

공공부문에서의 데이터 이용 및 관리 방법

†김진형^o †김영갑 †신성욱 ††변현진 ††신신애 ††이헌중 †정동원
†고려대학교 컴퓨터학과 ††한국전산원
{^okoolmania, stnoble, withimp, back}@software.korea.ac.kr†
{hjbyun, sashin, hjlee}@nca.or.kr††

A Data Use and Management Policy in the Public Sector

Jinhyung Kim^o Young-Gab Kim SungOok Shin Hyunjin Byun Shinae Shin Hunjung Lee Dongwon Jeong
Dept. of Computer Science & Engineering, Korea University
Nation Computerization Agency

요 약

국내 공공부문에서의 통합적인 다양한 서비스 제공을 위해서 여러 가지 노력과 연구가 진행되고 있다. 통합 서비스를 제공하기 위해서는 사용자에게 제공되는 정보들 간의 일관성 있는 구축, 관리 및 운용을 위한 DRM에 대한 연구가 필수적으로 요구된다. 이미 미국, 영국 등 선진국에서는 DRM 개발을 위한 연구가 계속적으로 진행되어 왔으며, 국내에서도 이러한 중요성을 인식하여 현재 한국형 DRM 개발을 위한 사전 연구를 진행하고 있다. 이 논문에서는 DRM 개발을 위한 주요 내용 중 응용, 사용자, 기관 등에 따라 달라지는 정보 접근을 체계적으로 관리할 수 있는 이용 관리 정책에 초점을 둔다. 다양한 측면에서 정보에 대한 접근성을 고려하고 이러한 규칙을 기술하고 관리하는 방법에 대하여 기술한다.

1. 서 론

DRM(Data Reference Model)은 데이터의 구축 및 운영을 위해 제공되는 지침이다.[1,2,3]. DRM은 유사한 정보를 다루거나 또는 서비스의 향상을 위해 데이터의 상호교환 및 공유가 요구되는 분야에서 개발해야 하는 아키텍처의 하부 구조에 해당한다. 이러한 중요성 때문에 이미 여러 선진국에서는 이에 대한 연구가 진행되고 있다[1,2,3].

따라서 국내에서도 DRM 개발을 위한 많은 선행 연구들을 진행하고 있으며, 특히 대국민 서비스를 위한 공공부문에서 활발하게 진행되고 있다[4]. 이러한 선행 연구 결과를 바탕으로 내년부터 본격적인 DRM 개발 사업이 진행될 것으로 보인다.

DRM을 개발하기 위해서는 메타데이터, 온톨로지, 데이터 사전, 데이터 보안, 이들에 대한 전체적인 메타 모델, 프로세스 등 다양한 기술과 구성요소를 필요로 한다.

국내외 관련 연구를 통해 이러한 다양한 기술과 구성 요소에 대한 정의는 국내 환경에 적합한 한국형 DRM 개발을 위해 필수적이며 사전 연구를 필요로 한다. 이를 위해 한국전산원에서는 고려대학교와 함께 한국형 DRM 개발을 위해 사전 연구를 진행하고 있다[4].

이 논문의 주제는 이 연구의 전체 주제 중 일부로서 다양한 관점에 따라 데이터의 이용에 대한 접근을 어떻게 제어하고 유지 관리할 것인가에 초점을 두고 있다. 이는 이용 측면에서의 데이터 보안과 유사하며 데이터의 가시성, 데이터의 접근 제어, 또는 데이터 이용 제어라는 용어와도 그 맥을 같이 한다. 이 논문에서는 이들 용어

들을 대표하여 데이터 이용 제어 혹은 데이터 접근 제어라고 표현한다.

이 연구의 결과는 전체 DRM에서의 데이터 이용에 대한 규칙을 정의하는데 이용될 수 있으며, 다양한 관점에 따라 동일한 데이터에 접근 제어를 가능하게 한다.

이 논문의 구성은 다음과 같다. 제 2장에서는 간략하게 공공부문 각 기관들 간의 데이터의 상호운용성 제공을 위한 전체 프레임워크와 DRM의 역할, 그리고 DRM의 주요 구성요소에 대하여 기술한다. 제 3장에서는 데이터의 접근 제어를 위한 이용 정책에 대하여 기술하고 또한 이용 정책의 표현과 활용 방법에 대하여 기술한다. 마지막으로 제 4장에서는 결론 및 향후 추가 연구 내용에 대하여 기술한다.

2. 데이터 상호운용을 위한 프레임워크와 DRM

2.1 전체 프레임워크

다양한 기관들간 또는 부처들간의 데이터 상호운용성을 통해 서비스의 다양화 및 질적 향상을 이루기 위해서는 크게 데이터 계층, 미들웨어 계층, 응용 계층 등 세 가지의 계층의 개발이 요구된다. 그림 1은 데이터 교환 및 공유를 통해 상호운용성을 제공하기 위한 전체적인 프레임워크를 추상적으로 보여준다.

그림 1에서, 데이터 계층(Data Layer)은 데이터베이스를 설계 및 구축하기 위한 규칙들과 각각 구축된 데이터베이스내의 데이터들간의 공유 및 교환을 위한 이질적인 문제를 해결하기 위한 규칙들의 집합이다. DRM은 이 계층에 해당한다.

미들웨어 계층(Middleware Layer)은 데이터베이스의 이형질성, 개발된 프로그램 간의 이질성 등의 문제를 해결하여 상호운용을 원하게 하는 계층이다. 플랫폼이 서

† 이 연구는 한국전산원 지원으로 수행되었음.

로 다르거나 개발 언어가 서로 다른 문제 또는 데이터베이스 관리 시스템이 서로 다른 문제 등을 해결하는 계층이다.

마지막으로, 응용 계층(Application Layer)은 해당 분야 즉 부처 및 기관에서 제공하고자 하는 서비스에 맞게 데이터를 가공하고 사용자에게 보여 줄 수 있는 인터페이스 환경을 생성하는 계층이다.

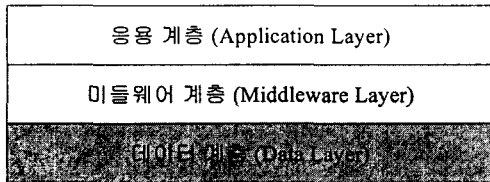


그림 1 상호운용을 위한 전체적인 프레임워크

2.2 DRM 구조 및 주요 구성요소

DRM은 데이터의 상호운용성을 극대화하기 위해 필요한 가장 중요한 구성요소로서 일반적으로 프로세스와 프로세스에 따라 생성해야 하는 모델들에 대한 규칙의 집합으로 구성된다.

프로세스는 크게 DRM을 정의하는 단계와 이를 활용하는 단계로 구분된다. 또한 DRM을 기능적 관점에서 분류하면 데이터(스키마 및 필드 정의)를 정의하는 기능, 메타데이터 정의하는 기능, 데이터의 이용 정책 관리 기능, 공유 및 교환을 위한 사상(Matching & Mapping) 기능, 분산 관리 기능 등이 있다.

DRM 구조는 컴포넌트 관점에 따라 분류할 수 있으며, 그림 2는 컴포넌트 관점에 따른 DRM의 일반적인 구조를 보여준다.

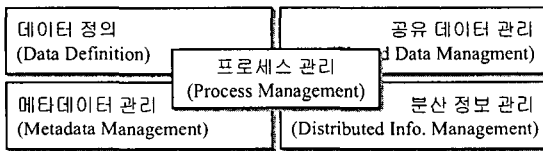


그림 2 DRM의 주요 컴포넌트

각각의 컴포넌트를 보다 세부적으로 정의하면 그림 3과 같다. 그림 3에서, 메타데이터 관리 컴포넌트는 국제표준인 ISO/IEC 11179에 따라 데이터 요소 즉, 필드를 정의하기 위해 요구되는 일련의 세부 구성요소와 정의된 데이터 요소의 집합인 MDR(Metadata Registry)로 구성되어 있다[5,6,7].

데이터 정의의 컴포넌트는 데이터베이스를 정의하기 위한 스키마와 필드를 정의하는 구성요소(Schema&Field), 데이터를 정의하기 위한 지침으로 이용되는 메타모델(Metamodel), 표준 필드인 데이터 사전 검색을 위한 구성요소(DD Search) 등으로 구성되어 있다. 특히 정의된 데이터를 이용하기 위한 정책을 기술하는 구성요소(Use

Policy Definition)를 지닌다. 이 구성요소는 공유 데이터 관리 컴포넌트의 접근 제어(Access Control) 구성요소와 함께 이 논문의 주제에 해당한다.

공유 데이터 관리 컴포넌트와 분산 관리 컴포넌트 또한 다수의 세부 구성요소로 구성된다. 분산 관리 컴포넌트가 지니는 세부 구성요소는 분산되어 있는 데이터에 대한 접근 및 대상 데이터베이스 구조에 맞게 정보를 생성해 주는 역할을 한다.

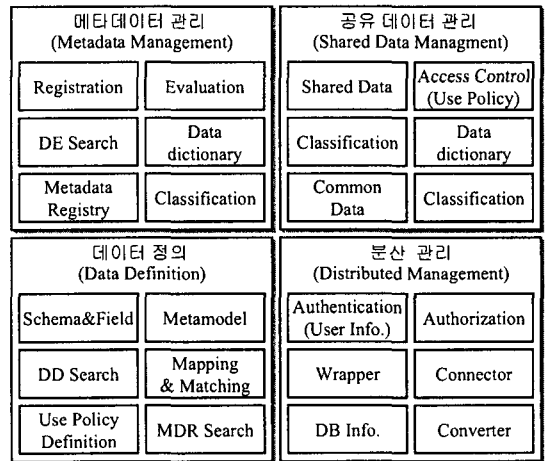


그림 3 주요 컴포넌트의 세부 구성요소

3. 데이터 이용 정책 정의의 관리 방법

3.1 관점에 따른 접근 제어

데이터에 대한 이용 정책은 일반적으로 그룹별로 접근을 허용하는 방법이다. 즉 사용자에 따라 특정 데이터에 대한 접근 및 이용이 허용 또는 거부된다. 그러나 다양한 공공부문의 경우 다양한 기관들이 서로 정보를 공유해야 하기 때문에 단순히 사용자에 따른 접근 제어는 적합하지 않다. 따라서 보다 다양한 관점에 따른 데이터 이용 정책을 고려해야 한다.

이 논문에서는 사용자를 포함하여 보안, 공유 등의 관점을 고려하여 동일한 데이터에 대한 다양한 접근 제어가 가능하도록 이용 정책을 정의한다. 미국의 SHADE[3]의 경우, shared, unique, universal 등으로 분류하고 있다. 그러나 국내 NEIS 문제를 보듯, 개인적인 정보 유출이라는 측면을 고려하지 않을 수 없다.

이 논문에서 이러한 단점을 해결하기 위해 데이터의 접근 제어를 위한 기관간 공유 관점, 개인 정보 보호(Privacy)를 위한 보안 관점, 각 기관별 일반적인 접근 제어 정의 관점, 그리고 제공하는 서비스에 따라 달리 정의할 수 있는 응용 관점으로 분류한다.

그림 4는 이 논문에서 제안하는 데이터 이용 정책, 즉 데이터 접근 제어를 위해 고려하는 관점간의 관계를 개념적으로 표현한 것이다.

그림에서 알 수 있듯이, 이용 정책 즉 데이터에 대한 접근 제어 규칙을 정의하기 위해서는 크게 공유, 보안,

조직, 응용의 네 가지 관점을 고려하게 된다. 따라서 한 데이터에 대한 접근 제어 규칙을 정의하기 위해서는 네 가지 관점을 수용하도록 정의해야 한다.

공유 관점은 기관간의 공유 레벨을 정의하는 것으로 *upperD, upperA, sibling, lowerD, lowerA* 등으로 분류한다. 각각 상위 직속 기관, 관련된 모든 상위 기관, 동일한 레벨의 관련 기관, 직속 하위 기관, 모든 관련 하위 기관을 의미한다.

보안 관점은 개인 정보의 보호를 위한 것으로 이미 NEIS에서 이 문제가 대두되었다. 보안 관점은 크게 *private, protected, public* 관점으로 분류된다. *private*은 개인 정보로서 개인만이 볼 수 있으며, 암호화되어 시스템에 저장되도록 한다. *protected*는 해당 관리 기관까지만 볼 수 있음을 의미한다. 마지막으로 *public*은 다른 일반 사용자들에게도 공개되는 정보이다.

조직 관점은 일반적인 데이터 접근 분류 방법으로서 *system, member, end-user* 등으로 분류하며 각각 시스템 관리자, 조직원, 일반 사용자를 의미한다.

마지막으로, 응용 관점은 서비스에 대한 이용을 제어하는 것으로 각각 서비스를 통해 제공되는 데이터에 대한 접근 제어를 가능하게 한다.

각 관점에 따른 데이터 접근 방법들은 상호배제, 나열, 의존 등의 관계를 지닌다. 이 논문에서는 이 부분에 대한 자세한 내용은 생략하기로 한다.

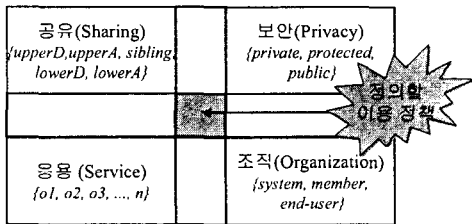


그림 4 이용 정책을 접근 제어 관점 분류

3.2 이용 정책 표현 및 활용

앞서 기술한 이용 정책들을 규칙을 이용하여 표현한다. 표현된 규칙들은 보다 빠른 처리를 위해 재구성되며 이를 이용하여 각 데이터에 대한 이용 즉 접근 제어를 수행하게 된다.

예를 들어, 데이터 집합 D 가 다수개의 데이터 집합 $d_1, d_2, d_3, \dots, d_n$ 으로 이루어져 있다고 하자. 이 때, d_1 은 모든 상위 기관들에 의해서만 공유되어야 하고 이 데이터 집합을 지니고 있는 조직의 시스템 관리자만이 정보를 접근할 수 있다고 하자. 이를 규칙으로 표현하면 다음과 같다.

$$(D.d_2) \rightarrow (upperA, system)$$

위에 표현된 규칙에 따라 d_2 를 접근하고자 할 때 액션 부분에 있는 접근 규칙에 따라 허용하거나 거부하게 된다.

4. 결론 및 향후 연구

이 논문에서는 DRM의 주요 구성요소 중 데이터에 대한 접근 정책을 분류하여 정의하고 표현하는 방법에 대하여 제안하였다. 제안된 데이터에 대한 접근 제어 분류는 접근 분류를 단순히 사용자 분류에 의존하여 제어하는 방법의 문제점을 보완할 수 있다. 예를 들어, 사용자 개인 정보 보호와 같은 문제를 해결할 수 있다.

향후에는 실제 다양한 도메인에 적용하고 해당 전문가들과의 협동 연구를 수행함으로써 분류한 관점에 대한 검증과 보완 작업이 이루어져야 한다. 마지막으로 보다 효율적으로 규칙을 적용하여 활용할 수 있는 알고리즘 개발 또한 요구된다.

참고문헌

- [1] Industry Advisory Council Enterprise Architecture Shared Interest Group, "Information and Data Reference Model : Standards Based Architecture to Support Federated Data Management," Version 2.0, 7 January 2003
- [2] Office of the e-Envoy, "e-Government Interoperability Framework," Version 6.0, 30 April 2004
- [3] 손태정 외, "국방 데이터 공유환경(SHADE) 기술구조 연구," 한국국방연구원 연구보고서, 정01-1728, 2001.07
- [4] 한국전산원, "공공부문 데이터 상호운용성 확보를 위한 기술방안 및 로드맵 수립," 2004
- [5] 김진관 외, "메타데이터 이질성 해결을 위한 MDR 기반의 메시지 변환 시스템," 한국정보과학회 논문지D, Vol. 31, No. 3, pp. 0232~0242, 2004.06
- [6] 신동길 외, "계층적 메타데이터 레지스트리 기반의 점진적 데이터 통합," 한국정보과학회지, 제30권, 제1호, 2003.04
- [7] ISO/IEC JTC1/SC 32, ISO/IEC 11179: Specification and Standardization of Data Elements, ISO/IEC JTC1, Part1~6