문헌에 대한 서비스는 더욱 어렵다.

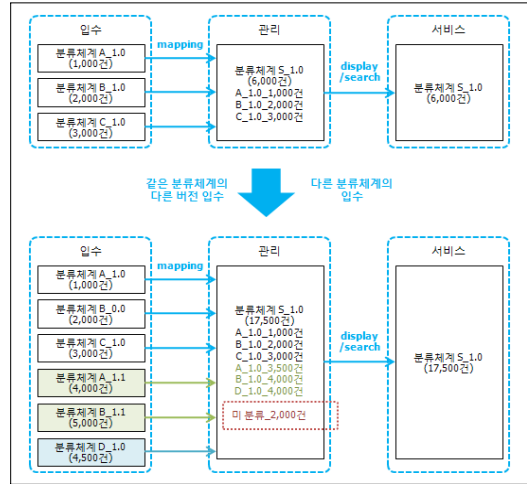


그림 1. 입수 분류체계 변경의 문제점

2.3.2 기존의 분류체계 관리의 문제점

[그림 1]은 입수 분류체계 관리에 대한 기존 방식의 문제점을 보여준다. 즉, 기존의 분류관리체계 S_1.0으로 입수된 각각의 분류체계 A_1.0, B_1.0, C_1.0에 대해 각각 1천건, 2천건, 3천건 입수되었다고 가정할 때, 이후 관리 분류체계는 변경되지 않고, 입수 분류체계만 각각 A_1.1, B_1.1, C_1.1로 변경되었다면 기존의 관리 분류체계로는 이를 다 수용할 수 없는 문제가 발생할 수 있으며, 이에 대한 관리적 어려움의 문제 또한 발생할 수 있다.

이러한 관리의 문제점을 요약하면 다음과 같다.

① 다른 분류체계/같은 분류체계의 새로운 버전의 입수에 따른 기관 표준 분류체계와의 새로운 매핑 작업을 상황이 발생할 때마다 수동으로 진행해야 하며, 일괄적으로 관리 데이터베이스를 업데이트 또는 마이그레이션을 통한 변환을 수행해야 한다.

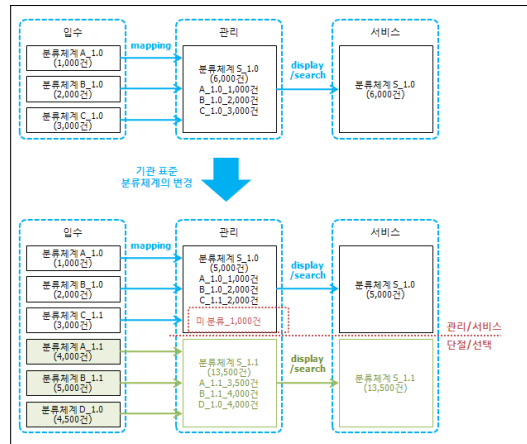


그림 2. 관리 분류체계 변경의 문제점

[그림 2]는 [그림 1]에서와 같이 기존의 관리 분류체계 S_1.0으로 각각의 입수체계가 A_1.0, B_1.0, C_1.0가 입수되었다고 가정할 때, 이후 관리 분류체계가 S_1.1로 버전 업 되었다면, 신규 입수 분류체계 A_1.1, B_1.1,

C_1.1는 자연스레 신규 관리 분류체계 S_1.1로 매핑 가능하지만, 기존에 분류된 6,000건에 대한 신규분류체계로의 매핑의 어려움과 동시에 새로운 입수 분류체계 13,500건에 대한 신규 관리 분류체계로의 매핑 또한 어려울 수 있다. 그리고 이러한 관리의 어려움은 주제 분류 서비스에도 많은 어려움을 초래할 수 있다.

표준 분류체계의 변경에 따른 기존 관리방식에서 나타날 수 있는 문제점을 요약하면 다음과 같다.

- ① 우선, 관리 분류체계가 변경됨에 따라 입수 분류체계가 같은 문헌에 대해서도 새로운 다른 분류체계로의 매핑(mapping)이 발생할 수 있으며, 기존 관리 분류체계로 분류된 문헌과 새로운 관리 분류체계로 분류된 문헌을 동시에 관리함으로써 관리체계의 혼선을 야기할 수 있다. 또한, 상기 방식의 문제점과 마찬가지로, 이를 해결하기 위한 기관 표준분류체계의 변경/확장 또한 임의적 분류 또는 기타 분류로 처리해야 함으로써 매우 어렵고, 마찬가지로 임의 분류된 문헌에 대한 서비스는 더욱 어려울 것으로 전망된다.
- ② 관리 분류체계의 상이함은 서비스에도 영향을 미치게 된다. 즉, 기존 관리 분류체계로 분류된 문헌에 대해 서비스 또한 같은 서비스 분류체계를 가지고 있다. 하지만, 새로운 관리 분류체계가 적용된다면, 서비스 또한 새로운 관리 분류체계를 준용하여 서비스 체계를 가지고 갈 것이다. 이 때, 기존의 관리 분류체계로 분류된 문헌에 대한 서비스는 미실행되거나, 실행을 위해 새로운 분류체계로의 대량의 데이터에 대한 일괄 업데이트를 수행해야만 할 것이다.

이에, 본 연구에서는 같은 입수 분류체계의 다른 버전 입수 및 다른 분류체계의 입수 문제, 그리고 다른 입수 분류체계의 입수 문제, 기관 표준 관리 분류체계의 변경 문제 등 상기에 기술된 다양한 문제에 대해 다음과 같은 해결 방안을 제시하고자 한다.

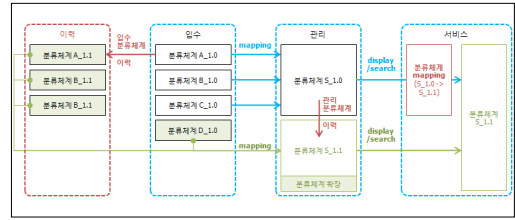


그림 3. 입수/관리 분류체계 개념도

우선 제안하는 방안에 대한 개념은 그림 3과 같이 입수 및 관리에 대한 분류체계 이력관리를 포함하여 입수-관리-서비스를 단일화된 프로세스로 정립하는 것이다. 그리고 기본적인 인터페이스를 도식화하면 [그림 4][그림 5]와 같고, 입수-관리와 입수/관리 분류체계의 이력 및 매핑 관계를 보여주는 관리자 뷰 인터페이스이다. 실질적으로는 관리자에게 입수/관리 분류체계 관리를 위한 mapping 및 tracing 도구를 제공함으로써 관리자가 보다 편리하게 분류체계를 구조적으로·체계적으로 관리할 수 있게 하고자 한다.

이를 위해, 분류체계 관리 데이터베이스 정의, 분류체계 등록/수정, 대·중·소 분류체계 C/R/U/D기능(CS 또는 웹상에서), 분류체계 인터페이스 관리 기능 등 다양한 기능을 실현할 필요가 있으며, 각 단계별 세부 설계는 다음과 같다.

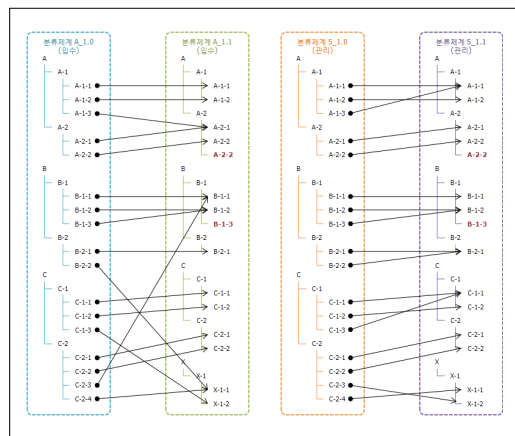


그림 4. 입수/관리 분류체계 이력관리 UI

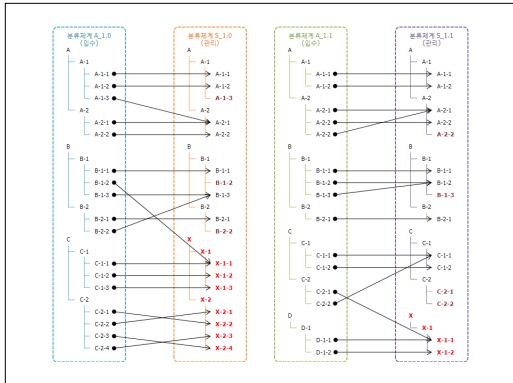


그림 5. 입수-관리 분류체계 매핑 관리 UI

III. 반자동(Semi-Automatic) 분류체계 관리

3.1 분류체계 파일의 자동 파싱

일반적으로 콘텐츠를 관리·서비스하는 기관에서는 여러 입수처로부터 다양한 형식의 콘텐츠를 입수하게 된다. 이 때 입수되는 Text, XML, MARC, MARC XML, Excel, Tagged Text, SGML, HTML, WORD, HWP 등과 같이 콘텐츠의 형식은 다양하며, 마찬가지로 분류체계의 입수 형식 또한 너무나 다양하고, 향후에는 더욱더 형식의 다양성이 증대될 것이다. 이에, 다양한 형식의 분류체계 입수 형태로부터 자동으로 분류체계 구조를 파싱하여 입수 분류체계 트리(tree)를 생성하여 주는 자동 파서에 대한 개발이 필요하고, 이는 기관에서 채택한 관리 표준 분류체계의 입수 시에도 반드시 필요할 것이다.

[그림 6]은 입수/관리 분류체계 파일의 파싱 프로세스의 흐름도이다. 다양한 분류체계 관리 파일로부터 <분류체계명>과 <분류체계코드>를 자동으로 추출한 후, 데이터베이스로 관리하게 되며, 입수 분류체계 이력 관리 및 관리 분류체계와의 매핑에 사용하게 된다. 이에 구현은 현재 개발 중인 DiCMS(Digital Contents Management System)의 입수 관리 프로세스와 동일함에 따라 이를 기반으로 구현이 가능할 것으로 판단되며, [그림 7]은 입수 관리 모듈인 eTracer을 보여주고 있다[8].

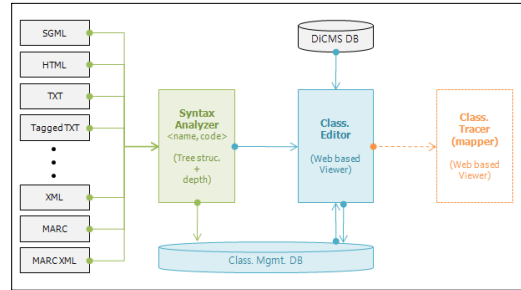


그림 6. 분류체계 파일의 자동 파싱 프로세스

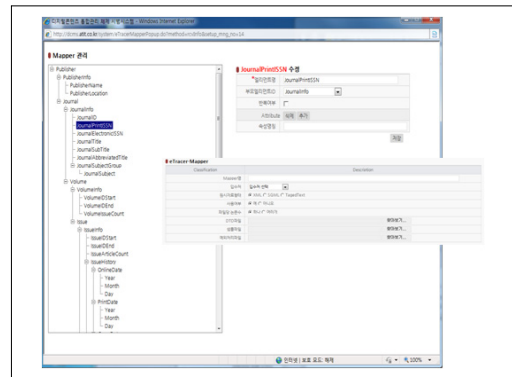


그림 7. eTracer의 입수파일 자동 파싱 화면

3.2 분류체계 간 매핑

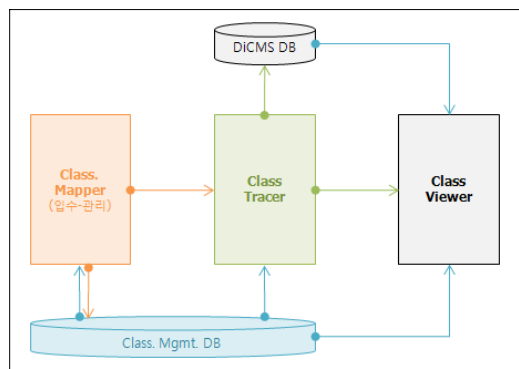


그림 8. 웹기반 분류체계 관리

파싱된 <분류체계명>, <분류체계코드>는 분류체계 DB를 통해 관리되며, 각 분류체계에 대한 name 및 code는 삽입, 수정, 삭제가 가능하다. 상기의 자동 파싱

프로세스에 의해 구성된 입수 분류체계와 환경설정에 의해 구성된 표준 관리 분류체계는 [그림 5]와 같이 웹 UI를 통해(Flex 기반 제작) display한다. 이렇게 구성된 인터페이스에 매핑할 분류체계를 간을 Flex를 기반으로 선으로 이음으로써, 두 분류체계를 연관시킬 수 있으며, 이 연관정보는 분류체계DB에 저장·관리되게 된다.

이러한 매핑 작업은 입수 분류체계-관리 분류체계 간, 관리 분류체계 및 버전 간에 수행되며, 이렇게 수행된 매핑 작업은 자동으로 입수 분류체계 버전 간, 관리 분류체계 간, 관리 분류체계-입수 분류체계 간 매핑 관계에 대해 view할 수 있게 하며, 이 또한 매핑 관계에 대한 삽입, 수정, 삭제가 가능하다. 이에 대한 구현 또한 상기의 [그림 7]과 같이 현재 개발 중인 DiCMS를 기반으로 구현이 가능할 것으로 판단되며, [그림 9]는 입수-관리간의 mapping 및 tracing 관리 기능을 보여주고 있으며, 이를 통해 분류체계 관리자는 새로운 관리 분류체계의 적용 및 새로운 입수 분류체계의 관리 분류체계로의 매핑 업무를 수행할 수 있을 것으로 판단된다. 이는 기존의 분류체계 관리/서비스에 대한 운영의 어려움 뿐만 아니라, 커다란 비용절감 및 최소한의 인력을 통한 운영이 가능해짐에 따라 실질적으로 분류체계 관리/서비스를 수행하지 못하고 있는 기관에서도 이것이 가능해 질 것으로 판단된다.

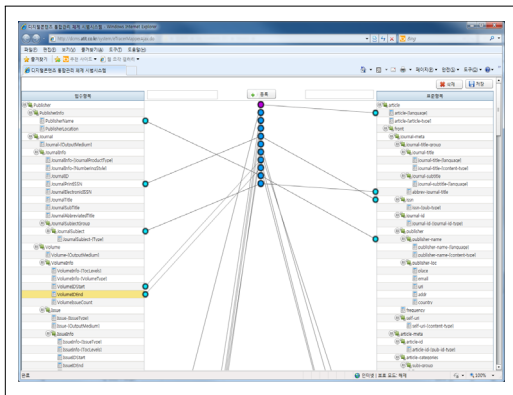


그림 9. 입수 콘텐츠 -관리 콘텐츠 간 반자동 관리 기능 화면

분류체계 관리DB는 각 입수·관리 분류체계별 <분류체계버전>, <분류체계버전 재정의일시>, <분류체계명>, <분류체계 Depth>, <분류체계코드> 및 <적용된 관리체계 버전>, <적용된 관리체계명>, <관리체계 Depth>, <관리체계코드> 등으로 구성되며, 각 기사 XML 상에는 분류체계 요소 중 관리요소를 포함하여 전체적으로 분류체계 요소를 기술하게 된다. 이때 관리 DB와 XML을 혼용하여(동시에) 운용하는 이유는 관리의 효율성을 극대화하고, 서비스 지연 및 액세스 부하를 최소화하기 위함이다.

IV. 결론 및 향후 연구

지금까지 새로운 분류체계 관리를 위한 여러 측면에서의 부분별 설계를 시도하였다. 우선, 입수/관리 분류체계의 입수에 있어서 기존에는 웹 상에서 view하거나, 문서 또는 데이터베이스 형태로 전달받아 이를 콘텐츠 관리/서비스 기관에서 문서 또는 데이터베이스 형태로 관리하고 있는 실정이다. 이는 인력/예산 낭비뿐만 아니라 관리적인 측면에서 많은 어려움을 초래하였다. 이에, 기존의 수동적인 입수절차를 탈피하여 파서(parser)에 의한 자동 파싱(parsing) 기능 및 이를 위한 수정/삽입/삭제 기능을 설계하였다. 이를 통해 향후에 입수·적용되는 분류체계에 대해 손쉽게 자동으로 입수·관리할 수 있는 기반을 제공할 수 있었다.

두 번째는 매핑에 있어, 입수된 분류체계와 각 콘텐츠 관리·서비스 기관이 보유한 관리 분류체계 간의 매핑에 있어서 수작업으로 상호간의 코드를 매핑하는 형태를 벗어나 Flex를 이용한 웹상에서의 자유로운 매핑(수정/삽입/삭제)을 제안하였다. 이를 통해 기존의 웹 기반 관리시스템에서 자유로이 매핑 작업을 수행하고, 이를 수정할 수 있을 것으로 예상된다.

끝으로 새로운 분류체계의 적용에 있어서 기존의 관리·서비스 체계에서는 새로이 입수되는 콘텐츠에 대해서만 새로운 분류체계가 적용가능하거나, 기존의 분류체계에 대해 일괄적인 갱신(update)을 통해서만 가능하다. 하지만, 제안된 방식을 통해서서는 입수-관리체계

와의 매핑 또는 관리체계 버전-업(version up) 시에, 새로운 콘텐츠에 자동으로 새로운 관리체계 코드가 적용되고, 기존의 콘텐츠에 대해서는 별도의 갱신이 없이 매핑 테이블 관리를 통해서 만으로도 관리 및 서비스가 가능해질 수 있다.

향후, 상기에 설계된 시스템에 대한 분류체계 반자동 관리시스템의 구현이 필요하며, 수동에 의한 분류체계 별 매핑과 같은 분류체계 이력관리 및 입수-관리 분류체계 간 매핑을 자동으로 할 수 있는 방안에 대한 연구 또한 요구된다. 이를 위해, 주제(주제어 포함) 자동 분류 시스템에 대한 연구 및 적용이 필요할 것으로 사료된다.

참 고 문 헌

[1] 홍성화, 서태설, "분류체계 일치를 통한 과학기술 정보 상호 교환 방법에 관한 기초 연구", 정보관리연구, 제35권, 제2호, pp.109-123, 2004.

[2] 오용선, 한정수, "디지털콘텐츠 기술 분류체계 및 분과기술 소개", 한국콘텐츠학회논문지, 제7권, 제4호, pp.36-71, 2009.

[3] 황다영, 김영인, 이병민, "기술융합 특성에 따른 새로운 분류체계의 제안", 기술혁신학회지, 제11권, 제4호, pp.592-612, 2008(12).

[4] 김비연 "DDC의 학제적 주제 분류에 관한 연구", 한국문헌정보학회지, 제45권, 제1호, pp.333-351, 2011.

[5] 백지원, "주제어 기반 분류에 관한 연구: 미국 공공도서관의 사례를 중심으로", 한국문헌정보학회지, 제44권, 제4호, pp.179-201, 2010.

[6] 백지원, "이용자 중심의 주제어 기반 분류를 위한 주제명 개발에 관한 연구: 지식조직체계 분석을 바탕으로", 정보관리학회지, 제28권, 제1호, pp.171-192, 2011.

[7] 진두석, 김광영, 이민호, 안설아, 이원구, 윤화목, 성원경, 최희운, "KISTI 지식총서: DCMS 정보관리 표준화 지침", 한국과학기술정보연구원, 지식

총서 ISBN: 978-89-6211-578-9, 2010.

[8] 조정길, "관계형 스키마로부터 중복성이 없는 XML 스키마로의 효율적인 변환 기법", 한국인터넷정보학회, 제11권, 제6호, pp.123-133, 2010.

[9] "산업기술표준분류/국가과학기술표준분류/미래유망기술(6T)분류/국가기술지도(NTRM)분류 / K I S T I 주 제 분 류 / C T 로 드 맵 , " <http://www.stepi.re.kr/http://www.itep.re.kr/http://www.mke.go.kr/http://www.smba.go.kr/http://kisti.re.kr/http://www.kocca.kr>

[10] Hjørland and Birger, "Information science: An interdisciplinary effort?," Journal of Documentation, Vol.64, No.1, pp.7-22, 2008.

[11] Szostak and Rick. "Classification, interdisciplinary and the study of science," Journal of Documentation, Vol.64, No.3, pp.319-332, 2008.

[12] Fewell, Rache, Macdonald, Logan, and Mainock, Loretta. "Bringing up baby bisac: wordthink classification. Converting a library collection to a word-based classification system," In Proceeding of 2009 CAL conference, 2009.

저 자 소 개

이 원 구 (Won-Goo Lee)

정희원



- 2000년 : 한남대학교 대학원 컴퓨터공학과 졸업(공학석사)
- 2005년 : 한남대학교 대학원 컴퓨터공학과 졸업(공학박사)
- 2005년 ~ 현재 : 한국과학기술정보연구원 선임연구원

<관심분야> : 데이터베이스, 지식관리, 과학데이터

신 성 호(Sung-Ho Shin)

정회원



- 2000년 : 경북대학교 경영학과 졸업(학사)
- 2002년 : 경북대학교 대학원 경영학과 졸업(석사)
- 2002년 ~ 현재 : 한국과학기술정보연구원 선임연구원

<관심분야> : 데이터통합, 데이터품질, IS평가

성 원 경(Won-Kyung Sung)

정회원



- 1989년 : 연세대학교 대학원 언어학과 졸업(석사)
- 1996년 : 프랑스 파리7대학교 전산언어학 졸업(박사)
- 2004년 ~ 현재 : 한국과학기술정보연구원 책임연구원

<관심분야> : 데이터베이스, 지식관리, 과학데이터

김 광 영(Kwang-Young Kim)

정회원



- 2001년: 부산대학교 전자계산학과(석사)
- 2011년 : 충남대학교 문헌정보학과(박사)
- 2001년 ~ 현재 : 한국과학기술정보연구원 선임연구원

<관심분야> : 정보검색, 개인화 검색, 아카이빙

이 민 호(Min-Ho Lee)

정회원



- 2000년 : 충남대학교 대학원 컴퓨터공학과 졸업(석사)
- 2006년 : 충남대학교 대학원 컴퓨터공학과(박사수료)
- 2001년 ~ 현재 : 한국과학기술정보연구원 선임연구원

<관심분야> : 시맨틱 웹, 빅데이터, 지식관리

정 도 현(Do-Heon Jeong)

정회원



- 2011년 : 연세대학교 문헌정보학과(박사수료)
- 2003년 ~ 현재 : 한국과학기술정보연구원 정보기술연구실 선임연구원

<관심분야> : 텍스트마이닝, 시맨틱웹, 정보검색

윤 화 목(Hwa-Mook Yoon)

정회원



- 1997년 : 공주대학교 대학원 전자계산학과 졸업(석사)
- 2008년 : 배재대학교 컴퓨터공학과 졸업(박사)
- 현재 : 한국과학기술정보연구원 정보기술연구실 책임연구원

<관심분야> : 데이터베이스, 정보검색, 온톨로지