

이용자 참여형 시맨틱 디지털도서관 아키텍처 설계

Designing an Architecture for Social Semantic Digital Libraries(SSDL)

오 삼 균(Sam-Gyun Oh)*

원 선 민(Sun-Min Won)**

초 록

정보기술 환경의 변화는 디지털 도서관에 혁신적인 변화를 촉구하고 있다. 본 연구는 온톨로지 시스템으로 구성되어 이용자 참여·협업을 적극 지원하는 디지털 도서관을 Social Semantic Digital Library(이하, SSDL)로 정의하고, 이러한 디지털 도서관에 요구되는 핵심사항과 아키텍처 모델을 제시하였다. SSDL은 메타데이터와 온톨로지를 기반으로 의미기반 정보처리를 지원함과 동시에 이용자의 참여를 통해 정보와 지식을 재생산하는 혁신적인 체제이다. 본 연구에서는 5개의 수평 계층과 2개의 수직요소로 구성되는 새로운 SSDL 모델을 제시하였다.

ABSTRACT

The change in information technology demands drastic change in digital library service. This study defines what a social semantic digital library should consist of and proposes a new architecture that incorporates core functions needed in designing a SSDL. The SSDL supports semantic information processing based on metadata and ontology and is an innovation system that allows SSDL users to participate in generating new knowledge by interacting with existing metadata and ontology structures. This study designed a SSDL model that consists of five horizontal and two vertical structures.

키워드: 시맨틱 라이브러리, 디지털 라이브러리, 웹2.0, 온톨로지, 메타데이터
Semantic Library, Digital Library, Web 2.0, Ontology, Metadata

* 성균관대학교 문헌정보학과 교수(samoh@skku.edu)

** 성균관대학교 문헌정보학과 대학원 과정(hidemix@naver.com)

논문접수일자 : 2007년 11월 21일 논문심사일자 : 2007년 11월 28일 게재확정일자 : 2007년 12월 8일

1. 연구의 배경 및 목적

물리적 도서관의 한계를 극복한 디지털 도서관은 정보통신 기술의 발전과 이용자의 기술적·문화적 성숙으로 혁신적인 변화에 직면하게 되었다. 온톨로지의 등장, 웹 2.0의 확산은 디지털 도서관이 의미중심 시스템으로 진화할 수 있는 계기가 되었다. 과거 일방·적인 정보제공 기능만을 수행한 정보시스템은 이와 같은 정보통신기술의 발전에 따라 이용자와 시스템 간의 상호작용을 실현할 수 있는 기반을 갖추게 되었다. 이에 더하여 단순한 데이터 자체에 지나지 않았던 정보가 메타데이터, 온톨로지의 활용으로 보다 유의미한 정보로 변모할 수 있었다. 또한 이 유의미한 정보는 이용자와 시스템 간의 상호작용에 있어 의미가 부여된 정보를 이용자에게 전달할 수 있는 기반이 되었다.

이처럼 급격하게 변화하는 현재의 정보기술의 발전에 따라 디지털도서관은 보다 지능적으로 변모하게 되었고 보다 의미 있는 정보를 전달하게 되었다. 이렇듯 의미기반 처리가 가능한 이용자 참여형 디지털 도서관을 본 논문에서는 SSDL(Social Semantic Digital Library)이라 명명한다. 국외의 경우 시맨틱 웹에 대한 관심이 높아지면서 SSDL에 대한 연구가 활발하게 진행되고 있으며 JeromeDL, SIMILE(Semantic Interoperability of Metadata and Information in unLike Environments) 등 SSDL 설계 및 구축에 관한 실증적 연구가 진행되고 있다. 그러나 아직 국내에는 시맨틱 웹에 대한 관심만큼 SSDL에 대한 관심이 높지는 않다.

시맨틱 웹의 유용성만큼이나 SSDL의 유용

성이 상당함은 국외의 연구와 프로젝트들을 통해 입증되고 있으며 그 가치는 날로 인정받고 있다. 국내에서도 SSDL의 구축과 그 유용성에 대해 연구하여 국내의 디지털 도서관에 적용함으로써 선진 기술을 확보해야 할 필요가 있다. 이에 본 연구는 현재 관심이 높아지고 있으며 그 가치가 인정되고 있는 SSDL의 구현을 위한 아키텍처 모델의 설계를 그 목적으로 한다. 본 연구의 2장에서는 SSDL의 개념에 대해 살펴보고 3장에서 SSDL에 적용될 수 있는 시맨틱 핵심 기술을 살펴본 뒤 4장에서 현재 진행되고 있는 시맨틱 웹 관련 사례를 통해 적용될 수 있는 최신 시맨틱 기술은 무엇이 있는지 살펴본다. 5장에서는 이들 정보기술을 적용하여 성공한 SSDL의 사례를 살펴보고 이상의 조사를 바탕으로 6장에서 SSDL에 요구되는 핵심 사항들을 분석한다. 마지막으로 7장에서는 이 분석과 조사된 기술들을 적용한 SSDL의 아키텍처 모델의 설계 내용을 기술하였다.

2. SSDL

정보기술의 발전과 기술 환경의 변화에 따라 도서관의 관점은 지협적 관점에서 광역적·의미적 관점으로 변화되었다. 도서관 관련 기술은 단순한 업무전산화에서 디지털정보자원관리, 글로벌 네트워크 환경 아래에서의 시스템 통합과 다양한 서비스 제공으로 발전되었으며 도서관의 물리적 한계 극복과 이용자 중심의 다양한 서비스 개발에 초점을 두고 있다.

그러나 정보자원 관리와 서비스 체계의 혁신을 위해서는 강력한 의미 표현 기능을 갖는 온톨

로지의 활용, 연관관계의 표현이 뛰어난 FRBR (Functional Requirements for Bibliographic Records)와 같은 개념모델을 활용하는 등, 보다 근본적인 접근이 요구된다(한성국, 이현실 2007).

차세대 웹으로 인식되고 있는 시맨틱 웹은 기계가독형 정의에 기반을 둔 정보의 연계를 통해 웹 자원을 지식화함으로써 정보의 효율적 검색, 통합, 재사용을 도모하는 새로운 기술이다(오삼균 2002). SSDL은 이러한 시맨틱 기술을 기존의 디지털 도서관의 설계에 도입하여 기능화된 정보자원 관리와 지능 에이전트에 의한 이용자 서비스 등을 제공하는 의미기반 디지털 도서관이다.

시맨틱 도서관은 분산 환경에서 독립적으로 운용되는 정보시스템간의 이질적 구조를 극복하여 의미적 상호운용성을 실현하고 의미적으로 통합되는 시스템의 구현에 역점을 둔다(한성국, 이현실 2007).

3. 시맨틱 웹의 핵심개념

시맨틱 웹은 집단지성의 실현을 지향하는 웹 2.0과 컴퓨터가 스스로 정보를 처리할 수 있는 환경을 제공하는 온톨로지 등의 의미기반 처리 시스템을 그 구성요소로 하고 있다. 또한 Ajax, RDF/OWL, TopicMaps 등과 같은 시맨틱 언어를 적극적으로 활용하기 때문에 보다 향상된 재사용성을 제공한다. 본 장에서는 시맨틱 웹의 핵심 개념을 살펴봄으로써 SSDL에 적용 가능한 시맨틱 기술에 대해 알아보려고 한다.

3.1 Uniform Resource Identifier(URI)

URI는 불변성 식별자를 부과하는 체제의 개념으로 사용된다. 자원에 할당되는 고유식별자의 불변성은 정보의 효율적 관리와 시맨틱 웹의 근간을 구성하기 위해 매우 중요한 특징이다. 정보의 효율적 관리 측면에서는 자원의 소재가 변경되더라도 식별자는 변하지 않아 자원을 제대로 찾아 주는 역할을 감당할 수 있다는 점, 시맨틱 웹의 근간 구성 측면에서는 메타데이터 요소와 주요 통제어에 URI를 부과하여 자원의 자유로운 재활용이 가능해진다는 점이 매우 중요하다.

URI의 기능이 제대로 활용되기 위해서는 각 URI의 값을 실제 자원과 연결된 URL로 변환시켜 주는 시스템이 필요하다. 현재 URI 변환시스템으로 가장 잘 알려진 시스템은 미국 CNRI(The Corporation for National Research Initiatives)에서 개발한 핸들(HANDLE)시스템이다. 또한 한국정보사회진흥원(NIA)에서 핸들과 유사한 UCI 서버를 활발히 운영하고 있으며, 더블린코어에서도 유사한 개념으로 http 형식의 URI 값을 제공하는 형태로 영속적인 URL이라는 개념의 PURL(Persistent URL)을 사용하고 있다(오삼균, 2002).

시맨틱 웹에서 URI 체제는 메타데이터 속성 및 클래스의 식별자로 활용되어 아직 정의되지 않은 속성에 대해서만 새로운 URI를 부과하고 이미 정의된 메타데이터 속성은 재사용함으로써 메타데이터 상호운용성을 개선할 수 있을 것이다.

3.2 온톨로지

시맨틱 웹은 컴퓨터가 이해할 수 있는 데이터의 생성을 추구한다. 컴퓨터가 데이터를 이해하고 데이터를 스스로 처리·가공할 수 있는 환경을 마련하기 위해서는 온톨로지로 정보를 표현하는 작업이 필요하다. 온톨로지는 실세계에 존재하는 모든 개념들과 그 개념 지식을 명세화 하기 위해 개념들 간의 속성, 의미 관계에 관한 정보를 정의하는 것이다(Gruber, 2001). 온톨로지는 웹상에서 애플리케이션 간에 지식을 공유하고 처리할 수 있도록 하는 중요한 역할을 한다. 또한 애플리케이션 시스템과 사람이 의사소통 할 수 있도록 도메인의 일반적인 이해와 공유를 제공한다. 지식온톨로지를 표현하는 대표적인 언어는 W3C의 표준으로 채택된 RDF/OWL과 ISO의 표준으로 채택된 TopicMaps(이하 토픽맵)가 있다.

3.2.1 RDF/OWL(Resource Description

Framework/Web Ontology Language)

RDF/OWL은 XML 스키마로 처리하기 힘든 메타데이터 요소들의 의미와 다른 요소들과의 관계를 컴퓨터가 이해 할 수 있도록 정보를 표현하는 개념의 틀을 제시한다. RDF는 자원, 속성, 속성값을 하나의 단위로 취급하는 'TRIPLE'을 그 핵심개념으로 삼는다. RDF에서는 자원을 주어로, 속성을 서술어로, 속성의 값을 목적으로 표현하여 자원과 자원의 속성에 대한 정보를 컴퓨터가 이해할 수 있는 형태로 표현한다.

RDF 스키마는 메타데이터 속성, 속성들 간의 상·하 관계, 클래스들 간의 상·하관계를 기계가독형으로 정의하기 위한 구문구조를 제

공하고 있다. RDF 스키마에는 미리 정해진 몇 개의 어휘들이 있고 이를 통해 RDF 구문에 쓰이는 어휘 사이의 관계를 의미적으로 정의하게 된다. RDF는 속성과 다른 자원과의 관계를 정의하는 방법은 제공하지 않는다. 이를 보완하여 웹 문서에 완전한 정보의 의미를 부여하기 위해 자원의 종류나 클래스를 표시하는 일이 RDF 스키마의 역할이다.

DAML+OIL은 주요 관계를 지원하는 표현력이 결여된 RDF와 RDF 스키마의 모델링 요소를 확장, 강화한 언어이다. OWL은 이 DAML+OIL에 기반을 둔 온톨로지 구축 경험을 토대로 개념의 일관성을 확보하여 클래스와 속성의 개념 및 그들 사이의 관계가 보다 명료하게 정의되도록 정리한 온톨로지 언어이다.

OWL은 추론시스템에 축적된 논리적 명제들을 정의하는 언어로서, 이러한 명제는 하나의 온톨로지에 근거하거나 또는 OWL이 명시하는 방식에 따라 수집한 다수의 분산형 온톨로지에 근거할 수 있다. OWL 온톨로지는 클래스와 속성, 그에 적용되는 갖가지 제약사항들의 집합으로서 클래스들 사이의 텍사노미 관계, 데이터의 속성, 객체의 속성, 클래스들의 인스턴스, 속성들의 인스턴스를 포함한다. OWL 명제는 클래스의 구성원들에 관한 사실과 구성원들 간의 관계를 기술하며, 이러한 명제의 집합으로 이루어진 온톨로지는 구문적으로 정의되지 않은 사실의 논리적 유추를 가능하게 한다.

3.2.2 토픽맵(Topic Maps)

RDF/OWL이 추론을 장점으로 가지고 있다면 토픽맵은 연계를 통한 관련된 모든 정보를 나열하여 브라우징을 통한 정보 발견이 보다

용이하다는 장점을 가지고 있다. TAO(Topic, Association, Occurrence)로 알려진 토픽맵의 기본 구성은 정보와 정보 간의 유의미한 관계를 통해 이용자에게 유용한 의미론적 정보를 제공할 수 있는 환경을 구성한다.

토픽은 현실세계에 존재하는 모든 것, 즉 추상적 개념을 기계가 이해할 수 있는 구체화된 객체로 바꾸어 표현한 것이다. 이러한 토픽은 인간 가독용 이름, 기본 이름을 가지며 연계, 즉 토픽 간의 관계에 따라 다른 역할을 가진다.

연계는 토픽맵 안에 정의된 토픽 간의 관계를 설정하는데 사용된다. 이를 통해 토픽맵 안에서 토픽이 갖는 문맥 정보를 제공한다. 토픽맵 안에 정의된 연계는 방향성을 갖지 않는다. 따라서 토픽에 대해 알지 못하는 사람이나 컴퓨터의 관점에서는 토픽 간의 관계가 분명치 않다. 이것을 분명하게 하기 위해 토픽맵에서는 관계를 설정할 때 역할을 활용하여 이 부분을 분명케 한다. 이러한 특징은 연계에 따라 이용자에게 제시되는 토픽의 성격이 달라, 보다 유의미한 정보를 제공할 수 있다는 점에서 주목할 필요가 있다.

어커런스는 토픽을 정보 자원과 연결시키는 역할을 한다. 토픽맵이 토픽에 관한 정보를 제공할 때, 각 토픽에 관한 어커런스로 토픽에 관한 어떤 정보가 존재하는지를 확인할 수 있다. 하나의 토픽은 많은 자원과 연결될 수 있지만 많은 자원이 동시에 많은 자원과 연결될 수는 없다.

토픽맵에서는 이들 기본 구성요소를 묶어 재사용할 수 있도록 한다. 즉, 유사한 범주에 속하는 각 요소를 타입(type)이라는 개념으로 묶어 새로운 토픽으로 간주하고 이에 연계와 어커런스를 통해 관계를 나타내고 정보 자원을 제공한다.

‘주제식별자’는 주제의 식별을 위한 기능이다. 주제가 직접 접근이 가능한 자원인 경우에는 이 주제식별자의 값은 자원의 URL 주소가 될 수 있다. 그러나 대부분의 자원인 사람이나 개념은 직접적으로 컴퓨터에 표현할 수 없기 때문에, ‘주제지시자’를 이용하여 그 개념을 표현하는 방식을 취한다. 주제지시자는 “주제의 식별을 위한 설명을 제공 하는 웹 페이지”이다. 주제지시자 또한 자원이므로 보통 주제 식별자로 쓰일 수 있는 URI 주소를 갖는다. 주제 식별자는 컴퓨터 처리를 위한 것이고, 주제지시자는 그 주제에 대한 사람의 이해를 돕기 위한 것이다.

3.3 웹 2.0(Web 2.0)

인터넷 기술이 성숙하게 되고 인터넷 이용자들이 기술적·문화적으로 성숙하게 됨에 따라 보다 이용자 중심적인 웹 기술이 요구되었다. 이러한 요구에 부응하여 90년대 말 닷컴 기업 붕괴에도 살아남은 기업들에서 보다 향상된 웹 기술을 개발하고자 하는 노력들이 있었으며, 이들 기업이 공통적으로 추구하는 웹 기술의 방향을 웹2.0으로 총칭하게 되었다(O'Reilly, 2005). 웹 2.0은 향상된 정보기술의 개발에만 국한되지 않고 웹 기반 정보 서비스에 총체적 기술현상으로 발전되었다.

웹2.0은 양방향의 읽기와 쓰기 웹으로써 이용자가 웹상에서 만나고, 협업하고, 공유할 수 있는, 웹기반 서비스가 지향하여야 하는 신개념을 제시한다. 블로그, 위키, RSS, Ajax 등은 웹2.0을 확산시킨 주요한 서비스와 기술들이며 위키피디아(Wikipedia), 디그(Digg) 등은 이들

기술을 활용한 웹2.0을 표방하는 사이트들로 웹서비스의 새로운 장을 열고 있다. 웹2.0은 개방과 공유, 참여와 협업의 이념을 바탕으로 정보서비스가 지향해야할 새로운 목표로 제시되고 있다. O'Reilly에서는 웹2.0의 7원칙을 다음과 같이 정의하고 있다.

- ① 웹은 정보의 전달, 공유, 활용을 위한 플랫폼으로서의 기능을 수행한다.
- ② 이용자의 참여, 지식의 공유를 통해 더욱 우수한 서비스를 생성한다.
- ③ 데이터는 차별화된 서비스의 원동력이다.
- ④ 이용자의 의견과 요구를 반영한, 실시간적인 지속적 갱신을 수행한다.
- ⑤ 단순하고 실용적인 프로그램을 신속하게 개발한다.
- ⑥ 다양한 정보기기에 정보서비스를 제공한다.
- ⑦ 다양한 응용프로그램과 서비스를 전달하여 풍부한 이용자 경험이 실현되도록 한다.

웹2.0에서 활용되고 있는 정보기술들은 도서관 업무와 서비스에 도입하여 응용될 수 있다. 현재 문헌정보시스템에도 웹2.0의 개념이 도입되어 라이브러리2.0(Library2.0)에 대한 논의가 진행되고 있다. Bonariabiancu는 웹2.0의 개념을 수용하여 개념도를 작성하고 다음의 8가지 원칙을 제시하였다(Bonariabiancu, 2006).

- ① 이용자 중심
- ② 기술 이해 환경
- ③ 이용자의 장기적 접근 허용
- ④ 하나 이상의 장치를 위한 콘텐츠
- ⑤ 단일 문헌정보시스템이 아닌 구성기반 소프트웨어
- ⑥ 지속되는 변화

- ⑦ 웹2.0 애플리케이션과 서비스의 이용
- ⑧ 개방형 표준

4. 시맨틱 웹 사례

시맨틱 웹 기술은 디지털 도서관에 적용하여 SSDL을 구현할 수 있는 가장 최신의 기술이다. 현재 이와 관련한 프로젝트가 활발하게 진행되고 있으며 연구 논문을 통해 그 실효성을 검증하고자 하고 있다. 본 장에서는 시맨틱 웹 기술과 관련한 사례를 통해 SSDL에 적용할 수 있는 최신 정보기술, 국제표준 및 모범사례들을 살펴보고자 한다.

4.1 FOAF(Friend Of A Friend)

Dan Brickley와 Libby Miller는 시맨틱 웹 기술에도 적용할 수 있으며 인물 정체성 관리상의 문제를 해결하기 위한 메타데이터로 FOAF를 제시하고 명세를 작성하였다. FOAF는 RDF를 사용하여 클래스와 속성들을 정의하였다. FOAF에서는 다음과 같은 상황을 설명할 수 있는 구조를 제공한다.

- ① 그룹(group)에 소속된 개인(person)
- ② 조직(organization)에 의해 생성된 문서(document)
- ③ 한 개인이 알고 있는 다른 개인

FOAF에서 개인의 표현은 단순하게 표현된 FOAF 파일의 집합과 연계를 통해 복잡하게 표현될 수 있다. 즉, 개인으로서의 표현, 조직에 속한 개인으로서의 표현, 다른 개인과 관련된 개

인으로서의 표현을 통해 복잡한 개인의 정보가 표현되는 것이다. 그러나 이러한 개인이 동일하게 체인지는 각 FOAF 파일에서 표현되지 않는다. 여러 FOAF 파일에서 정체성은 이메일에 의해 식별된다. 개인은 여러 이메일 주소를 가질 수 있지만 이메일은 반드시 하나의 개인만을 가질 수 있다. 따라서 이메일 주소가 같은 개인은 동일한 인물로 간주될 수 있는 것이다. 즉, FOAF에서는 두개의 개체를 설명하는 파일에서 동일한 이메일 주소가 검색되어지면 그 두개의 개체들이 동일인물이거나 적어도 상관관계를 가지고 있다고 간주된다. 개인에 대한 설명에 이메일이 없다면 차선책으로 개인 홈페이지의 URI를 이용하는 방법을 사용할 수 있다.

이와 같은 시도들은 도서관의 경우 전거파일 에 적용될 수 있으며 명칭관련 관리가 용이할 수 있는 시스템을 마련하는 기반으로 작용할 수 있을 것이다(Graves, 2007).

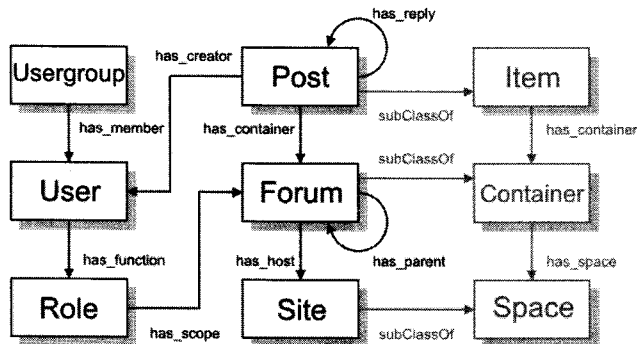
4.2 SIOC(Semantically-Interlinked Online Communities)

SIOC는 온라인 커뮤니티가 보유한 정보를 의

미적으로 연결하기 위한 시도이다. DERI(Digital Enterprise Research Institute)와 NUI galway (Galway of the National University of Ireland)의 주관으로 진행되는 이 프로젝트는 온라인 커뮤니티 내에 존재하는 정보를 다른 커뮤니티와 의미적으로 연결하기 위해 커뮤니티의 특징, 포스트의 세부 내용을 의미적으로 표현하는 것을 주요 목표로 하고 있다. SIOC는 이를 위해 커뮤니티 간 관계 표현을 위한 SIOC 온톨로지와 SIOC 추출/검색을 위한 애플리케이션으로 구성되어 있다.

SIOC 온톨로지는 RDF/OWL을 이용해 표현되며 Forum, Post, Role, Site, User, UserGroup의 6개 클래스와 속성으로 구성되어 있다. 또한 SIOC 온톨로지는 FOAF, RSS, 더블링크어 등의 메타데이터를 이용하여 확장할 수 있다.

그림 1은 SIOC의 온톨로지 구조로 6개의 주요 클래스를 통해 커뮤니티와 커뮤니티에 포스팅 된 정보를 표현하고 3개의 서브 클래스를 통해 다른 타입의 커뮤니티와 포스팅 된 정보를 표현한다. 서브 클래스의 사용으로 다른 커뮤니티에 전제된 포스팅에 대한 정보를 의미적으로 연결할 수 있게 된다.



(그림 1) SIOC 온톨로지

4.3 Social Information Architecture

시맨틱 기술의 발전과 웹2.0의 확산으로 이용자 참여를 바탕으로 한 정보 서비스의 큰 흐름이 나타나게 되었다. 이용자 간 정보 교류를 도와주고 촉진시키는 소프트웨어의 총칭인 social software, 정보·자원에 대한 이용자 분류의 총칭인 social tagging, 정보 구성에 대중의 의도를 반영해야 한다는 개념의 social information architecture 등 집단지성이 고려되는 기술현상에 집단의 의미를 부여하여 'social'이란 어두를 사용하고 있다. 이는 차세대 웹의 흐름이 이용자 중심에서 참여하는 대중으로 변하고 있음을 시사하고 있다고 할 수 있다. 즉, 이용자의 잠재된 요구를 충족시키기 위해서는 표출된 대중의 요구를 반영할 수 있어야 한다는 것이다.

이들 기술현상 중 Gene Smith가 제안한 Social Information Architecture(이하 SIA)는 SSDL의 IA를 구성함에 있어 필요한 의미 기반 인터페이스를 제공할 수 있다. SIA의 기본이 되는 개념은 이용자의 집단적 참여에서 찾을 수 있다. 즉, 특정 정보에 대한 이용자들의 의견과 평가를 순위, 관심사 등의 장치를 통해 정보에 반영하고, 클러스터링, 협력적 필터링 등의 유의미한 정보를 제공할 수 있는, 이용자 참여에 기반을 둔 알고리즘을 개발하는 것이다.

Gene은 SIA 설계를 위한 5가지 원칙을 제시하였다.

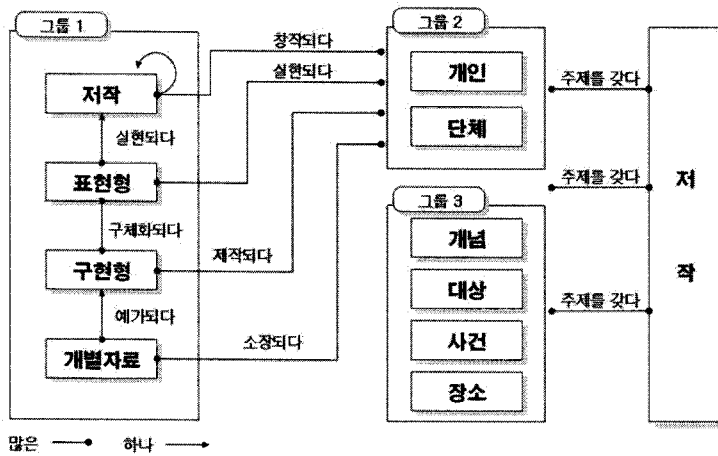
- ① 이용자연대의 다른 수준을 허용
- ② 피드백 순환의 모니터와 변경
- ③ 교환: 투명성 대 학습화
- ④ 보다 큰 시스템으로의 참여
- ⑤ 새로운 행동, 집합체, 디스플레이를 디자인

4.4 FRBR(Functional Requirements for Bibliographic Records)

IFLA에서는 다양한 매체, 애플리케이션, 이용자 요구에 맞는 서지정보와 관련 정보체계를 분명하게 정의하기 위해 개체-관계(entity-relation) 기반의 개념틀인 FRBR을 발표하였다(IFLA, 1998). FRBR은 이용자의 서지 레코드 이용 목적을 탐색(find), 식별(identifier), 선정(select)과 확보(obtain)의 4가지 요소로 정의하고 이러한 목적을 달성할 수 있는 최소 수준의 개념틀을 제시하였다(한성국, 이현실 2007).

FRBR은 그림 2와 같이 3개의 그룹으로 구성되며 각 그룹은 논리적 연관관계를 갖는다. 그룹 1은 서지 레코드에 명시되거나 기술된 지적, 예술적 산물을 표현하는 개념 수준의 저작, 표현형, 구현형과 물리적 실체를 지시하는 개별 자료로 구성되며 이들은 그룹 2의 개인과 단체에 의해 어떠한 책임이 지워져 있는지 표현되며 그룹 3은 저작에 반영된 개념, 대상, 사건, 장소 등의 주제를 정의한다. 이로써 FRBR은 기존의 서지레코드가 표현하지 못했던 저작과 행위자, 주제 간의 다양한 관계를 의미적으로 표현할 수 있게 하였다.

SSDL 설계에 있어 FRBR은 소장 도서의 의미론적 이용을 위해 적용할 수 있는 좋은 모델로 인식되어 왔다. 기존의 물리적 도서관장서는 모두 MARC를 사용하여 정리, 관리되고 있다. 이에 현재 LC나 OCLC에서 기존의 MARC 기반 서지 정보를 FRBR로 자동 변환하는 알고리즘을 개발하고 있으며 MARC의 FRBR 자동 변환에 따라 보다 용이하게 SSDL의 기반을 구축할 수 있을 것이다.



(그림 2) FRBR의 구조

5. SSDL 사례

본 장에서는 SSDL의 대표적 사례인 JeromeDL과 시맨틱 기술의 제고를 목적으로 하는 프로젝트인 SIMILE를 살펴봄으로써 본고에서 설계하고자 하는 SSDL 구조의 바탕으로 삼고자 한다.

5.1 JeromeDL

Gdansk 공과대학 중앙도서관과 DERI는 디지털 도서관에 시맨틱 요소를 적용하여 웹상에서 쉽게 문서를 출판하고 여러 문서 형식의 지원, 저장, 검색을 실현하기 위한 협동 프로젝트로 JeromeDL을 진행하였다.

JeromeDL이 제시한 SSDL 요구사항은 다음과 같다.

- ① 전통 도서관 유산의 지원
- ② 이용자 중심적 브라우징 지원
- ③ 효과적인 검색
- ④ 보안과 계정 제한

- ⑤ 자원의 다양한 형식 지원
- ⑥ 다른 디지털 도서관 시스템과의 커뮤니케이션 지원
- ⑦ 시맨틱 웹과 커뮤니케이션, 정보 관리 연구의 최근 결과를 이용

JeromeDL은 이상의 요구사항을 실현하고자 자원의 이용, 즉 열람, 검색, 브라우징을 위한 미들웨어를 온톨로지로 구현하였고 이용자의 프로파일을 의미적으로 관리하기 위해 FOAF를 사용하였다.

FOAF는 JeromeDL에서 보다 의미적인 정보를 제공하는 데에도 사용되었다. Social Semantic Collaborative Filtering은 FOAF를 기반으로 한 의미적 자원 필터링 방법이다. 이는 FOAF를 통해 연결된 이용자 간의 연계와 순위 선정을 통해 의미적으로 유용한 자원을 선별해 낼 수 있는 알고리즘을 제공한다.

JeromeDL은 시스템 관리와 서지정보, 커뮤니티에 온톨로지를 사용하였다. FOAF와 FOAF-

중간층은 온톨로지를 사용하여 자원 간 관계를 MARC21, 더블링크어 등을 사용하여 확장적으로 기술하고 여러 시맨틱 요소들을 사용하여 최상위층에서의 시맨틱 서비스를 가능케 한다.

최상위층에서는 커뮤니티 연대와 주석화, 협력적 필터링 등의 웹2.0 요소들을 기반으로 한 유틸리티를 사용할 수 있다. JeromeDL은 다양한 디지털 도서관 서비스를 통합하기 위해 관리도구인 KOS(Knowledge Organization System)을 사용한다. KOS를 통해 사서는 서비스를 관리할 수 있다. Ajax로 구현된 인터페이스와 제시되는 통제어를 통해 손쉽게 변경과 추가가 가능하다. JeromeDL은 유연하고 확장적인 구조를 가지고 있다. 즉, 자원을 분류하는 방법이 있어 기존의 분류법인 DDC와 웹2.0의 분류법인 폭소노미(folksonomy)를 함께 사용하며, 자원의 서지정보를 기술하는 방법에 있어서도 온톨로지 목록인 MarcOnto와 일반 메타데이터인 MARC21, 더블링크어 등을 함께 사용할 수 있도록 한 것이다.

5.2 SIMILE

SIMILE(Semantic Interoperability of Metadata and Information in unLike Environments) 프로젝트는 MIT 도서관과 MIT CSAIL(Computer Science and Artificial Intelligence Laboratory)의 공동 프로젝트로 개인, 공동체, 기관이 보유한 디지털 자산, 스키마/어휘집/온톨로지, 메타데이터 그리고 서비스 간의 상호운용성을 제고하여 이용자에게 서비스를 제공하는 방법을 찾는 것을 목적으로 하고 있다.

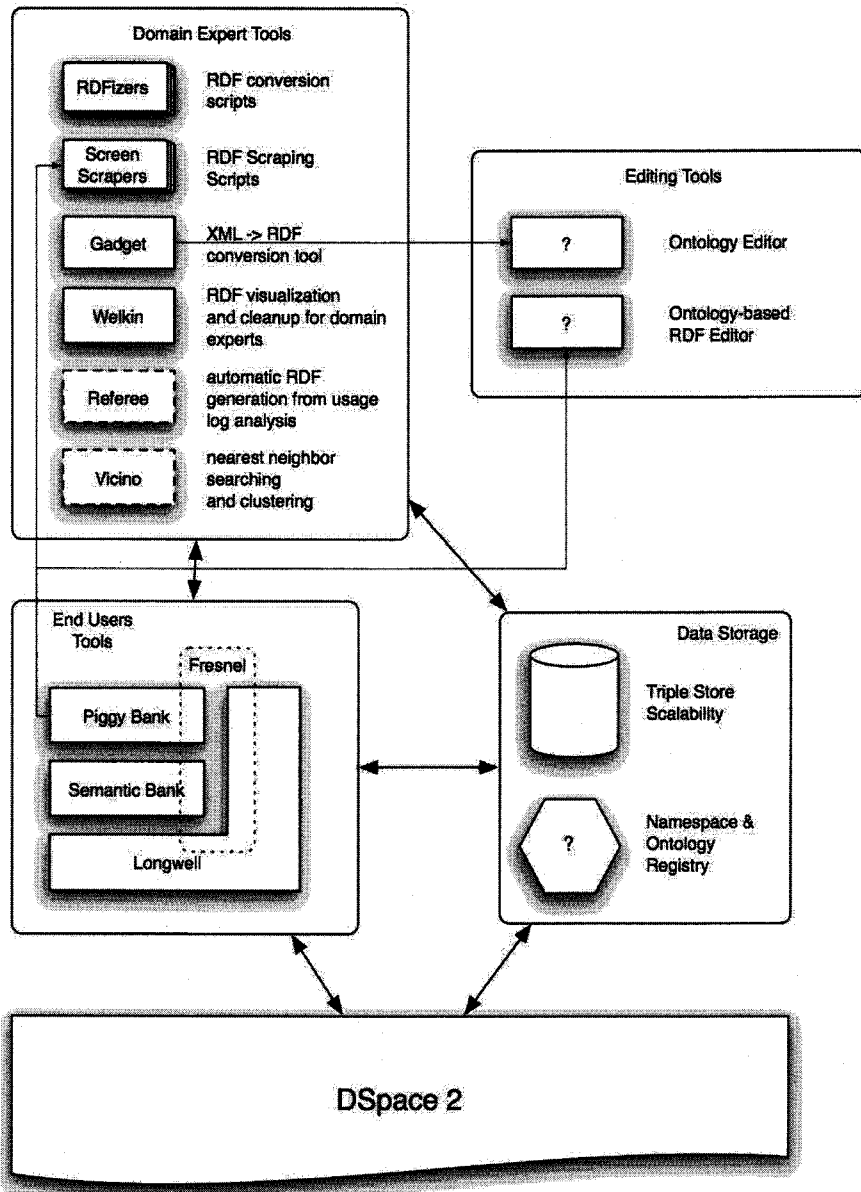
SIMILE 프로젝트는 RDF와 시맨틱 웹 기술을 사용하여 무분별한 스키마와 메타데이터를 지원하는 방법을 연구하고 있으며, 웹 표준 기반 디지털 자산 보급 아키텍처를 적용하는 것을 목표로 하고 있다. 보급 아키텍처는 특정 디지털 가공에 유용한 관점과 서비스의 소비를 위한 유용한 관점의 묶음을 더한 메커니즘을 제공하는 것이다.

이러한 SIMILE 프로젝트는 MIT의 디지털 저장소인 DSpace의 기능적 향상에 시맨틱 요소를 적용하여 의미론적으로 보다 유용한 서비스를 제공하기 위해 시작되었다. 2년에 걸친 초기 연구에서 얻은 결과는 2년 동안의 후속 연구에서 Dspace 뿐만 아니라 이와 유사한 플랫폼에도 적용할 수 있는 프로토타입을 개발하는데 활용되고 있다.

그림 5는 SIMILE 프로젝트에서 시맨틱 웹 기술을 이용하여 개발된 어플리케이션과 이들을 사용하여 설계한 프레임워크를 보여주고 있다. SIMILE 프로젝트는 크게 전문 도구 영역, 이용자 도구 영역, 데이터 저장소, DSpace 메타데이터 엔진의 4부분으로 나누어 프레임워크의 기초를 구성하였다. 각 영역에 대한 상세한 내용은 다음과 같다.

■ 전문 도구 영역

전문 도구 영역은 정보처리, 검색, 브라우징, 네비게이션, 디스플레이를 위한 도구를 RDF와 시맨틱 웹 기술을 사용하여 개발하는 역할을 담당하고 있으며 이 영역에서 개발된 도구들은 의미기반 디지털 자원의 생성과 연계, 온톨로지 구현 및 구축 등 RDF 온톨로지의 구축과 온톨로지 기반 정보 탐색을 가능하게 하는 것



(그림 5) SIMILE 시맨틱 어플리케이션과 프레임워크

을 목적으로 개발되었다. 현재 RDF 온톨로지 구축 도구인 RDFizers, Screen Scrapers, Gadget, Referee 등이 개발되었으며 브라우징, 네비게이션, 디스플레이 도구인 Exhibit, Timelines,

welkin 등이 개발된 상태이다. 이들 도구를 사용하여 새로운 자원뿐만 아니라 기존의 자원들도 RDF 형식의 의미기반 정보로 변환하여 사용할 수 있다.

■ 사용자 도구 영역

사용자 도구 영역은 RDF와 시맨틱 기술을 사용하여 디지털 자원의 브라우징, 검색, 네비게이션을 위한 상호교류 도구를 개발하는 역할을 하며 RDF 기반의 유용한 사용자 인터페이스를 개발하는 것을 목적으로 하고 있다. 현재 Piggy Bank, Semantic Bank, Longwell, Fresnel 등의 어플리케이션이 개발되었으며 이는 전문 도구 영역에서 개발된 어플리케이션으로 생성·변환된 의미기반 데이터와 온톨로지를 이용자와 연결하는 역할을 수행하고 있다.

■ 데이터 저장소

데이터 저장소는 효과적이며 확장 가능한 저장 그리고 데이터의 관리 및 탐색을 위한 기반을 마련한다. 이를 위한 도구로는 RDF의 트리플을 반영한 Triple Store Scalability와 이름 공간, 온톨로지 레지스트리를 사용하고 있다.

■ DSpace 메타데이터 엔진

기존의 DSpace를 확장하여 의미기반 시스템으로 재구성하기 위해 DSpace 메타데이터 엔진이 사용되었다. 이 메타데이터 엔진은 정보 저장소와 정보 탐색 구성요소에 시맨틱 기술과 시맨틱 웹 도구를 적용하여 기존의 DSpace에 축적된 자원을 의미적으로 다룰 수 있도록 하고 있다.

SIMILE 프로젝트는 온전한 디지털 도서관의 아키텍처를 구성하여 시맨틱 기반 서비스를 진행한 프로젝트는 아니다. 그러나 4 개 영역을 기본 구조로 하여 영역내의 다양한 도구의 개발과 그 도구의 실제적 적용을 통해 디지털 도

서관에 적용될 수 있는 기술들의 실효성과 효율성을 실증적으로 보여주었다. 개별화, 모듈화된 부분을 필요에 따라 구성요소로 사용할 수 있으며 요구에 따라 독립적인 운용이 가능하도록 하였다는 점에 주목할 필요가 있으며 디지털 도서관의 시맨틱으로의 진화를 가능하게 하고 있다는 점에서 SSDL 아키텍처 설계하는 데 참고할 가치가 있다.

6. SSDL의 요구사항

이용자 참여형 시맨틱 도서관은 아직 우리에게 생소하다. 그 기본 틀을 디지털 도서관에 두고 있지만 그 기능과 요구사항은 분명히 다르다. 아직 시맨틱 도서관에 대한 많은 연구가 진행되지 않았으며 시맨틱 도서관에 대한 개념 정의도 분명하지 않다. 따라서 SSDL에 요구되는 이용자, 관리자, 사회적 요구와 구현 시 필요한 기술적, 구조적 요구에 대해 디지털 도서관의 요구사항과 현재의 정보기술, 그리고 이를 이용하는 이용자들의 변화추세를 고려하여 SSDL 구현 시 사회적으로 요구되는 기능적 요구사항과 SSDL 자체에 요구되는 구조적 요구사항을 살펴보고자 한다.

6.1 기능적 요구사항

SSDL은 기존 디지털 도서관의 단순한 확장이 아닌 사회적, 의미적으로 확장된, 의미 기반의 총체적인 정보시스템이다. 달리 표현하면, SSDL은 자원과 자원 간의 유의미한 연관관계를 기반으로 잠재된 이용자의 정보요구까지 총

족시킬 수 있는 새로운 정보시스템인 것이다. 따라서 자원을 의미적으로 해석하고 이를 이용하는 이용자들에게 대한 의미 기반의 정보서비스가 이루어져야 할 것이다. 이러한 맥락에서 SSDL에 요구되는 핵심사항들을 살펴보고자 한다.

■ 의미기반 서비스 제공

이용자들의 보다 향상된 정보 탐색을 지원하기 위해 의미기반 검색과 브라우징이 제공되어야 하며 다른 SSDL과의 협력적 업무를 위해 이종 시스템 간 상호운용성 보장과 OAI-PMH와 같은 프로토콜 수준에서의 협력적 메타데이터 교환이 가능해야 한다. 무엇보다도 제공되는 이들 서비스는 의미기반 정보 처리, 정보 교환이 고려되어야 한다. 이를 위해서는 정보자원과 업무 서비스가 기계가독형 온톨로지로 표현될 수 있도록 기술되어야 한다.

■ 다양한 접근점 제공

단순한 정보 제공이 아닌 정보를 찾아 갈 수 있는 다양한 접근점이 제공되어야 한다. 의미기반 검색의 결과는 컴퓨터가 판단한 결과로 인간이 의도한 바와는 다른 결과를 줄 수 있다. 또는, 따라서 보다 정보의 연계에 기반한 총체적인 접근점을 제시하여 사용자 스스로가 탐색을 통해 자신의 의도에 맞는 정보를 찾을 수 있도록 해야 한다.

■ 집단 지적 활동 지원

도서관 이용자뿐만 아니라 도서관 관계자도 상호 협력적인 지적 활동을 할 수 있는 공간을 지원해야 한다. 커뮤니티를 활용한 협업작업, 정보공유 등 다양한 지적 활동을 지원해야 한

다. 집단 지적 활동은 정보의 재생산에 기여할 것이며 이용자의 참여를 통해 생성된 새로운 지식은 전문가 분석에 의한 온톨로지와 통합될 것이다. 또한 이 통합된 지식은 의미 있는 정보를 이용자에게 제시하는 기술적 방법에 있어서도 큰 협력자가 될 것이다.

■ 정보자원의 효율적 관리

도서관이 소장하고 있는 자원뿐만 아니라 메타데이터, 온톨로지, 이용자 프로파일, 집단지성 등 SSDL 운영과 관련된 자원을 효과적 그리고 효율적으로 관리해야 한다.

6.2 구조적 요구사항

SSDL은 최신의 기술을 사용하여야 할 뿐만 아니라 지속적인 기술혁신을 적용해야 한다. 따라서 SSDL은 구현이 용이하지 않으며 구현된 후에도 도태되기 쉽다. SSDL 구축 시 고려해야 할 핵심적 요구사항들을 분석하여 구축에 필요한 원칙을 제시한다.

■ 모듈화

시맨틱 기술뿐만 아니라 현재 사용되고 개발되는 정보기술의 생명주기는 매우 짧다. 이러한 정보기술의 특징은 전체 시스템의 기술적 낙후를 가져오며 결국에는 시스템의 도태를 초래한다. 이러한 문제점을 해결하기 위해서는 하나의 시스템으로 운영되는 것 보다는 여러 구성요소로서의 소프트웨어를 집합적으로 사용하는 시스템으로 운영되는 것이 필요하다. 다시 말하면, 시스템을 구성하는 각 부분을 모듈화 하여 개별 구성 요소의 기술적 도태에 따

라 변경을 가능케 하는 것이다. 이는 전체 시스템이 도태되는 것을 막을 뿐만 아니라 필요에 따라 모듈을 개선하고 변경하는 것만으로 시스템이 변하는 것을 가능하게 해 준다.

■ 의미기반 네트워크

하나의 도서관에서 모든 자원을 보관하거나 저장하고 이를 서비스 할 수는 없다. 이미 많은 도서관의 시스템들은 네트워크를 통해 직접적 그리고 간접적으로 연결되어 있으며 상호 협력적인 업무를 네트워크를 통해 협업하고 있다. 그러나 이는 단순히 기관과 기관 시스템과 시스템의 물리적 연결일 뿐이며 그 상호 협업 동안 의미가 부여된 정보의 교류는 미미하다. SSDL은 상호호환성을 가진 메타데이터와 상호운용성을 보장하는 온톨로지를 구축하여 의미적 연합이 가능한 의미기반 네트워크를 구성해야한다. 이는 시스템간의 자원공유만이 아니라 서비스 및 업무 연합도 지원할 수 있어야 한다. OAI-PMH와 같은 표준 프로토콜과 SIOC와 같은 네트워크 온톨로지가 이러한 네트워크를 구성하는데 유용할 것이다.

■ 자원 및 저작권을 보호

의미기반 네트워크를 구현하기 위해서는 시스템이 개방되어 있어야 한다. 이러한 환경에서는 보안 문제가 중요하다. SSDL은 도서관 업무와 관련된 정보와 이를 이용하는 이용자에 대한 정보를 모두 포함하고 있으므로 이를 외부의 공격으로부터 보호하는 것이 필요하다. 뿐만 아니라 저작권은 도서관 시스템의 가장 큰 보호 대상 중 하나이므로 이 또한 고려되어야 한다.

7. SSDL(Social Semantic Digital Library) 설계

7.1 SSDL의 기본 방향

SSDL의 핵심은 도서관 서비스의 시맨틱의 하부구조를 제대로 구성하는 것과 이용자 참여를 통해서 그 시맨틱을 확장해 나가는 것이다. SSDL에서 중요한 시맨틱은 '연계시맨틱(Association Semantics)'과 '참조시맨틱(Reference Semantics)'이라고 본다. 연계시맨틱을 보장하기 위해서 현재의 MARC나 더블린코어와 같은 메타데이터를 기계가독형으로 표현해야 하고 메타데이터 간의 보다 긴밀한 연결 관계를 구축해야 한다. 이러한 연계시맨틱을 제대로 수립하기 위해서는 W3C에서 개발한 RDF/OWL과 ISO에서 개발한 토픽맵 온톨로지 언어를 활용하여 데이터베이스 구조 자체를 변경해서 새로운 시맨틱이 수시로 추가될 수 있는 체계를 마련해야 한다.

그리고 참조시맨틱을 보장하기 위해서는 토픽맵 분야에서 정립한 공적주제식별체계를 국가적 차원에서 갖추어야 한다. 이를 위해서는 공적주제식별자와 공적주제지시자를 구별하여 기계는 식별자를 사용하고 사람은 주제지시자의 설명을 통해서 모든 메타데이터 요소, 주제, 인명 등이 충돌이 없이 통합될 수 있는 체계를 마련해야 한다.

참조시맨틱에 대한 개념은 지식의 무결성을 강조한 Steve Pepper의 주장에서도 그 맥락을 같이 하고 있다. 메타데이터, 온톨로지, 텍사노미에 대한 관심의 증대와 월드 와이드 웹에서 시맨틱 웹으로의 진화 속에서 주변의 급박한

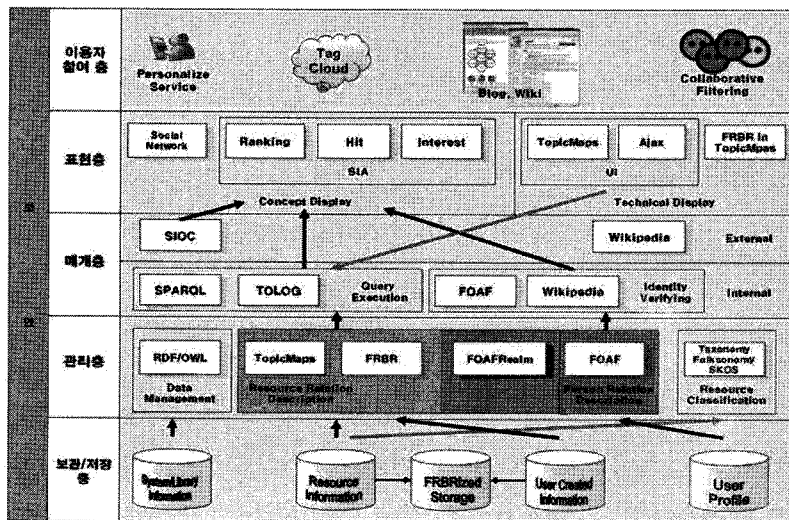
신호들을 볼 수 있고, 무결한 지식의 핵심 개념에도 내포되어 있다. “무결한 지식”이라는 단어는 각각 분리된 시스템과 저장소 사이에서 자동적으로 수집되고 합쳐질 수 있는 지식과 정보의 개념을 표현하는데 쓰인다. 무결한 정보를 위한 전제조건은 주제중심의 콘텐츠 조직이다. 이는 글로벌 식별체계 문제의 해결책으로 우선적으로 고려되고 있다. 무결한 지식의 목적은 정보의 참조무결성을 개선하고 지식의 재사용과 공유를 촉진하는 데 있다.

최근 디지털 도서관의 시맨틱 적용을 위한 시도가 있었다. Gdansk 공과대학 중앙도서관과 DERI의 JeromeDL, MIT의 SIMILE가 그 시도의 주역들이며 아직도 그 연구가 계속되고 있다. 이중 JeromeDL은 FOAFRealm과 MarcOnto, SIOC 등의 온톨로지를 사용하여 SDL을 넘어 이용자의 보다 광범위한 참여가 가능한 사회적 SSDL을 구현하고 있다.

국외의 사례와는 대조적으로 국내의 디지털

도서관은 아직 정보검색·탐색 기능을 제공하는 정보시스템의 기본적 역할을 수행하는데 머물고 있다. TopicMaps를 사용하여 지식인에 대한 정보를 제공하는 동대문구정보화도서관의 시맨틱 적용 시도가 있었지만 일부 시스템에 적용된 사례로 아직 국내에서는 시맨틱의 디지털 도서관 적용에 대한 연구가 미진한 것으로 보인다.

이에 본 연구에서는 국내에서의 시맨틱 기술이 디지털 도서관에 적용되어야 할 필요가 있음을 인식하고 앞서 살펴본 SSDL에 요구되는 기능적, 구조적 핵심 사항들을 중심으로 그림 6과 같은 SSDL의 아키텍처 모델을 설계하였다. 본 아키텍처 모델은 시스템 구축 전략, 시스템의 구조, 핵심 구성요소, 적용 기술 등을 보여준다. 본 연구에서 설계한 SSDL의 아키텍처 모델은 5개의 수평 층과 2개의 수직 요소로 구성되며 각 기능을 요약하면 다음과 같다.



(그림 6) SSDL 구조

7.2 SSDL의 자원, 메타데이터 및 온톨로지 식별체계 등록기

디지털도서관의 자원을 유기적으로 연결하고 지속적인 서비스를 유지하기 위해서는 모든 자원에 고유의 식별자를 부과하여서 저장장소의 위치의 변경에 관계없이 정보자원에 접근할 수 있는 체계를 마련해야 하며, 디지털도서관에서 사용되는 있는 모든 메타데이터 요소에 고유의 식별자를 부과하여 디지털도서관 시스템 간의 통합검색이 완벽하게 이뤄져야 하고, 주요 인물, 사건, 주제어에도 식별자를 부과하여 이름에 의거한 모호성을 철저히 배제시킬 시스템을 갖추어야 한다.

• 정보자원 식별자 등록

정보자원에 식별자를 부착시켜서 운용하는 체계는 DOI(Digital Object Identifier)와 PURL(Persistent URL)이 선구적인 사례이고 한국에서는 한국정보사회진흥원(NIA)이 지원하는 UCI 사업이 가장 좋은 사례이다. Handle을 활용한 통합 SSDL을 위한 자체 Resoultion 서버를 운용할 수도 있고, NIA의 서비스를 이용할 수도 있겠다. 중요한 것은 확고한 정보자원 지속적 접근을 허용하는 체계를 갖추어야 SSDL이 성공할 수 있다는 것이다.

• 메타데이터 요소 등록

인터넷의 활성화로 말미암아 다양한 메타데이터 스키마와 애플리케이션 프로파일이 생성되었다. 각 DL에서 사용하는 메타데이터 요소들 간의 상호운용성이 보장되지 않으면 진정한 통합 검색은 불가하다. 메타데이터 요소명으로

통합해서는 안 되고 각 요소에 대한 정의는 상세하게 모호성이 없도록 이뤄져야 하며, 이렇게 정의된 각 요소에 고유의 식별자를 부과함으로써 메타데이터의 통합은 요소의 식별자에 근거하여 이뤄져야 한다. 각 메타데이터 요소에 대한 정의는 ISO 11179에 근거한 더블링크어와 NSDL 레지스트리가 가장 좋은 모범사례이다. 이와 유사한 메타데이터 레지스트리를 구축하여 모든 DL 메타데이터가 체계적으로 통합될 수 있는 토대를 마련해야 한다.

• 온톨로지 인스턴스 등록

자원과 메타데이터 요소에 대한 식별자 뿐 아니라 주요 인물, 사건, 개념들에 고유의 식별자를 부과함으로써 DL 의미를 구축할 필요가 있다. 이렇듯, 등록된 것들을 웹상에 공표하면 유익한 공적주제(Published Subjects) 시스템이 되는 것이다. 이렇게 공적주제에 대한 기반 구조를 마련하면 모든 자원을 주제 중심으로 Collocation할 수 있는 체계를 확립할 수 있다. 예를 들어, 모든 주요 인물에 고유의 식별자가 부과되면 어떤 기관에서 그 인물에 대해서 기술하든지 실수 없이 모든 정보가 결합(merging)될 수 있는 체계를 마련할 수 있는 것이다. 사람은 그 인물에 대한 설명을 읽고서 자신이 기술하고자 하는 인물과 같은가를 확인하고 같다면 이미 할당된 그 인물에 대한 식별자를 그대로 차용해서 사용함으로써 식별자 기반으로 모든 정보가 통합될 수 있는 체계의 초석을 마련하는 것이다. 한 이름이 여러 의미를 지닐 수 있기 때문에 이름은 좋은 식별자가 될 수 없다. 모든 온톨로지 인스턴스에도 식별자를 부과함으로써 메타데이터 요소에 근거한 정보의 통합 능

력을 한층 격상시킬 수 있다. 똑같은 논리가 주요 사건과 주제에도 적용될 수 있다. 이와 같은 공적주제는 Topic Map과 연계하여 ISO에서 시작된 URI를 이용하여 의미적 식별체계를 구축하는 방법이다. 이것은 국립중앙도서관 웹 아카이빙 OASIS의 후원으로 지속적으로 개발되고 있으며, W3C의 시맨틱웹 Activity 그룹으로부터 많은 관심을 받고 있다. W3C에서도 이러한 주제식별자를 RDF와 OWL 기반 온톨로지 생성에 적용할 수 있는 것으로 보고 있다.

• 텍사노미 등록

도서관에서는 이미 개발된 분류체계와 시소러스의 자원이 풍부하다. 이런 자원을 활용하지 않는 것은 SSDL에서도 큰 낭비를 초래할 것이다. 시소러스를 기계가독형으로 표현할 수 있는 SKOS(Simple Knowledge Organization System)가 개발된 것은 크게 환영할 일이다. SSDL에서는 기존의 분류체계와 시소러스를 SKOS 기반으로 변환하여서 한 곳에 집약시킨다면 보다 원활한 주제적 접근을 도와 줄 것이다. 여기서 추구하는 것은 단순 어휘에 고유식별자를 부착시켜서 등록하는 수준을 넘어서 통제어휘들 간의 상·하 관계를 등록할 수 있는 기반을 마련하는 것이다. SKOS 기반으로 구축된 NSDL(미국 과학기술디지털도서관)에서 개발한 레지스트리와 유사한 기능을 갖출 필요가 있다.

SKOS는 기계가 읽을 수 있는 방법으로 지식 결합시스템을 표현하기 위한 간단하지만 강력한 프레임워크를 제공할 목적으로 개발된 것이다. SKOS 코어는 개념 스키마의 기본 구조와 내용을 표현하는 모델을 제공한다. '개념 스

키마'는 여기서 개념의 집합으로 정의되며, 임의로 개념들 간 의미 관계에 대한 구조를 표현한다. 시소러스, 분류 스키마, 주제명 표목표, 분류(taxonomies), folksonomies(태깅, 꼬리표달기), 메타데이터 구성요소 그리고 통제 어휘의 다른 유형은 모두 개념 스키마의 예이다. 개념 스키마에는 용어사전과 전문용어사전에도 포함되어 있다.

개념 스키마(통제 어휘)가 이와 유사한 구조를 가진다면, RDF를 이용하는 것은 데이터를 시맨틱웹 어플리케이션이 다른 RDF 데이터에 연결하거나 또는 합치는 것을 허용할 것이다. 실제로, 이것은 분산화 방법으로 데이터자원이 웹 전반으로 분산될 수 있지만, 새롭고 예기치 못한 방법으로 여전히 어플리케이션에 의해서 유효하게 구성되고 통합된다는 것을 의미한다. SKOS 코어는 RDF의 어플리케이션이다. 현 레지스트리가 SKOS 구조 기반의 개념 스키마 기능을 포함하여 확장된다면 분산된 자원의 관리가 더욱 효율적으로 이루어지고 다양한 형태의 자원의 관리가 가능할 것이며 의미론적 식별이 가능하게 될 것이다.

• 온톨로지 스키마 등록

온톨로지 스키마 레지스트리는 국가적 DL 의미망을 위한 식별자 등록의 가장 복잡한 단계이다. 여기에서는 토픽맵이나 RDF/OWL 기반으로 생성된 모든 온톨로지를 다양한 형태로 등록할 수 있는 기능을 제공해야 한다. 무엇보다도 현재까지 개발된 모든 온톨로지 스키마를 전제로 등록할 수 있어야 하고, 또한 개념 단위별도 등록할 수 있어야 한다. XML로 표현된 토픽맵과 RDF/OWL 온톨로지를 그대로 등록

함으로써 다양한 기계가독형 온톨로지 스키마를 다른 사람들이 사용할 수 있는 기반을 제공해야 한다. 물론 이 마지막 단계의 온톨로지를 정의할 때 등록된 어휘 PSI와 기계가독형 텍스트 노미를 재활용할 수 있다.

7.3 SSDL의 보관/저장층

최하위층은 SSDL 시스템 및 도서관 운영과 관계된 정보, 디지털 자원, 기존 도서관 정보, 이용자의 참여로 생산된 정보, 이용자 프로파일 등 SSDL 운영에서 생산되는 모든 정보를 보관하고 저장하는 기능을 수행한다. 이 계층에서는 상위 계층인 관리층의 데이터 관리 개념들과의 유기적인 연계를 위해 메타데이터를 포함한다. 특별히 디지털 자원 및 기존 도서관 정보와 이용자 참여로 생산된 정보는 FRBR 개념들을 적용하여 변환된 메타데이터를 함께 갖는다. 보관/저장층의 역할은 디지털 도서관의 이용과 운영을 위한 정보와 데이터를 저장하고 이들 정보와 데이터 간의 관계를 설명하는 메타데이터를 생성하여 관리층에서의 유의미한 관리를 가능하게 하는 기반을 마련하는 것이다.

7.4 SSDL의 관리층

관리층은 최하위층의 데이터를 온톨로지화하여 관리하는 기능을 수행한다. 시스템 운영 및 도서관 운영 관련 데이터는 RDF/OWL을 통해 지식베이스화 되고 업무의 효율성과 이용자 참여에 따른 피드백 순환의 효율성을 제고하는 기반이 된다. 디지털 자원, 기존 도서관 자원 및 이용자 참여로 생산된 자원은 토픽맵과 FRBR,

FOAFRealm을 통해 유의미한 관계가 설정되고 상위층인 매개층에 정보를 전달한다. 특히 이용자 참여로 생산된 자원과 디지털, 기존 도서관 자원의 생산 및 보존 활동에 관련된 인물에 대한 모든 정보는 FOAFRealm을 통해 의미적으로 관련성이 부여되고 관리된다. 또한 이용자 프로파일과 저작 생산 활동에 관계된 모든 인물은 FOAF를 통해 관련성이 부여되고 이는 직접적으로 매개층의 식별 확인 장치와 연결되어 최신의 인물정보를 획득할 수 있게 된다. 관리층은 동적으로 레지스트리에서 정보를 취하기도 하고 레지스트리 정보를 갱신할 수도 있다.

디지털, 기존 도서관 자원은 보다 다양한 접근점을 제공하기 위해 기존 분류법 및 이용자 참여 분류를 통해 분류되고 기존의 시소러스를 시소러스 기반 논리적 관계 설정 언어인 SKOS로 변환하여 관리되고 이용된다. 관리층에서의 데이터의 온톨로지 기반 관리는 이용자의 정보 이용과 도서관 운영에 의미론적으로 유용하며 효율적인 정보를 제공하고 매개층을 통해 유의미한 정보를 이용자에게 전달할 수 있는 기반을 마련하는 역할을 한다.

7.5 SSDL의 매개층

매개층에서는 표현층의 이용자 인터페이스를 통해 입력 받은 질의를 처리하고 연계된 외부 기관의 자원을 입수·처리하는 기능을 한다. RDF 기반인 FOAF와 FOAFRealm의 처리와 전문(full text) 검색을 위하여 RDF 기반 질의처리 언어인 SPARQL을 사용할 수 있고 표현층과 관리층에서 사용된 토픽맵의 처리를 위하여 토픽맵기반 질의처리 언어인 토픽맵질

의어(TMQL 또는 tolog)를 사용하여 의미기반 검색을 가능케 한다.

이메일을 식별자로 사용하는 FOAF의 약점을 위키피디아와 연동시켜 인물에 대한 식별을 보완하고 강화시킨다. 또한 위키피디아에 등록되지 않은 인물에 대한 정보는 FRBR와 토픽맵의 연동을 통해 온톨로지 레지스트리에 등록토록 한다. 외부 정보는 타 기관, 타 커뮤니티와 의미적 연계를 가능케 하는 SIOC를 활용하고 위키피디아와 연계하여 보다 의미적이고 효율적인 정보 생산을 도모한다. 매개층의 중심 기능은 내부적 요구를 수용하여 내·외부 정보를 검색, 탐색할 수 있는 기반을 마련하는 것이다. 매개층은 의미 관계가 정의된 정보를 효과적으로 검색하여 이용자에게 전달하기 위해 반드시 요구되는 기능을 수행한다.

7.6 SSDL의 표현층

표현층은 시스템 자체로는 최상위층으로 이용자와의 상호교류를 담당한다. 이용자에게 SSDL의 정보를 검색, 브라우징, 열람 할 수 있도록 인터페이스를 제공하는 것이다. 표현층은 개념적 표현과 기술적 표현으로 나뉘어 개념적 표현 부분에서는 생산자 간의 사회적 관계를 나타내는 Social Network 장치와 이용자의 관심도, 순위, 선택 수 등 평가 피드백이 반영된 SIA 기반 정보구성을 통해 이용자에게 정보를 의미적으로 필터링하여 제공한다.

기술적 표현 부분에서는 유의미한 관계를 기반으로 정보를 표현하는 토픽맵과 토픽맵으로 변형된 FRBR 기반 서지 정보를 표시하여 의미적으로 유효한 정보를 표현하고 이용자 서버

를 사용하여 페이지 변환을 줄이는 Ajax를 적용하여 이용자로부터 질의를 받고 정보에 대한 피드백을 받는다. 표현층은 기존 디지털 도서관의 이용자 인터페이스가 수행한 단순한 질의 응답 기능을 넘어 정보의 의미론적 구성하고 의미관계 제시, 이용자 평가 수용 등 이용자의 내재된 요구에 응답하고 수용하는 상호 보완적인 이용자 인터페이스로써 그 의미가 크다.

7.7 SSDL의 이용자 참여층

이용자 참여층은 개인화 서비스, 태그 클라우드, 블로그, 위키, 협력적 필터링 등 도서관 시스템 운영과 정보 자원 관리에 사용자의 의견과 요구가 반영될 수 있는 인터페이스와 공간을 지원하는 기능을 한다. 개인화 서비스는 최하위층의 이용자 프로파일과 연계되어 FOAF를 통해 의미적으로 관리되고 태그 클라우드는 관리층의 폭소노미에 반영된다. 블로그와 위키를 통해 생산된 정보 역시 최하위층의 이용자 참여로 생산된 정보에 저장되어 토픽맵, FRBR로 관계가 정의되고 매개층을 거쳐 표현층에서 다시 이용자에게 제공된다.

SSDL의 주요한 이슈인 지식의 재생산과 이용자 중심 서비스 개발은 협력적 지성활동, 이용자 피드백과 같은 이용자의 참여가 필수적이므로 이용자 참여 층은 SSDL이 표방하는 목적에 가장 중요한 요소이다. 디지털 도서관과 차별성을 주는 SSDL의 핵심이다.

7.8 SSDL의 보안층

보안은 모든 자원이 디지털화 되어 있으며

개방된 SSDL에서 간과할 수 없는 요소이다. 네트워크 시스템에서의 보안 유지 방법은 프로토콜 보안과 같은 원천적인 방법에서부터 키보드 해킹 방지, 디지털 서명키 등 보안 소프트웨어를 사용한 시스템적인 방법 등 다양하나 OAI-PMH, Z39.50, REST 등 개방형 시스템에서 사용되고 있는 프로토콜을 사용하여 시스템 개방을 방해하지 않으며 보안을 유지하는 것이 바람직하다. SSDL의 보안을 유지하기 위한 방법은 전문분야에서의 보다 많은 논의가 필요하다.

이상의 SSDL 아키텍처의 각 계층은 최하위층인 보관/저장층이 그 기반이 되어 이용자 참여층에 이르러 시스템을 완성하고 있다. 본 연구에서 설계한 SSDL은 이용자 참여층의 피드백으로 보관/저장층의 데이터가 갱신되므로 순환 구조라 할 수 있다. 또한 RDF/OWL과 Topic Maps 등을 이용하여 데이터 간의 의미관계를 표현하여 의미 기반 서비스를 제공할 수 있도록 하고 SIOC를 사용하여 외부 정보시스템의 정보를 공유할 수 있는 체계를 마련하는 것이다.

8. 결론

SSDL은 아직 우리에게 생소한 시스템이다. 이와 관련한 국내의 연구가 아직 미진한 상태이며 외국의 경우에도 이제 막 그 첫발을 떼고 있는 상태이다. 그러나 전통적 서지 목록 체계와 문헌정보 시스템이 FRBR, 온톨로지와 같은 의미기반 시스템으로 전환되고 있으며 웹2.0과 시맨틱 웹의 급속한 발전으로 도서관이 라이브러리2.0으로의 진화에 직면해 있다. 이러한 일련의 변화를 총체적으로 담은 도서관이

SSDL이다. 본 연구에서는 이러한 SSDL에 요구되는 핵심사항들을 살펴보고 이를 분석하여 SSDL의 모델을 설계하였다.

SSDL은 메타데이터와 온톨로지를 기반으로 정보와 지식을 표현하며 이를 이용자에게 제공할 뿐만 아니라 기존의 단순하게 연결되어 있었던 기관과 기관, 시스템과 시스템을 의미 기반 처리 시스템을 통해 보다 의미적으로 연결 될 수 있도록 한다. 또한 급속도로 확산된 웹2.0은 이용자를 더 이상 수용자의 위치에 머무르지 않게 하여 이용자 스스로가 시스템에 참여하고자하는 의지를 북돋워 주었다. 이를 통해 SSDL은 이용자 참여가 가능한 보다 적극적인 이용자 중심 시스템으로 진화하였다. 즉, SSDL은 보다 적극적인 이용자 중심, 보다 유효한 의미 중심, 보다 다양한 상호교류 중심의 시스템인 것이다.

본 연구는 이러한 SSDL을 구현하기 위하여 보관/저장층, 관리층, 매개층, 표현층, 이용자 참여 층의 5개 층으로 구성되는 아키텍처 모델을 제시하였다. 보관/저장층은 메타데이터가 포함된 디지털 자원의 저장고로써 의미적 상호 호환성과 지식 추출의 기반이 되는 계층이며 관리층은 이러한 보관/저장층의 의미적 데이터의 관계를 정의하여 의미기반 서비스를 가능토록 지원한다. 매개층은 이용자와 시스템의 자원을 연결하여 이용이 가능토록 할뿐만 아니라 외부 시스템과의 커뮤니케이션을 가능토록 한다. 표현층은 정보를 체계적으로 표현하여 이용자의 정보 이용·활용을 가능케 하며 이를 통해 발생한 새로운 지식은 이용자 참여 층을 통해 시스템에 반영되어 새로운 서비스, 새로운 정보/지식이 생성될 수 있도록 한다.

본 연구에서 제안한 SSDL의 아키텍처 모델은 최근의 정보기술의 수준을 고려한다면 이와 유사한 시스템들이 일반화되고 있으므로 시스템의 구현에는 어려움이 없을 것으로 판단된다. 아직 SSDL에 대한 연구가 미진한 상태이지만

라이브러리 2.0 등 관련 연구와 더불어 새로운 문헌정보시스템에 대한 요구가 확산되고 있는 만큼 본 연구를 기초로 SSDL 구현에 관한 연구가 활성화되기를 기대한다.

참 고 문 헌

- 오삼균, 2002, 시맨틱웹 기술과 활용방안, 『정보관리학회지』, 19(4): 297-319.
- 유정연, 이규철, 2004, TM-S: 서비스를 위한 Topic Maps 기반의 온톨로지 및 질의 언어 설계, 『한국정보과학회 학술발표논문집』, 2004년도 봄(B): 109-111.
- 한성국, 이현실, 2007, 시맨틱 라이브러리를 위한 아키텍처 참조 모델, 『정보관리학회지』, 24(1): 75-101
- Barbara Tillett, 2004, "what is FRBR?: A Conceptual Model for the Bibliographic Universe", Library of Congress Cataloging Distribution Service, [online]. [cited 2007.11.25] <<http://www.loc.gov/cds/downloads/FRBR.PDF>>
- Gruber, T. 2001. "What is an Ontology?", [online]. [cited 2007.11.23] <<http://www.ksl.stanford.edu/kst/what-is-an-ontology.html>>
- Graves, constabaris, & Brickley, 2007, "FOAF: Connecting People on the semantic Web", KERIS 종합목록 주요소식, 14: 1-5 [online]. [cited 2007.11.26] <http://www.riss4u.net/webzine/final_webzine/200707/document07_07_01.pdf>
- IFLA Study Group on the functional Requirements for Bibliographic Records Final Report, 1998, "Functional Requirements for Bibliographic Records Final Report." [online]. [cited 2007.11.23] <<http://www.ifla.org/VII/s13/frbr/frbr.pdf>>
- Jorge Cardoso, Amit Sheth, 2006, "THE SEMANTIC WEB AND IT'S APPLICATIONS", Semantic Web Services, Processes and Applications, 3: 3-33
- MacKenzie Smith, 2005, "Project SIMILE: Semantic Interoperability of Metadata and Information in unLike Environments", Massachusetts Institute of Technology [online] [cited 2007.12.22] <http://simile.mit.edu/funding/mellon_2005.pdf>
- S. R. Kruk, T. Woroniecki, A. Gzella, M.

- Dabrowski, and B. McDaniel, 2007, "The anatomy of a Social Semantic Digital Library", [online]. [cited 2007.11.27]
 <<http://www.marcont.org/marcont/pdf/jeromedl-ecdl2007.pdf>>
- S. R. Kruk, S. Decker, 2005, "FOAFRealm: Making Social Collaborative Filtering Real", [online]. [cited 2007.11.27]
 <http://www.marcont.org/marcont/pdf/eswc2005_foafrealm.pdf>
- S. R. Kruk, B. Haslhofer, P. Piotrowski, A. Westerski, T. Woroniecki, 2006, "Role of Ontologies in Semantic Digital Libraries", [online]. NKOS Workshop, [cited 2007.11.23]
 <<http://www.ukoln.ac.uk/nkos/nkos2006/presentations/kruk.pdf>>
- Pepper, Steve. 2006. Toward Semantic Superhighway: A Manifesto for Published Subjects. [online]. [retrieved 2007.12.24]
 <<http://www.ibiblio.org/hhalpin/irw2006/spepper.html>>
- 웹자원**
- BRICKS Community.
 <<http://www.brickscmmunity.org/>>
- Findability.org.
 <<http://www.findability.org/archives/000166.php>>
- FOAF Project Homepage.
 <<http://www.foaf-project.org/>>
- Library 2.0 memo map.
 <http://www.flickr.com/photo__zoom.gne?id=113222147&size=m>
- SIOC Project Homepage.
 <<http://sioc-project.org/>>
- SIOC Core Ontology Specification.
 <<http://rdfs.org/sioc/spec/>>
- Social Information Architecture workshop
 <<http://www.slideshare.net/gsmith/social-information-architecture-workshop/>>
- SIMILE Homepage
 <<http://simile.mit.edu/>>