

연구정보시스템의 메타데이터 보안 요소 연구

박묘경*, 이희조
고려대학교 컴퓨터정보통신대학원
e-mail:gatsby20@korea.ac.kr*, heejo@korea.ac.kr

Secure Integretion of Metadata in Research Information Systems

Myo-Kyung Park, Heejo Lee

*Graduate School of Computer and Information Technology,
Korea University

요 약

우리나라는 정보기술에 대한 투자 확대와 과학기술성과확대를 위한 노력을 지속적으로 해오고 있다. 하지만 아직도 수요자중심의 중장기적이고 기초·원천적인 국가연구개발사업의 실질적 설계 및 실행이 미흡한 실정 속에서 각각 개별 기관의 연구정보관리로 인해 원활한 연구관리가 이루어지는데 어려움이 크기에 중앙관리 및 정보공유를 위한 메타데이터 센터의 필요성이 가중되고 있으나 그에 따른 보안의 중요성도 커지고 있다. 따라서 본 논문에서는 정보통신연구진흥원, 한국과학기술평가원, 한국산업기술평가원, 한국과학재단 등 주요 정부기관별 종합연구관리시스템의 메타데이터들을 비교·분석을 통해 그 관리에 필요한 보안 요소들에 대해 연구하고 향후 종합연구관리 시스템 구축은 물론 국가 메타데이터 표준안 제정시에도 고려되어야 할 보안 방안에 대해 영역별로 제시하였다.

1. 서론

지식기반경제에서 국가의 경쟁력은 지식을 창출하고 창출된 지식을 확산시켜 가치창출에 활용하는 능력에 달려있다고 볼 수 있다. 그러나 연구 관리의 투명성에 대해서는 그 동안 많은 비판들이 끊임없이 제기되어 왔으나 구체적인 해결안들이 실행되지 못하였다. 국가연구개발사업이 관리하고 있는 개별 과제관리기관들이 협의회 등을 구성하여 부처 간, 기관 간 업무협력을 꾀하고 있지만, 각 부처 간의 이해관계와 정보보안, 시스템의 표준화 등에 있어 이견으로 인한 현재까지 연구개발 종합정보의 체계적인 유통체제가 마련되지 못하고 있다[1].

본 논문은 지속적인 연구개발예산의 증가에도 불구하고 국가차원의 연구개발 종합조정이나 체계적 전략연구기획에 한계만이 아니라 이러한 국가연구개발사업의 연구결과(성과)의 활용에 있어서 체계적으로 활용되고 있지 못하다는 기존의 연구[1]에 이어서 연구관리시스템이 각기 유기적으로 활용되기 위해 가장 필요한 메타 데이터의 보안을 위해 고려되

어야 할 요건에 대해 연구하고자 한다.

따라서 2장에서는 국가연구개발사업 관리시스템이 각 부처별 기관별로 중복적으로 운영되고 있으며 부분적으로 활용되고 있는 연구 관리시스템과 연계 활용되지 못하고 있는 기존의 실정에 대해 살펴보고, 3장에서는 메타데이터의 정의와 그 필요성에 대해 알아보고, 4장에서는 정부 주요 연구관리기관의 연구정보시스템을 비교·분석하여 이를 근거로, 표준·공통 핵심정보인 정보보안 요소들과 관리 대책에 대해 살펴본다. 마지막으로 5장에서는 결론 및 향후 과제를 제시한다.

2. 우리나라 정부기관의 관리 시스템 실정

국가연구개발사업에 관련한 다양한 법령과 부처별 규정이 존재하고 있음에도, 이들은 대체적으로 공공기술의 관리 가운데 주로 행정적 관리에 관한 사항에만 치중하며, 공공기술의 연구개발에서도 그에 소요되는 국가예산의 통제에만 중점을 두고 있다. 외관상 연구개발성과의 활용을 강조하면서도 이

를 위한 필수적 전제조건에 대한 이해 내지 제도적 장치가 미흡하다[2],[4].

국가연구개발사업 종합관리시스템은 이미 시스템을 확보하고 있는 기존의 대형 과제관리기관으로 한국과학기술기획평가원, 한국산업기술평가원, 한국과학기술진흥재단, 정보통신연구진흥원, 한국과학재단, 농촌진흥청 등의 기관들은 새로운 시스템을 따로 설치하기보다는 에이전트 시스템과 인터페이스만을 설치함으로써 시스템 간 연동이 가능토록 하였다. 연계기관간 콘텐츠 표준을 위해서 기존에 개발되어 있는 정보의 메타데이터의 등록을 유도하고, 신규로 개발되는 정보에 대해서는 사전에 메타데이터를 표준화함으로써 효과적인 정보유통이 이루어지도록 하는 새로운 개념의 표준화 실행 체계가 구축되었다. 즉, 한국과학기술정보연구원의 MDR과 같은 국가차원의 메타데이터센터(MDC:Metadata Center)를 설립 및 이를 통하여 기존의 부처별 특성에 따라 사용하고 있던 표준양식 및 분류체계를 그대로 사용하더라도 효과적인 정보공유가 가능토록 도모하는 초기단계에 있다[1].

3. 메타 데이터의 정의와 필요성

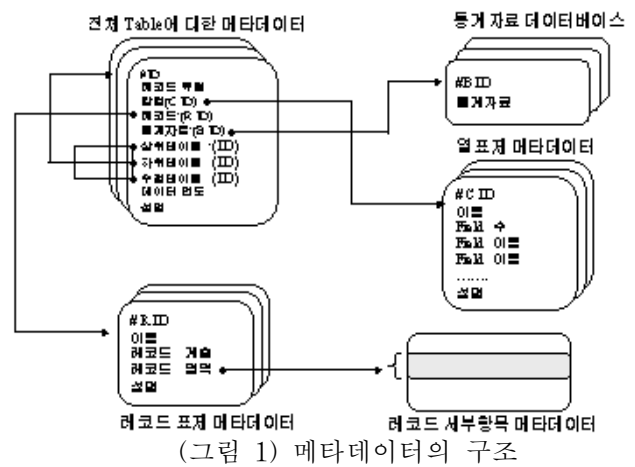
<표 1> 공공 정보의 공공 이용을 위한 정보자원 관리 방안

우선순위	관리계층	관리 지침
1	제도	부처의 정보소유권 관련 정책마련, 유통문서표준정책 마련, 정보접근절차에 관한 규정 마련
	관리	정보공유의 중요성 환기 및 인센티브 마련, 통합 네트워크의 구축과 연동 계획 마련, 정보공유를 위한 가치분류체계 마련
	기술	고속LAN확보, 통신표준확보, 보안시스템확보
2	제도	통신보안관련 규정의 마련, DB표준과 메타데이터 구축, 등급체계를 통한 정보의 가치분류
	관리	정보분류체계 개발과 강제, 메타데이터 작성 및 관리, 이형 DB 관리체계 확립, 데이터합병 방안 마련
	기술	DBMS확보, 보안시스템의 확보, 광과일저장매체 확보,
3	제도	정보시스템 사용환경 개선
	관리	자료관리절차의 확립, 기관수준의 자료통합성 유지,
	기술	Viewer의 확보
4	제도	정보수집허가절차의 마련, 정보의 정확성 확보, 디지털 형식 정보의 수용 및 접수
	관리	자료의 표준화 관리, 공무원의 정보활용 교육
	기술	저장매체확보

위의 표를 살펴보면 고속 LAN확보 등의 네트워크 연결인 1순위를 제외하고 살펴보면 무엇보다 통신보안관련 규정의 마련, DB표준과 메타데이터 구축, 정보분류체계 개발과 강제, 메타데이터 작성 및 관리, 보안이 가장 우선순위의 것을 알 수 있다.

여기서 메타데이터는 정보 시스템, 데이터베이스, 파일 등에 포함된 데이터의 내용, 구조, 형식 및 의미를 시스템적으로 정의한 정보이다. 메타데이터는 일반적으로 ‘데이터에 대한 데이터(data about data)’라고 하며, 대상이 되는 정보 자원의 속성과 특성 및 다른 자원과의 관계를 기술하여, 이용자의 관점에서는 검색을 돕고 관련기관에서는 정보원 제어와 관리를 돕는 역할을 한다.

즉, 메타데이터(metadata)는 실제로 저장하고자 하는 데이터(예를 들면 비디오, 오디오, 텍스트 등)는 아니지만, 이 데이터와 직접적으로 혹은 간접적으로 연관된 정보를 제공하는 데이터를 나타내는 말이다.



메타데이터는 데이터에 대한 공동적인 이해를 가능케 함으로써, 시간과 공간을 초월하여 데이터로의 접근, 사용, 공유를 가능하게 한다. ISO/IEC, ANSI를 비롯한 여러 표준화 위원회에서 자원의 공유를 위해 메타데이터 표준안을 마련하고 있다. 현재 통일된 하나의 메타데이터 표준은 존재하지 않는다. 이는 다양한 환경과 요구에 맞는 적절한 메타데이터 포맷이 요구되며, 이에 따라 각 분야에서 사용되는 메타데이터 구조도 특정 분야의 요구에 맞게 매우 상세하고 전문적으로 개발되고 있기 때문이다. 오늘날 수십 여종의 메타데이터가 사용되고 있으며, 대표적인 것으로 도서관에서 오래전부터 목록용으로 사용되었던 MARC(Machine Readable Cataloging)가 있으며, 인터넷상의 전자자원을 기술하기 위한 더블린 코어(Dublin Core) 메타데이터, 온라인 서점과 관련된 ONIX(ONline Information eXchange), 전자도서관을 위한 MODS(Metadata Object Description Schema) 등이 있다. 이 밖에도 다양한 메타데이터들이 각각의 목적과 용도에 따라 개발되었고, 또

한 개발 중에 있다. 메타데이터를 구체적으로 활용한 예로는, 미국의 각 기관에서 생성된 통계 자료의 공유를 위한 미 통계국(Bureau of the Census)의 통계정보시스템(SIS: Statistical Information System) 등이 있다[5],[8].

이렇듯 인력 및 기관 정보시스템의 데이터 교환 및 상호호환성을 위한 공통 및 핵심 포맷인 메타데이터가 필요한 것이다.

4. 우리나라 주요 연구관리기관의 메타데이터 관리 현황 및 보안 요소

<표 2> 정보통신연구진흥원의 메타 데이터 관리 현황

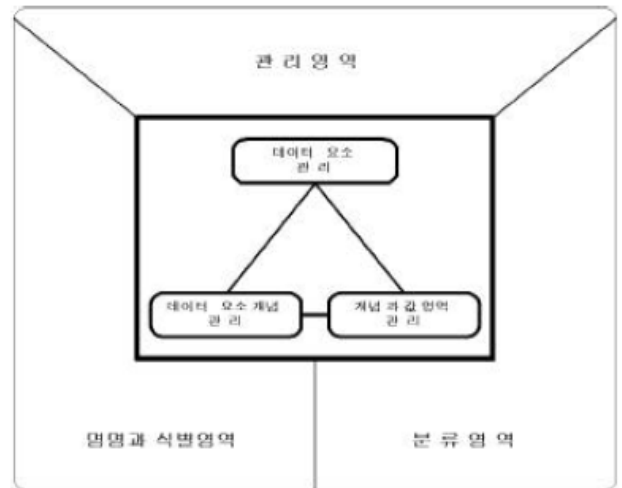
항목	적합성	테이블 이름	필드이름	데이터 형태	최대 길이	입력 데이터 예시	특이 사항
과제명	O	과제마스터	연구과제명	char	500	한국 SW벤처..	
수행기간 -연속과제 시작/종료	O	과제마스터	연구기간개시/종료	date	8	19970501	
-당해연도 시작/종료	O		당해연구기간개시/종료	date	8	19990501	
	O		연구기간개시/종료	date	8	20000430	
수행기관	△	과제주연구기관	사업자등록번호	char	10	3128210822	
연구책임자 -성명	O	과제연구인원	책임자_이름	char	40	홍길동	
-주민번호	O		책임자_주민번호	char	14	123456-78	
-E-mail	O		책임자_메일	char	100	hong@core	
-전화번호	O		책임자_전화번호	char	50	a.ac.kr 02-123-1234	
요약정보 -연구목표	O	과제연구내용	연구목표	char	4000	본 연구에서	
-연구내용	O		연구내용	char	4000	제시한 SW..	
-연구성과	△						
키워드 -한글	O	과제키워드	키워드	char	300	벤처캐피탈, Software...	
-영문	O						
기술분류코드	O	과제마스터	기술분류코드	char	10		한/영 혼용 사용
분석코드 -경제사회목적	X						각 과제담당자가 수작업으로 처리
-연구개발단계	X						
-참여기어분류	X						
-기술수명주기	X						
-연구수행주체	X						

위의 표처럼 정보통신부의 대표적인 연구관리기관인 정보통신연구진흥원의 메타데이터의 경우 자체 운영중인 과제관리시스템에서 대부분 전산처리가 가능하다. 하지만 연구성과 등의 경우 연구계획서의 연구내용에서 추출하여야 하며, 분석코드의 경우 DB상에서 관리되어 왔으나, 현재까지의 연구개발사업 조사·분석·평가 자료제출의 경우 과제담당자들이

수작업을 통하여 처리하고 있다.

한편, 한국과학기술평가원의 경우, 복잡한 과제 코드체계 등을 시스템 개발과 함께 전반적인 재설계를 수행중이다. 한국산업기술평가원은 전체적으로 과제관리 정보시스템 도입과 함께 DB schema를 재편성 중이다. 특히 새로 설계한 과제고유번호를 도입하여 기존의 분류코드 형식에서 일련번호 형식으로 변경할 예정이다. 또한 분석코드 부분을 웹을 통해 입력하고 있다. 한국과학재단은 역시 현재는 과제정보를 통합된 단일 테이블로 관리하고 있으나 시스템 업그레이드와 함께 DB schema를 재편성 중이다[1],[4].

이렇듯 상호운용성이라는 메타데이터의 특성을 뒷받침하기 위해 가장 필수적인 보안에서 고려되어야 할 요소들에 대해 크게 ①관리 영역, ②명명 및 식별 영역, ③분류 영역, ④데이터 요소 개념 관리 영역, ⑤개념과 값 영역 관리 영역, ⑥데이터 요소 관리 영역으로 세분화하여 제시하고자 한다[6],[7].



(그림 2) 메타데이터의 보안 영역

①관리 영역

관리영역은 하나의 데이터베이스와 레지스트리를 구성하는 구성요소들의 관리에 대한 특면을 지원한다. 즉, 보안요소에 대한 책임을 지는 조직과 관련이 있다.

- 관리된 구성요소(administered component)

각 기관 내지는 기업 간에 공유를 위하여 정의와 명세를 요로 하는 데이터 요소 구성요소를 의미한다.

- 등록 기관(registration authority)

메타데이터의 등록에 대한 권한을 가지는 조직을 의미한다. 이러한 등록 기관 각각에 대하여 관리된 구성 요소는 반드시 다른 관리된 구성요소들과 구별하

기 위한 식별자와 버전을 가진다.

②명명과 식별 영역

이 영역은 해석방향, 지역성, 기술적 요소 등에 의하여 발생 할 수 있는 등록된 메타데이터의 여러 이름을 관리하여 식별한다.

③분류 영역

등록된 DB에 대하여 분류법(taxonomy),네트워크, 존재론(ontology)과 같은 하나 이상의 분류방법이 사용된다. 분류된 구성요소는 객체 부류, 특성 부류, 표현 부류, 데이터 요소 개념 또는 개념적 영역을 위한 일반화이고, 하나 이상의 분류 방법 내에 적용된다.

④데이터 요소 개념 관리 영역

이 영역의 목적은 데이터 요소의 개발에 관련된 개념에 대한 정보를 유지 보수하는 것이다. 즉, 실시간으로 갱신되는 각 메타데이터의 업데이트와 함께 보안을 염두에 둘 때 유의해야 부분 중에 하나이다. 이 영역의 주된 구성요소는 데이터 요소 개념으로 조합되는 개념(객체부류)과 특성이다.

- 객체 부류(object class): 실세계의 생각, 추상 또는 사물들의 집합으로 명백한 범위와 의미, 그리고 속성, 행위들이 같은 법칙에 의하여 정의된다.

- 특성(property): 한 객체 클래스 내의 구성요소가지는 일반적인 특성을 가리킨다.

- 데이터 요소 개념(data element concept): 객체 부류와 특성의 조합은 새로운 개념을 만들기 때문에 데이터 요소 개념은 객체 부류와 특성의 정의와는 다른 새로운 정의를 갖는다. 데이터 요소 개념 역시 분류된 구성요소로서 분류 방법 내에 존재할 수 있다.

⑤개념과 값 영역 관리 영역

메타 데이터에서 이 영역은 데이터 요소 개념을 지원하는 개념적 영역과 데이터 요소를 지원하는 값 영역의 관리를 맡고 있고 물리적, 논리적 코드 집합으로 표현할 수 있다.

- 개념적 영역(conceptual domain)

- 값 영역(value domain)

⑥데이터 요소 관리 영역

이 영역은 메타데이터 요소를 관리하는 목적에 사용된다. 데이터 요소는 실질적 또는 추상적인 사물에 대한 정보를 정형화된 표현을 제공한다. 다른 요소들보다도 데이터 요소는 데이터 요소 개념의 재사용 가능하고 공유 가능한 표현이 될 수 있다는 점에서 메타 데이터의 가장 큰 장점이자 그만큼 약점이 될

수도 있다는 점에서 종합연구관리시스템 구축 시 가장 염두에 두어야 할 것이다.

5. 결론 및 향후 과제

이상 앞서 살펴본 바를 종합해보면 종합연구관리시스템간의 상호연동을 가능케 하는 키워드인 메타데이터의 필요성 및 보안이 무엇보다도 시급하다.

본 논문은 몇 개의 정부 주요 연구관리기관의 연구관리시스템을 비교 및 분석하여 현재의 메타데이터 보안 요소의 조건 및 필요성을 파악함으로써 그 절차와 ①관리 영역, ②명명과 식별 영역, ③분류 영역, ④데이터 요소 개념 관리 영역, ⑤개념과 값 영역 관리 영역, ⑥데이터 요소 관리 영역 등의 구성요소와 그 하위개념들을 제시하고자 하였다.

또한 메타데이터 보안 요소를 제시함으로써 아직 미 제정된 국내 메타데이터 표준안 제정과정에서의 효율성과 온라인 서점이나 도서관등 사용기관의 특성에 따라 설계된 특화된 메타데이터들처럼 연구정보시스템이라는 고유의 특징을 지지하고 보안 유지할 수 있는 메타데이터의 구현 및 보안 관리에 이바지하기를 기대한다.

참고문헌

[1] “연구 보고서-연구개발지식의 종합관리시스템 구축 방안 연구” pp. 198-213 과학기술부 2003

[2] “국가정보화백서-「행정기관의 정보화책임관 지정·운영에 관한 지침안」” 한국전산원 1997-2000

[3] 최홍석 이정준 구상희 “정부 정보자원관리모델에 관한 연구” 한국전산원 2000

[4] 김민형 “PBS 중심의 출연예산관리시스템 현형과 과제” 과학기술정책연구소 1999

[5] 김태중 “Web 데이터베이스의 효율적인 검색을 위한 분류체계 및 초록 작성 지침” 한국데이터베이스진흥센터 1998

[6] “표준화 연구 보고서-메타데이터의 명세 및 메타모델의 표준화” pp. 161-178 한국데이터베이스진흥센터 2003

[7] 한선화 손용욱 황보열 권용수 김태희 “웹기반 과학기술 통계, 지표의 자동 생성 시스템의 설계 및 구현” 정책분석평가학회보 제 8권 제 2호 1998

[8] 한선화 “과학기술정보시스템을 위한 메타데이터의 활용” 한국과학기술정보연구원 2000