

# 태그 조직화를 위한 소셜 메타데이터 프레임워크 구축\*

## Construction of Social Metadata Framework for Organizing Social Tags

이 승 민 (Seungmin Lee)\*\*

### 목 차

- |                             |                            |
|-----------------------------|----------------------------|
| 1. 서 론                      | 4. 소셜 인포메이션 아키텍처 구축        |
| 2. 소셜 메타데이터의 조직화            | 4.1 분석대상의 선정               |
| 2.1 소셜 메타데이터의 의미            | 4.2 소셜 메타데이터 분석 결과         |
| 2.2 소셜 메타데이터 조직화를 위한 접근방법   | 4.3 소셜 인포메이션 아키텍처 프레임워크 구축 |
| 3. 소셜 메타데이터의 조직화 방안         | 5. 결 론                     |
| 3.1 기술사항으로서의 소셜 메타데이터       |                            |
| 3.2 조직화 구조로서의 소셜 인포메이션 아키텍처 |                            |

### 초 록

소셜 메타데이터는 이용자의 자발적인 참여를 통해 정보자원에 대한 풍부한 기술사항을 생성한다는 장점을 지니고 있지만, 조직화 된 구조를 적용하기 어렵다는 특성으로 인해 여러 가지 한계 또한 보이고 있다. 이에 본 연구에서는 소셜 메타데이터의 유형 가운데 태그를 중심으로 LibraryThing에서 활용되고 있는 태그의 의미를 분석하고, 이를 기반으로 기술의 대상이 되는 정보자원의 서지적 카테고리를 제공함으로써 태그의 의미적 조직화를 위한 대안적인 방안을 제안하였다. 서지적 카테고리를 구조적으로 제공하고 이를 통해 태그 부여단계에서 태그의 의미적 조직화를 유도하기 위해 소셜 인포메이션 아키텍처를 적용하여 태그의 조직화를 위한 소셜 메타데이터 프레임워크를 구축하였다. 이는 소셜 메타데이터의 조직화를 위한 개념적인 기반을 제공함으로써 향후 태그를 자동화 된 방식으로 조직하는데 활용할 수 있는 의미적 기반을 마련해 줄 것으로 기대된다.

### ABSTRACT

Although social metadata has strengths in creating amount of user-contributed resource descriptions, its function is limited because of its non-systematic characteristics. This research proposed an alternative approach to semantic organization of social metadata. It analyzed the semantics of tags created in LibraryThing in order to provide bibliographic categories for describing information resources. Social information Architecture is adopted in generating the bibliographic categories so that social metadata framework can be constructed. This framework can provide the conceptual foundations for semantically organizing social metadata and is expected to be applied to the existing approaches to automatically organize social metadata.

키워드: 소셜 메타데이터, 소셜 인포메이션 아키텍처, 태그, 서지적 카테고리, 기술사항, LibraryThing  
Social Metadata, Social Information Architecture, Tag, Bibliographic Category,  
Resource Description, LibraryThing

- \* 본 연구는 숙명여자대학교 교내연구비지원에 의해 수행되었음(과제번호 1-1403-0150).  
본 연구는 2014년도 한국문헌정보학회 추계 학술발표대회에서 발표한 '소셜 인포메이션 아키텍처에 기반한 소셜 메타데이터 프레임워크 구축'을 수정·보완한 것임.
- \*\* 숙명여자대학교 문과대학 문헌정보학과 조교수(ableman@sookmyung.ac.kr)  
논문접수일자: 2014년 10월 22일 최초심사일자: 2014년 11월 3일 게재확정일자: 2014년 11월 6일  
한국문헌정보학회지, 48(4): 91-113, 2014. [http://dx.doi.org/10.4275/KSLIS.2014.48.4.091]

## 1. 서론

웹 2.0으로 진화된 웹은 웹 상의 공간이라는 개념을 넘어 많은 사람들이 다양한 목적으로 정보를 생산하고 이에 접근하며 공유하는 지식 플랫폼으로 자리잡고 있다. 이는 정보를 수동적으로 제공받는 공간에서 웹의 이용자들이 직접 정보를 생성하고, 공유하고, 협동작업을 수행할 수 있는 이용자 중심의 공간으로의 진화를 의미한다. 이를 통해, 현재의 웹은 정보를 일방적으로 제공하고 표현하는 대신에 사람들 사이의 연결을 통해 이용자 상호간에 정보를 공유하고 정보를 생산할 수 있는 새로운 정보 환경을 구축해 오고 있다.

이러한 웹 2.0으로의 진화는 이용자의 정보활동에 있어서의 패러다임의 변화를 이끌어 내고 있다. 웹 2.0 환경에서 이용자는 정보활용을 위한 플랫폼을 공유함으로써 다양한 이용자들이 생성한 콘텐츠를 이용하는 정보의 소비자일 뿐만 아니라 콘텐츠의 생산자로서의 역할 또한 수행하고 있다. 이러한 이용자 중심의 정보활동으로의 진화에는 정보의 생산과 소비에 이용자가 직접 참여하는 방식이 큰 역할을 담당하고 있으며, 이는 소셜 메타데이터(social metadata)라는 새로운 형태의 정보 기술사항을 이끌어 내고 있다.

소셜 메타데이터는 정보의 검색, 웹 정보자원의 기술 등에 이용자가 직접 참여함으로써 이용자들의 정보활동을 지원하는데 있어서 여러 가지 장점을 지니고 있지만, 모호한 용어의 사용, 비체계적인 구조, 개인에게 특화된 정보의 생성 등 여러 문제를 야기시키기도 한다. 이러한 이유로 소셜 메타데이터의 효용성에 대해

서는 현재까지도 계속해서 논란이 지속되고 있지만, 이용자가 직접 참여하는 소셜 메타데이터는 현재의 정보 검색 및 활용에 있어서 커다란 영향을 미치고 있는 것이 사실이다.

이러한 측면에서, 소셜 메타데이터가 지닌 장점을 극대화하고 단점을 보완하기 위해 많은 연구가 수행되어 왔다. 이들 연구의 대부분은 이용자가 생성한 소셜 메타데이터를 체계적인 방식으로 조직화하는데 중점을 두고 있다. 특히, 태그(tag) 등과 같은 열거적인 소셜 메타데이터의 의미를 분석하고 이를 구조화하여 검색의 효율성을 높이기 위한 방안이 다양한 측면에서 제시되어 왔다. 하지만, 대부분의 연구는 이용자가 부여한 태그 자체를 독립적인 단위로 고려하고 있기 때문에, 정보자원에 대한 풍부한 기술사항의 제공이라는 태그의 본질적인 기능보다는 태그 자체를 검색의 접근점으로 한정하는 검색적인 측면에 치중하고 있다. 또한, 이용자가 부여한 태그의 의미적 모호성, 비밀관성, 비통제성 등의 문제를 해결하지 않은 상태에서 구조화 된 체계를 적용하고 있기 때문에, 현재까지 제시되어 온 소셜 메타데이터의 조직화 방안은 큰 효과를 거두지 못하고 있다.

이러한 문제를 해결하고 소셜 메타데이터를 보다 효과적으로 활용하기 위해, 본 연구에서는 각각의 태그를 독립적인 단위로 고려하는 것이 아닌 정보자원에 대한 전체적인 기술사항으로써의 소셜 메타데이터에 중점을 두고 이를 조직화하기 위한 방안을 제안하고자 한다. 본 연구에서 제안하는 접근방법은 이용자가 태그를 생성하는 단계에서 정보자원의 기술을 위한 서지적 컨텍스트를 제공하고, 이를 통해 구조화 된 체계를 적용할 수 있는 기반을 마련하는 것이다.

이를 위해, 본 연구에서는 기술의 대상이 되는 정보자원의 서지적 컨텍스트와 정보자원의 기술사항으로써의 소셜 메타데이터 사이를 의미적으로 연결시켜 주기 위해 소셜 인포메이션 아키텍처(social information architecture)를 적용하여 소셜 메타데이터의 조직화를 위한 개념적인 프레임워크인 소셜 메타데이터 프레임워크(social metadata framework)를 구축하고자 한다. 이를 통해, 이용자가 자발적으로 참여하여 생성한 소셜 메타데이터를 웹 상의 정보자원에 대한 기술사항과 결합시킴으로써 보다 풍부한 정보자원의 기술을 유도할 수 있는 기반을 마련하고자 한다.

## 2. 소셜 메타데이터의 조직화

### 2.1 소셜 메타데이터의 의미

소셜 미디어의 대중화와 함께, 현재의 웹은 정보의 생성·관리·소비에 이용자가 직접 참여할 수 있는 웹 2.0 환경으로 진화해 왔다. 웹 2.0 이전의 웹이 일방적인 정보 제공의 형태였다면, 웹 2.0에서는 일반 이용자들이 웹 환경에 자발적으로 참여하여 다양한 정보를 생성하고 이를 공유함으로써 웹 상에서 이전과는 다른 다양한 정보활동을 수행하고 있다. 이용자들은 자발적인 참여를 통하여 웹 상에 새로운 정보를 생성하기도 하지만, 기존의 정보자원에 이용자 중심의 새로운 기술사항을 추가함으로써 정보의 활용을 지원하기도 한다. 정보자원에 새로운 정보를 추가하는데 있어서는 여러 가지 방식이 적용되고 있지만, 이 가운데 가장 일반적

인 방법 중의 하나로 소셜 메타데이터를 들 수 있다(강신재 and Ding 2008).

소셜 메타데이터는 웹 상에 존재하는 정보자원에 대한 기술사항으로써의 콘텐츠를 의미한다. 전통적으로 정보자원은 전문가 혹은 정보자원의 저자 등에 의해서 메타데이터의 형식으로 기술되고 조직되어 왔다. 하지만, 웹 2.0 환경으로의 진화와 함께, 이용자들은 태그, 평가, 주석 등의 형식으로 정보자원에 대한 기술사항을 직접 생성할 수 있게 되었으며, 이러한 기술사항으로써의 이용자 참여 콘텐츠는 일반적으로 소셜 메타데이터로 인식되고 있다. 이외에도, 웹 2.0 상에서의 소셜 메타데이터는 웹 상에 존재하는 다양한 정보자원에 수록된 콘텐츠와 관련된 키워드를 부여하는 행위로 설명되기도 한다(강신재 and Ding 2008).

현재 소셜 메타데이터는 ‘이용자 참여 콘텐츠(user-contributed contents)’ 혹은 ‘이용자 참여 주석(user-contributed annotations)’ 등 여러 가지 명칭으로 사용되고 있으며, 대부분의 경우 태그의 형식을 취하고 있다. 일반적으로 태그는 “이용자의 마음속에 존재하는 개념과 웹 상의 정보자원 사이의 관계를 형성해 주는 모든 단어”로 정의할 수 있다(Guy and Tonkin 2006). 정보자원에 태그를 부여하는 정보활동은 ‘협동적 태깅(collaborative tagging)’ 혹은 ‘소셜 태깅(social tagging)’으로 불리우고 있으며, 이는 이용자들이 의한 메타데이터의 생성이라는 측면에서 소셜 메타데이터의 범주에 포함되고 있다(Kakali and Papatheodorou 2010).

초창기의 소셜 메타데이터는 주로 태그의 형태로 보급되었으며, Delicio.us, Flickr 등과 같은 사진 중심의 웹사이트, YouTube 등과 같은

동영상 중심의 웹사이트에서 널리 적용되어 왔다. 이후 Amazon.com 등과 같은 상업적 웹사이트에서도 소셜 미디어의 특성을 접목시키면서 태그의 적용이 활발하게 이루어지게 되었고, 특정 정보자원에 대한 이용자 코멘트, 평가, 추천 등의 방식으로 소셜 메타데이터의 적용이 확대되고 있다. 현재는 도서관계에서도 소셜 메타데이터를 적용하여 정보자원에 대한 풍부한 기술사항을 제공할 수 있는 기반을 마련하고 있다. 이러한 측면에서, 소셜 메타데이터는 메타데이터의 생성에 이용자가 직접 참여하는 메타데이터의 확장된 개념으로 다루어지기도 한다.

하지만, 소셜 메타데이터는 전통적인 메타데이터와는 여러 가지 측면에서 차이를 보이고 있다. 소셜 메타데이터는 전문가에 의해 생성되는 것이 아니라 이용자의 직접적인 참여를 통해 부가되는 정보이며, 따라서 이용자 측면에서 정보자원에 대한 기술사항을 제공해 준다. 이러한 이용자 중심적인 특성으로 인해서 다른 이용자가 해당 정보자원을 검색하고 이용하는 데 도움을 주는 새로운 유형의 정보로 인식되고 있다. 이를 통해, 소셜 메타데이터는 이용자들 사이의 커뮤니케이션을 촉진하고 커뮤니티를 구축하거나 기관의 서비스를 확장하는데 활용되기도 한다 (OCLC 2012). 이외에도, 이용자가 부여한 콘텐츠를 이용하여 특정 정보자원과 관련된 내용의 조직이나 활용을 지원해 주는 기능을 수행하기도 한다(Alemu, Ross and Chandler 2012).

이와 같이, 소셜 메타데이터는 정보자원의 기술, 검색 등에 있어서 여러 가지 장점을 지니고 있다. 이용자들의 자발적인 참여를 통해서

이루어지기 때문에 집단지성을 효과적으로 적용할 수 있으며, 전문가에 의해 생성된 메타데이터를 다수의 이용자들이 보완하고 추가함으로써 풍부한 기술사항을 제공해 준다는 장점이 있다. 또한, 메타데이터의 생성에 별다른 비용이 소요되지 않으며, 이용자들의 자발적이고 즉시적인 참여를 통해 이루어지기 때문에 최신의 기술사항을 유지할 수 있다는 장점을 지니고 있다. 이와 함께, 소셜 메타데이터는 전통적인 메타데이터가 지닌 복잡한 구문규칙 등에 영향을 받지 않는 자유로운 형식의 메타데이터로 설명할 수 있으며, 콘텐츠 중심적인 메타데이터를 대규모로 생성할 수 있다.

이러한 장점에도 불구하고, 소셜 메타데이터는 여러 가지 한계 또한 지니고 있다. 대표적인 문제로는 태그로 사용되는 용어의 의미적 모호성을 들 수 있다. 이용자가 개인의 관점에서 통제되지 않은 어휘를 자유롭게 선택하기 때문에, 태그로 부여된 하나의 단어가 여러 가지 의미로 사용되기도 하며, 여러 단어가 동일한 의미를 표현하기도 한다. 따라서, 동일한 정보자원을 대상으로 하더라도 태그를 부여하는 이용자의 수준이나 관점에 따라서 다른 의미의 태그가 적용될 수 있으며, 검색의 접근점으로 사용되기에는 부적절한 용어가 사용될 수 있다는 문제가 나타나고 있다. 이는 정보자원에 대한 기술사항으로써의 태그의 질을 저하시키는 원인으로 작용하고 있다. 이와 함께, 여러 단어를 연결하여 하나의 태그로 부여하는 경우 이를 어떻게 의미적으로 구분해야 하는지 등의 문제가 지속적으로 제기되고 있다. 이러한 이유로 인해서 이용자가 자율적으로 부여한 태그를 체계적으로 조직하기 어렵다는 한계를 보이고 있

으며, 이는 소셜 메타데이터로써의 태그의 효율성 저하라는 문제를 가져오고 있다.

## 2.2 소셜 메타데이터 조직화를 위한 접근방법

소셜 메타데이터는 이용자의 자발적인 참여를 통해 특정 정보자원과 관련된 풍부한 기술 사항을 제공해 주고 다양한 접근점을 생성할 수 있는 방안을 마련해 준다. 하지만, 이용자 참여의 방식은 표준화 된 방식이 아니라 이용자 개개인이 자신에게 도움이 되는 방식으로, 그리고 자신에게 의미가 있는 용어를 이용하는 방식이 일반적이다. 이러한 개인에게 특화된 방식은 소셜 메타데이터의 효율성을 저하시키는 주된 이유로 인식되고 있다. 이러한 문제를 해결하기 위해, 개인 중심의 비정형화 된 용어를 체계적인 방식으로 구조화하기 위한 여러 가지 방안이 제시되어 왔다.

이 가운데, 이용자가 부여한 다양한 태그를 자동화 된 방법으로 분류하여 검색의 접근점으로 활용하기 위한 방안이 여러 연구자들에 의해서 제안되어 왔다. Wray and Eklund(2010)는 정형화 된 개념 분석 방식(Formal Concept Analysis)을 적용한 협동적인 태깅 시스템(Collaborative Tagging Systems)을 이용하여 이용자가 부여한 태그를 의미적으로 분류하기 위한 시도를 하였다. 이 연구에서는 예술작품과 관련한 정보자원을 대상으로 하고 있으며, 태깅 시스템에서 예술작품의 범주를 구체화하고 이를 기반으로 태그를 자동화 된 방식으로 분류하는 접근방법을 제안하였다. 하지만, 이 접근방법에서는 상호 관련된 태그가 서로 다른 카테고리에 속하게 된다는 문제를 해결하지는 못하고

있다. Kakali and Papatheodorou(2010)는 폭소노미(folksonomy)의 측면에서 이용자가 부여한 소셜 태깅을 자동화 된 방식으로 분류하고자 하였다. 이 연구에서는 OPACIAL 시스템에서 추출한 태그들을 대상으로 하여 폐쇄화된 네비게이션을 이용한 검색 인터페이스를 구축하고, 이를 통해 다양한 형태의 태그를 의미적으로 구분하는 방식을 적용하고 있다. 이는 이용자가 부여한 태그의 형태론적인 측면에 중점을 두고 있으며, 이를 기반으로 태그의 의미를 분석하고 이를 주제어 색인에 적용하는 접근방법을 취하고 있다. 하지만, 이들 연구들은 이용자가 부여한 태그의 의미적 모호성을 제거하지 않은 상태에서 태그의 의미를 분석함으로써 태그를 분류한 결과의 정확성이 저하된다는 한계를 보이고 있다.

이와는 다른 측면에서, 주제어 색인으로서의 태그의 효율성을 분석하는 연구가 수행되어 왔다. Al-Khalifa and Davis(2007)는 자동화 된 색인 시스템에서 추출한 키워드와 폭소노미 사이의 관계를 분석하였다. 자동화 된 방식으로 추출된 색인과 이용자가 부여한 태그 사이의 관계를 분석한 결과, 이들 사이에는 높은 수준에서의 일치도가 나타나고 있으며, 이용자가 생성한 태그의 검색 접근점으로써의 활용도 또한 높은 것으로 나타나고 있다. 이와 유사한 측면에서, Thomas, Caudle and Schmitz(2009)는 이용자가 부여한 태그를 의미적으로 분류하기 위하여 소셜 태그를 통제어휘 및 자동화 된 색인과 비교한 결과를 제시하였다. 이 연구에서는 자동화 된 방식으로 추출한 색인과 통제어휘에 소셜 태그를 추가하면 검색 및 브라우징에 있어서의 효율성을 높일 수 있다는 결과

를 확인함으로써 의미적으로 조직화 한 태그의 효용성을 주장하고 있다. Lin et al.(2006)은 소셜 태그, 정보자원의 서명, 시소러스 사이의 상관관계를 분석하기 위해서 이용자가 부여한 태그를 자동적으로 추출한 정보자원의 서명 및 MeSH와 비교를 수행하였다. 그 결과, 11%의 태그만이 MeSH와 일치하는 것으로 나타났다. 이러한 결과는 이용자가 태그를 부여하는 목적과 MeSH에서 사용하는 통제어휘의 근본적인 목적의 차이에 기인하는 것으로 분석되었다. 이들 연구를 종합해 보면, 이용자가 부여한 태그의 의미적 모호성을 제거하지 않은 상태에서의 태그의 조직화는 상황에 따라 다른 결과가 나타난다는 것을 보여주고 있다.

소셜 메타데이터에 대한 연구는 태그 자체의 분류나 조직화 등에 대한 연구가 주를 이루고 있으나, 소셜 메타데이터가 이용자의 정보활동에 미치는 영향에 대한 연구도 지속적으로 수행되어 왔다. Golder and Huberman(2006)은 Delicio.us 웹사이트에서 태그의 사용 빈도를 측정하였는데, 태그는 개인적인 용도로 부여되는 키워드 형태가 대부분을 차지하고 있다는 것을 밝히고 있다. Bischoff et al.(2008)은 웹사이트, 음악자료, 이미지 자료 등 다양한 형태의 정보자원에 이용자들이 태그를 부여하는 행태를 분석하였다. 분석 결과, 태그는 정보자원에 새로운 정보를 부가하고 있으며, 대부분의 태그는 정보자원의 내용을 정확하고 신뢰성있게 기술해 주고 있음을 밝히고 있다. 또한, 이용자가 생성하는 대부분의 태그는 검색어로 사용될 수 있으며, 이용자들의 정보검색 행위에 부합하는 특성을 지니고 있는 것으로 나타났다. Kipp(2007)은 CiteULike 웹사이트 상에 존재

하는 정보자원을 기술하는데 있어서, 이용자, 저자, 전문가들이 사용하는 언어를 비교·분석하였는데, 이용자들이 부여하는 태그는 저자의 키워드 및 목록의 주제 기술 언어와 밀접하게 관련이 있다는 것을 보여주고 있다. 이들 연구를 종합해 보면, 이용자들은 태그를 이용하여 웹 상에서 다양한 정보활동을 수행하고 있으며, 이용자가 부여한 태그는 정보자원에 대한 새로운 정보를 추가해 주는 기능을 수행하고 있음을 알 수 있다. 반면, 정보자원에 대한 기술사항으로써의 태그는 정보자원을 검색하는데 있어서 다양한 접근점을 제공해 주는 도구로 활용될 수 있지만, 조직화되지 않고 비체계적인 방식이라는 특성은 태그의 효용성을 저하시키고 있다는 점이 공통적인 문제로 지적되고 있다.

이와 같이, 이용자가 부여한 태그를 체계적으로 조직화하기 위한 방안이 여러 연구를 통해서 제안되어 왔다. 현재까지 제안되어 온 접근방법들은 이용자가 부여한 태그들을 자동화된 방식으로 분류하는데 중점을 두고 있다. 이와 함께, 부여된 각각의 태그를 독립적인 단위로 인식하여 이를 검색의 접근점으로 활용하기 위한 방안이 제안되어 왔다. 하지만, 이용자가 부여하는 태그는 용어의 의미적 모호성을 지니고 있기 때문에, 상당수의 태그들이 의미적으로 명확하지 않으며 일반화되지 않은 개인적인 관점으로 그 사용이 국한되기도 한다(Warner and Brussee 2008). 이러한 비일관성을 해결하지 않은 상태에서 각각의 태그를 분류하는 방식은 여러 가지 문제를 일으킬 수도 있다(Mika 2007; Tonkin 2006; Voss 2006).

### 3. 소셜 메타데이터의 조직화 방안

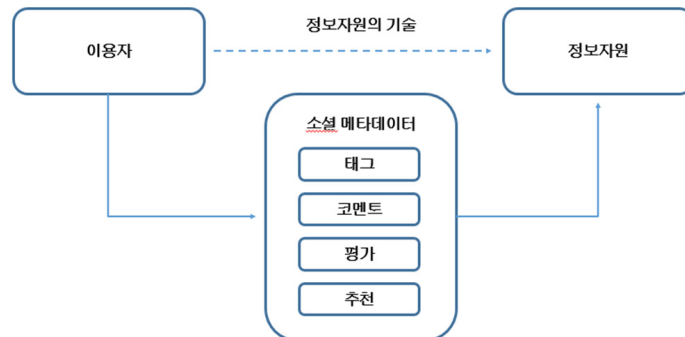
#### 3.1 기술사항으로서의 소셜 메타데이터

소셜 메타데이터는 이용자들이 직접 참여하여 자유로운 방식으로 생성하기 때문에, 태그의 부여를 위해 사전에 규정된 구조가 없으며 모든 용어들이 열거되는 형식을 지닌다. 하지만, 소셜 메타데이터는 특정 정보자원이 지니고 있는 여러 가지 측면을 이용자의 참여를 통해서 풍부하게 기술해 주고, 이를 통해 검색의 접근점을 확장하는 기능을 수행하고 있다. 이러한 이유로, 소셜 메타데이터는 전통적인 메타데이터의 범주에 포함되는 것이 일반적이다.

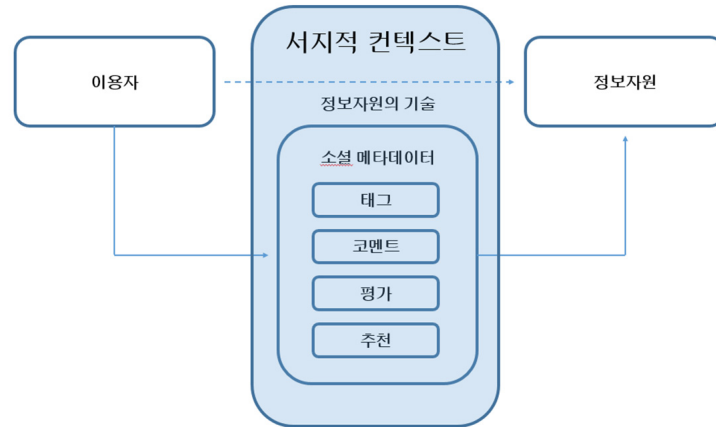
하지만, 메타데이터와 소셜 메타데이터는 조직화의 측면에서 차이를 보이고 있다. 기관 중심으로 생성되는 메타데이터는 체계화 된 구조를 지니고 있으며 정형화 된 방식으로 생성되기 때문에, 이용자들은 특정 체계 안에서 신뢰성 높고 일관성있는 접근점을 활용할 수 있다. 또한, 계층적인 구조 혹은 체계화 된 구조를 적용할 수 있기 때문에 메타데이터로 기술된 요

소의 의미가 명확하게 제공된다는 장점이 있다. 반면, 메타데이터는 전문가에 의해서 생성되기 때문에, 정보자원에 대한 기술사항이 전통적인 서지사항 중심으로 이루어진다는 한계가 있다. 이와는 달리, 소셜 메타데이터는 이용자가 자발적으로 참여하여 생성되기 때문에 이용자 관점에서 다양한 기술사항을 제공해 줄 수 있으며, 이로 인해 기관 중심의 정형화 된 기술사항을 보완하여 이용자 중심의 기술사항을 충분하게 제공해 준다는 장점이 있다. 반면, 소셜 메타데이터는 전문적 지식이 없는 이용자들이 의해서 생성되기 때문에, 정보자원에 대한 기술사항이 체계적으로 조직되지 못하고, 택소노미 혹은 계층적 구조가 효과적으로 적용되기 어렵다는 한계를 보이고 있다.

메타데이터와 소셜 메타데이터 사이의 이러한 차이로 인해서, 소셜 메타데이터를 생성하는 과정은 메타데이터의 생성과는 다른 방식으로 진행된다. 메타데이터는 사전에 규정된 구조화 된 체계에 정보자원으로부터 추출한 값을 입력하지만, 소셜 메타데이터는 오픈된 공간 안에 태그, 평가, 코멘트, 추천 등의 형태로 생성된다(〈그림 1〉 참조).



〈그림 1〉 소셜 메타데이터의 생성 방식



〈그림 2〉 소셜 메타데이터 생성을 위한 서지적 컨텍스트 부여

이와 같이, 메타데이터와 소셜 메타데이터는 각각 고유한 특성 및 장·단점을 지니고 있지만, 이들은 특정 정보자원에 대한 기술사항을 제공해 준다는 공통점을 지니고 있다. 따라서, 메타데이터의 체계화 된 구조와 소셜 메타데이터의 풍부한 기술사항이라는 두 가지 방식을 접목시키게 되면 메타데이터와 소셜 메타데이터는 상호 보완하는 기능을 수행할 수 있으며, 보다 풍부한 서지적 기술사항을 제공해 주는 도구로 사용될 수 있을 것이다. 또한, 이러한 매개적인 접근방식은 이용자의 자율성을 보장함과 동시에 조직화 된 체계를 적용할 수 있는 방식이 될 수 있을 것이다.

하지만, 메타데이터의 서지적 구조를 소셜 메타데이터의 열거적인 방식에 그대로 적용할 경우 여러 가지 문제가 발생할 수 있다. 따라서, 본 연구에서는 소셜 메타데이터에 정보자원과 관련된 서지적 컨텍스트를 설정하고, 이용자들이 소셜 메타데이터를 부여하는 단계에서 체계화 된 조직 구조를 적용하는 방식을 적용하고자 한다. 이는 정보자원에 부여된 소셜 메타데이터를 조

직화하는 방식이 아니라 이용자들이 소셜 메타데이터를 부여하는 단계에서의 조직화를 시도하는 접근방법으로, 소셜 메타데이터를 체계적으로 조직화 할 수 있는 기반을 마련해 주는 대안적인 접근방법으로 생각할 수 있다.

소셜 메타데이터를 부여하는 과정에서 서지적 컨텍스트를 설정하기 위해서는 메타데이터의 구조화 된 체계와 소셜 메타데이터를 융합할 수 있는 프레임워크의 구축이 필요하다. 이를 위해, 본 연구에서는 소셜 인포메이션 아키텍처(social information architecture)를 적용하여 기술의 대상이 되는 정보자원의 특성을 반영한 서지적 컨텍스트를 제공하고, 이를 통해 소셜 메타데이터를 의미적으로 조직하기 위한 기반을 마련하고자 한다.

### 3.2 조직화 구조로서의 소셜 인포메이션 아키텍처

이용자가 정보를 활용하는 환경에서, 이용자와 시스템 사이의 상호작용은 특정 컨텍스트 안



에서 이루어진다(Saracevic 2010). 인포메이션 아키텍처(information architecture)는 이러한 컨텍스트를 제공해 줌으로써 정보자원을 효과적으로 활용할 수 있는 기반을 마련해 주고 있다.

인포메이션 아키텍처는 여러 가지 측면에서 설명할 수 있는데, 일반적으로 인포메이션 아키텍처는 정보를 이용하는 활동의 모델이나 개념을 정립함으로써 복잡한 정보체계에 대한 일관성있는 구조를 구축하기 위한 정보의 카테고리화로 설명할 수 있다. 이는 웹사이트를 구조화하고 분류하는 기술이나 분야를 의미하며, 이용자들이 정보를 검색하고 관리할 수 있는 효과적인 방식을 제공해 준다(Paul 2014).

하지만, 웹 2.0 환경의 정착과 함께 정보환경의 소셜화가 이루어지면서, 인포메이션 아키텍처의 개념 역시 소셜화 된 방향으로 진화하고 있다(Smith 2008). 이러한 소셜화 된 인포메이션 아키텍처는 소셜 인포메이션 아키텍처(social information architecture)의 명칭으로 불리고 있다. 소셜 인포메이션 아키텍처는 아직까지는 그 기능이나 의미가 명확하게 정립되어 있지는 않다. 하지만, 소셜 인포메이션 아키텍처는 인포메이션 아키텍처의 하위 개념으로 시작되었으며, 기존의 인포메이션 아키텍처에 사회화라는 개념이 추가된 것으로 볼 수 있다(Paul 2014). 특히, 소셜 미디어 및 공유 사이트의 등장은 인포메이션 아키텍처에 소셜화라는 새로운 측면을 부가하게 된 주된 이유가 되었다.

소셜 인포메이션 아키텍처는 구조화 된 정보를 시각적으로 표현하는 것 이상의 의미를 지니고 있다. 소셜 인포메이션 아키텍처는 정보자원을 중심으로 이용자 자신이 수행하는 정보활동을 이해할 수 있도록 정보자원과 관련된 컨텍스

트를 명확하게 제공해 줄 수 있다(EPA 2014). 즉, 정보자원을 활용하는데 있어서 이용자 개인이 자신들의 개인적인 목적을 충족시킬 뿐만 아니라 이를 통해서 다른 이용자와 상호작용을 할 수 있도록 지원하는 기능을 수행하고 있다. 이러한 측면에서, 소셜 인포메이션 아키텍처는 웹 상에서 이용자들이 다른 사람들과 상호작용을 할 수 있는 효과적인 공간을 설계하는 과정으로도 설명할 수 있다. 따라서, 소셜 인포메이션 아키텍처에서는 정보의 일반적인 제공을 위한 구조화가 아니라 정보의 공유, 대화, 상호작용, 이용자 사이의 협동 등이 중요한 의미를 지니게 된다(Brown 2010).

소셜 인포메이션 아키텍처의 이러한 특징은 소셜 메타데이터를 체계적인 방식으로 구조화하는데 있어서 여러 가지 가능성을 지니고 있다. 특히, 소셜 인포메이션 아키텍처에서 제공하는 조직화 시스템은 정보자원과 관련된 컨텍스트를 웹사이트에 추가함으로써 이용자들이 정보자원에 대한 기술사항으로써의 태그를 부여할 때 이를 사전에 의미적으로 조직화할 수 있는 구조를 제공해 줄 수 있다. 이에 본 연구에서는 정보자원에 대한 전체적인 기술사항으로서의 태그의 집합을 하나의 단위로 설정하고 태그의 부여 단계에서 소셜 인포메이션 아키텍처를 적용하여 정보자원과 관련된 서지적 컨텍스트를 제공해 줌으로써 태그의 조직화를 이루는 방식을 적용하고자 한다. 이용자가 부여하는 태그를 정보자원에 대한 종합적인 기술사항으로 고려할 경우, 이용자가 개인적인 목적으로 부여하는 태그를 정보자원의 기술사항에 적절한 방식으로 통합할 필요가 있으며, 이를 통해서 태그를 생성한 이용자 자신뿐만 아니라 다른 이용자에

계도 보다 다양한 관점에서 풍부한 기술사항을 제공해 줄 수 있게 된다(Smith 2007).

하지만, 소셜 인포메이션 아키텍처를 통한 서지적 컨텍스트의 제공이 이용자가 부여한 태그 자체의 의미적 호소성을 완전하게 해결할 수는 없으며, 이용자의 태그 부여 행위 자체를 통제할 경우 이용자의 자율성이라는 소셜 메타데이터의 특성을 저해할 수 있다. 따라서, 태그의 조직화와 태그 부여의 자율성을 매개할 수 있는 구조가 필요하며, 이용자가 참여하기 쉬운 방식의 구조화 된 체계의 마련이 필요하다. 이는 구조화 된 체계임과 동시에 이용자들이 자유롭게 소셜 메타데이터를 생성할 수 있는 유연성을 지니고 있어야 한다. 소셜 인포메이션 아키텍처는 정보자원에 대한 서지적 컨텍스트를 부여함으로써 이용자의 자율성과 태그의 조직화를 매개하는 매개적인 구조를 제공해 줄 수 있다.

## 4. 소셜 인포메이션 아키텍처 구축

### 4.1 분석대상의 선정

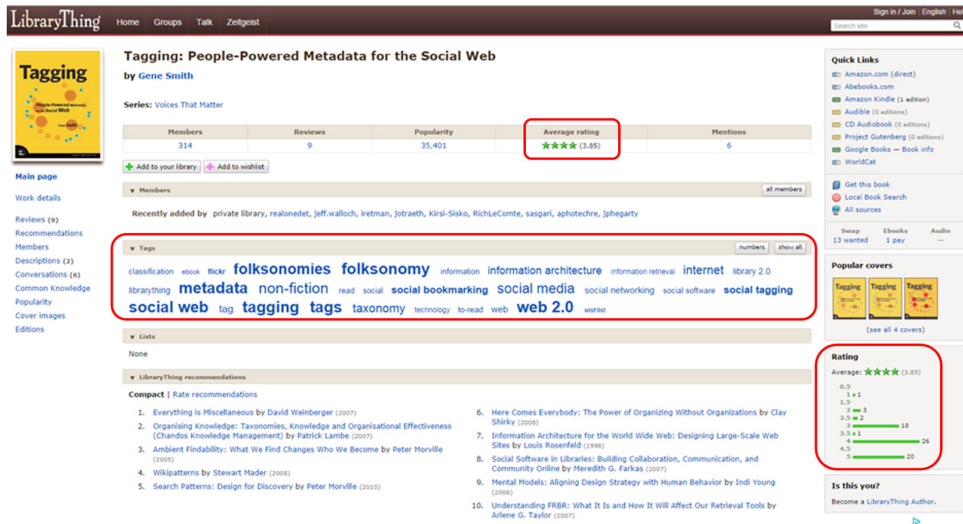
소셜 메타데이터의 환경에서 소셜 인포메이션 아키텍처의 적용 가능성을 확인하기 위해, 본 연구에서는 소셜 메타데이터의 유형 가운데 태그를 적용하고 있는 웹사이트를 선정하고 태그를 의미적으로 분석함으로써 이들 태그를 실효성있게 조직화 할 수 있는 방안을 제안하고자 한다. 웹사이트의 선정에 있어서는, 소셜 메타데이터의 기능에 이용자의 참여도가 높게 나타나고 부여된 태그의 활용도가 높은 LibraryThing

(<http://www.librarything.com>)을 연구의 대상으로 삼아 분석을 수행하였다.

LibraryThing은 2005년부터 시작된 도서 목록정보 제공 웹사이트로, 소셜 미디어 방식을 적용하여 정보자원에 대한 목록정보를 공유하기 위한 목적으로 구축되었다. 주로 단행본 자료에 대한 메타데이터 형식의 정보를 제공해 주고 있으며, 전 세계적으로 수많은 이용자들이 참여하여 양적으로 풍부한 소셜 메타데이터가 생성되고 있다. LibraryThing에서는 이용자들의 실질적인 참여가 활발하게 이루어지고 있어 질적으로도 높은 수준의 소셜 메타데이터가 생성·활용되고 있는 것으로 판단된다.

LibraryThing에서 이용자가 생성하는 소셜 메타데이터는 크게 태그(tag), 평가(rating), 리뷰(reviews) 등으로 구분해 볼 수 있다. 이 가운데, 평가는 독립적인 항목으로 별점의 형식으로 부여되고 있으며, 통일된 형식으로 이용자들에게 정보를 제공해 주고 있기 때문에 조직화의 필요성이 크게 나타나지 않고 있다. 이용자의 추천 여부 역시 독립적인 항목으로 구성되어 있으며, 추천을 하는지 하지 않는지로 구분된 두 가지의 선택이 통일된 방식으로 나타나기 때문에, 이용자의 의견이 일관성 있는 방식으로 나타나고 있다. 반면, 태그는 이용자들이 부여하는 다양한 형태의 단어 혹은 구 등이 단순히 나열되는 형태로 나타나고 있다. 이는 LibraryThing 웹사이트 내에 존재하는 정보자원에 대한 검색의 접근점으로써의 기능을 수행할 수는 있으나, 접근점의 선택에 있어서는 많은 비효율성을 내포하고 있다.

이에 본 연구에서는 LibraryThing에서 제공하는 소셜 메타데이터 가운데 태그에 중점을 두



〈그림 3〉 LibraryThing에서 제공하는 소셜 메타데이터의 예

고 분석을 수행하였다. 또한, 정보자원의 유형 가운데 단행본을 분석의 대상으로 삼았으며, 주제 분야를 메타데이터 분야로 한정하였다. 정보자원의 유형과 주제 분야의 한정은 부여된 태그의 의미를 보다 명확하게 구분하기 위한 것이며, 이를 통해서 부여된 태그의 의미를 보다 명확하게 분석할 수 있다.

#### 4.2 소셜 메타데이터 분석 결과

소셜 인포메이션 아키텍처를 적용한 태그의 조직화를 위해서는 정보자원에 대한 서지적 컨텍스트의 구축이 필요하다. 이를 위해, 본 연구에서는 LibraryThing에 부여된 태그를 양적인 측면 및 의미적인 측면으로 구분하여 분석을 수행하였다. 이를 통해, 이용자들이 정보자원을 기술하기 위해 실제적으로 부여한 태그의 다양한 측면을 확인할 수 있으며, 이를 기반으로 이용자 측면에서의 서지적 컨텍스트를 제공할 수

있는 기반을 마련할 수 있다.

##### 4.2.1 소셜 메타데이터의 양적인 분석

LibraryThing 내에 존재하는 정보자원 가운데 메타데이터 분야의 정보자원에 대해 부여된 태그를 정량적으로 분석한 결과는 다음과 같다.

LibraryThing 내에서 메타데이터 분야에 속하는 정보자원은 총 108개로 나타났으며, 이들 정보자원에 대해서는 총 1,567개의 태그가 부여되어 있다. 또한, 각각의 정보자원에는 최소 0개에서 최대 221개까지 태그의 양적인 범위가 형성되어 있다.

108개 정보자원에 대한 평균 태그의 수는 14.51개로 나타났다. 하지만, 이 가운데 태그가 부여되지 않은 정보자원의 수가 21개에 이르고 있으며, 이를 제외한 87개의 정보자원으로 범위를 한정하면, 정보자원에 대한 평균 태그의 수는 18.01개로 나타났다.

각 정보자원 당 부여된 태그의 비율을 살펴

보면, 1개~5개 사이의 태그가 부여된 정보자원은 총 33개로 전체의 30.56%를 차지하고 있다. 이외에 6개~10개의 태그가 부여된 정보자원이 총 18개(16.66%)로 높은 비율을 차지하고 있으며, 26개~30개의 태그가 부여된 정보자원도 11개(10.18%)로 높은 비중을 차지하고 있다(〈표 1〉 참조). 이 가운데, 21개 이상의 태그가 부여된 정보자원은 'textbook', 'LIS415', 'LS575' 등의 태그가 부여되어 있어 대학에서 교재로 사용되거나 인지도가 높은 정보자원으로 판단되며, 다른 정보자원에 비해 태그의 생성에 많은 사람들이 참여하고 있는 것으로 나타났다. 반면, 태그가 부여되지 않았거나 태그의 개수가 1개~5개인 정보자원은 오래된 정보자원이거나 메타데이터 분야 내에서도 특수한 주제를 다루는 전문서적인 경우가 많은 것으로 나타났다.

이상의 분석 결과를 종합해 보면, 이용자들이 직접 생성한 태그는 정보자원에 대해 양적으로 풍부한 기술사항을 제공해 주는 것을 알 수 있다. 하지만, 태그가 부여되지 않은 정보자원의 비율이 20% 가깝게 나타난 반면, 30개 이상의 태그가 부여된 정보자원의 비율 또한 7.41%에

이르고 있어, 양적인 측면에서의 태그의 부여는 정보자원에 따라 양극화 현상을 보이고 있는 것으로 나타났다. 또한, 부여된 태그의 수가 5개 이하인 정보자원의 비율이 50.00%에 이르고 있어, 이러한 경우 정보자원에 대한 기술사항으로서의 태그의 의미가 저하된다고 볼 수 있다. 하지만, 부여된 태그의 수가 적다하더라도, 기존의 서지사항에서 제공하는 기술사항과는 다른 측면에서의 기술사항을 제공할 경우 이는 정보자원에 대한 부가적인 정보를 제공해 준다는 측면에서 의의가 있는 것으로 볼 수 있다.

#### 4.2.2 소셜 메타데이터의 의미적 분석

태그 등과 같은 소셜 메타데이터는 기존의 전통적인 서지사항과는 다르게 이용자 개인의 측면에 집중한 기술사항이 두드러지게 나타나는 경향이 있다. LibraryThing에 부여된 태그 역시 기존의 서지사항에서 제공하는 기술사항과는 달리 이용자 중심으로 정보자원에 대한 기술사항이 제공되는 측면이 강하게 나타나고 있다. LibraryThing에서 이용자가 특정 정보자원에 대해 부여한 태그를 의미적으로 분석한 결과는 〈표 2〉와 같다.

〈표 1〉 LibraryThing에 부여된 태그의 양적 분석

태그 수	권 수	비율(%)
0개	21	19.44
1개~5개	33	30.56
6개~10개	18	16.66
11개~15개	8	7.41
16개~20개	5	4.62
21개~25개	4	3.71
26개~30개	11	10.18
30개 이상	8	7.41
총계	108	100.00

〈표 2〉 LibraryThing 태그의 의미적 구분

의미적 카테고리	태그 카테고리	실례
내용적 측면	저자 관련 태그	haynes_david, zeng
	저자 관련 정보	australian author, librarian, Library Technician
	장르	non-fiction, general, adult non-fiction
	주제 분야	cataloging, information science, library science
	컨텐츠 관련 키워드	TEI, XML, MODS, EAD, metadata, archives
	판차사항	3rd edition
형태적 측면	페이지수	280, 375
	자료 유형	ebook, ebook-toc, texts, paperback, softcover
	발행일자	2006, 2012, 21st century
	발행기관	ALA
	발행지역	North shore, Manchester, USA
	언어	English
운영적 측면	소장 기관	OU SLIS, City University London, @seattle
	자료 위치	central shelf, my office, Shirley's desk, @office
	이용자 수준	professional
	저작권 관련	@UBSMB
	용도	textbook, textbooks, LIS415, LS575, workstuff
개인적 목적 측면	이용자 추천	goodreads
	개인 목적	myreadinghistory, orders, checked in, borrowed, uncategorized, books reviewed, sold, wishlist, Already read, books-i-own, to-read, own, amazon purchases, signed by author
	개인 특화 태그	EL, 5c, 5d, breakfast at tiffanys, LocUnknown

LibraryThing에서는 이용자가 부여한 태그들이 의미적인 구분 없이 통합된 형태로 열거되는 방식으로 제공되고 있다. 하지만, 이들 태그들은 모두 정보자원이 지닌 다양한 측면들을 기술해 주고 있으며, 이를 그 기술의 대상이 되는 측면에 따라 의미적으로 구분해 보면 크게 정보자원의 '형태적 측면', '내용적 측면', '운영적 측면', '개인적 목적 측면' 등 4개의 의미적 카테고리로 구분할 수 있다(〈표 2〉 참조). 이렇게 구분된 각각의 측면들은 기술의 대상이 되는 정보자원과 관련된 서지적 컨텍스트를 제공해 줄 수 있으며, 이를 통해 태그가 지닌 의미를 구분해 줄 수 있다.

각각의 카테고리는 정보자원에 대해 실제적

으로 부여된 태그에 따라서 다시 총 20개의 세부적인 하위 카테고리로 구분할 수 있다. 이들 가운데는 저자 관련 사항, 출판 관련 사항, 주제 관련 사항 등 도서관계에서의 서지적 기술사항과 일치하는 항목들도 존재하지만, 이용자 추천이나 용도, 소장 기관 등과 같이 기존의 기술사항과는 다른 측면을 기술해 주는 카테고리도 다수 존재하고 있다.

이러한 카테고리 측면에서의 특징뿐만 아니라, 각각의 카테고리로 구분된 기술 내용에 있어서도 기존의 서지항목과 태그 사이에는 차이를 보이고 있다. 예를 들면, '장르'는 기존의 통제어휘에서 사전에 규정해 놓은 정형화 된 구분을 적용하는 것이 아니라, 'general', 'adult non-

function' 등과 같이 이용자의 관점에서 정보자원의 장르를 자유로운 형식으로 기술해 주고 있기 때문에, 해당 정보자원의 이용자들에게 보다 실제적인 관점에서의 기술사항을 제공해 주고 있다. 정보자원의 저자와 관련된 사항에 있어서도, 'australian author', 'library technician' 등과 같이 저자와 관련된 세부적인 사항들에 대한 정보를 제공해 주고 있다.

태그를 통해서 제공되는 이러한 세부적인 정보들은 각각의 태그가 독립적인 단위로 인식될 경우 어떤 의미를 지니고 있는지 파악하기 어렵다. 하지만, 이들 태그들을 서지적 컨텍스트

내에서 의미적으로 구분할 경우 이들 각각의 태그가 지닌 의미가 구체적으로 파악될 수 있으며, 이를 통해 해당 정보자원에 대한 보다 명확한 기술사항으로 적용할 수 있게 된다.

이와 같은 태그의 의미적 구분을 기반으로 태그의 조직화를 지원하기 위해서는 각각의 태그 카테고리가 전체적인 태그의 집합 내에서 차지하는 비율을 확인하는 과정이 선행되어야 한다. <표 2>에서 구분된 태그 카테고리가 LibraryThing 내에서 사용되는 비율을 확인한 결과는 다음 <표 3>과 같다.

LibraryThing에 부여된 총 1,567개의 태그

<표 3> LibraryThing 태그 카테고리의 비율

의미적 카테고리	태그 카테고리	태그 수	비율(%)
내용적 측면	저자명	35	2.23
	저자 관련 정보	10	0.64
	장르	17	1.08
	주제 분야	474	30.25
	컨텐츠 관련 키워드	166	10.59
	판차사항	12	0.77
소계		714	45.56
형태적 측면	페이지 수	9	0.57
	자료 유형	57	3.64
	발행일자	46	2.93
	발행기관	18	1.15
	발행지역	18	1.15
	언어	14	0.89
소계		162	10.34
운영적 측면	소장 기관	58	3.70
	자료 위치	134	8.55
	이용자 수준	24	1.53
	저작권 관련	21	1.34
	용도	87	5.55
소계		324	20.68
개인적 목적 측면	이용자 추천	15	0.96
	개인 목적	154	9.84
	개인 특화 태그	198	12.64
소계		367	23.42
계		1,567	100.00

가운데, 정보자원의 '내용적인 측면'과 관련된 태그가 총 714개(45.56%)로 가장 높은 비율을 차지하고 있다. 이 가운데, 해당 정보자원이 포함되는 주제 분야를 기술한 태그가 총 474개(30.25%)의 비율을 차지하고 있으며, 해당 정보자원에 수록된 구체적인 콘텐츠를 기술한 항목 또한 총 166개(10.59%)로 높은 비율을 나타내고 있다. 이들 항목들은 해당 정보자원이 수록하고 있는 콘텐츠를 기술해 주는 것으로서 전통적인 서지사항에서도 제공해 주는 것이지만, LibraryThing에서는 이용자 관점에서 기존의 기술사항에 비해 보다 풍부하고 자세한 내용을 제공해 준다는 차이를 보이고 있다. 특히, 정형화 된 통제어휘를 사용하지 않기 때문에 이용자의 관점에서 보다 실제적인 기술사항을 마련해 준다는 특징을 지니고 있다. 저자 정보 역시 기존의 서지사항에서도 제공해 주는 정보이지만, 해당 정보자원의 저자와 관련된 구체적인 정보를 제공해 주는 것으로 나타났다. 하지만, 저자정보(45개, 2.87%)와 관련해서 부여된 태그의 비율이 상당히 낮게 나타나고 있어, 저자와 관련된 태그는 독립된 카테고리 설정하기보다는 '내용적 측면' 카테고리에 속하는 것이 보다 합리적인 것으로 판단된다.

정보자원의 형태와 관련된 태그는 페이지 수, 자료의 유형, 자료의 발행과 관련한 사항, 언어 등으로 구분할 수 있다. 이들 태그의 대부분은 기존의 서지사항에서 제공하는 기술사항과 유사한 항목으로 구성되어 있지만, 실제로 기술된 내용에서는 기존의 서지사항과는 달리 보다 구체적이고 세부적인 사항들을 제공해 주고 있다. 이들 태그 가운데, 정보자원의 발행과 관련된 태그의 의미는 발행일자(46개, 2.93%),

발행지역(18개, 1.15%), 발행기관(18개, 1.15%) 등으로 구분할 수 있다. 하지만, 이들 각각의 카테고리는 그 발생빈도가 상당히 낮게 나타나고 있어 이들 역시 독립된 카테고리 설정하기보다는 정보자원의 현재의 형태를 발행한 '형태적 측면' 카테고리로 통합하는 것이 보다 효율적인 것으로 보인다.

정보자원의 운영과 관련된 측면을 기술해 주는 태그는 소장기관, 자료위치 등 해당 정보자원의 소장과 관련된 태그, 이용자 수준, 용도 등 해당 정보자원의 이용과 관련된 태그, 저작권과 관련한 태그 등으로 구분할 수 있다. 이 가운데, 해당 정보자원에 수록된 콘텐츠를 이용할 수 있는 이용자의 수준이나 해당 정보자원의 용도 등은 기존의 서지사항에서 제공하지 않는 기술사항으로, 태그에서 제공해 주는 고유한 기술사항으로서의 특징을 지니고 있다. 또한, 정보자원의 운영과 관련된 측면을 기술해 주는 태그는 총 324개(20.68%)로 나타나, 이용자들은 기존의 서지사항에서 제공하는 기술사항 이외에 해당 정보자원의 콘텐츠적인 수준이나 용도 등에 대한 기술사항에 대한 요구가 높게 나타남을 알 수 있다.

해당 정보자원에 대한 개인적인 이용 목적, 이용자 추천 등은 기존의 서지적 기술사항에서는 제공해 주지 못하는 항목이다. 이러한 유형의 태그들은 태그를 부여한 개인에게 특화된 것으로, 다른 이용자들에게는 큰 의미를 지니지 못하기 때문에 전체적인 태그의 활용도를 저하시킨다는 주장이 제기되기도 한다. 하지만, 이들 태그를 참조하여 부가적인 정보를 입수할 수 있는 이용자들이 존재할 수 있기 때문에, 이들 역시 해당 정보자원에 대한 포괄적인 기술

사항으로 간주하는 것이 바람직한 것으로 보인다. 또한, 이러한 개인적인 목적을 위한 태그는 그 생성 건수가 367개(23.42%)로 상당히 높게 나타나고 있어, 이는 이용자들이 태그를 부여하는 본질적인 목적을 반영하고 있는 것으로 나타났다. 이러한 측면에서, 개인적인 이용으로 의미가 국한되는 태그들은 '개인적 목적 측면' 카테고리를 마련하여 개인화 된 태그의 생성을 지원하는 것이 필요하다고 볼 수 있다.

부여된 태그의 의미적 분석을 통해 나타난 결과를 정보자원에 대한 기술사항이라는 측면에서 보면, 소셜 메타데이터로 이루어지는 정보자원의 기술 역시 전통적인 서지적 구조와 유사한 형태로 조직화 될 수 있는 가능성을 지니고 있다. 이에 본 연구에서는 태그의 의미적 분석을 기반으로 하여, 이용자 중심으로 생성되는 전체로서의 태그의 집합에 개인화 된 서지적 기술사항을 조직화 할 수 있는 소셜 인포메이션 아키텍처를 적용함으로써 태그의 조직화를 위한 기반을 마련하고자 한다.

#### 4.3 소셜 인포메이션 아키텍처 프레임워크 구축

소셜 인포메이션 아키텍처는 이용자가 생성한 소셜 메타데이터를 유연성있고 동적인 방식으로 조직화 할 수 있는 효과적인 방안으로 적용될 수 있다. 소셜 메타데이터로서의 태그에 적용되는 소셜 인포메이션 아키텍처는 기술의 대상이 되는 정보자원이 속한 서지적 컨텍스트를 구체적으로 표현해 줄 수 있는 구조적인 특성을 지니고 있다.

소셜 메타데이터로서의 태그에 소셜 인포메

이션 아키텍처를 적용하기 위한 첫 번째 단계는 태그의 의미적 카테고리를 프레임워크 구축의 기본적인 구조로 마련하는 것이다. 이는 각각의 태그가 지닌 의미를 분석한 결과를 기반으로 하고 있으며, 기술의 대상이 되는 정보자원의 특성에 따라 이들 기본요소에 세부적인 요소들이 추가되어 정보자원에 대한 전체적인 기술사항을 형성하게 된다.

〈표 3〉에서 제시한 카테고리 가운데, '내용적 측면' 카테고리는 정보자원에 수록된 콘텐츠와 관련된 태그를 수록할 수 있는 카테고리이며, 정보자원의 저자, 서명, 주제 등을 표현해 주는 태그를 부여할 수 있다. 또한, 해당 정보자원의 내용에 적합한 이용자의 수준 등에 대한 기술사항을 포함할 수도 있다. '형태적 측면' 카테고리는 정보자원의 물리적 형태 및 해당 정보자원을 현재의 형태로 발행한 기관 혹은 개인과 관련된 사항을 수록할 수 있는 카테고리, 정보자원의 발행사항, 크기, 정보자원의 형태, 웹 정보자원의 접근점 등을 수록할 수 있다. '운영적 측면' 카테고리는 전통적인 서지사항에서는 포함되지 않는 항목이지만, LibraryThing에 부여된 태그를 분석한 결과, 정보자원의 소장과 관련된 태그의 비중이 상당히 높게 나타나고 있다. 또한, 정보자료를 이용하는 목적을 기술하는데 사용된 단어들은 그 의미가 다른 항목과 중첩되어 나타나는 경우가 많기 때문에, 이를 하나의 카테고리로 설정하여 정보자원의 운영 및 활용적 측면과 관련된 태그들을 의미적으로 명확하게 구분해 주는 것이 필요한 것으로 나타났다. 이러한 이유로 '운영적 측면' 카테고리를 설정하였으며, 이 카테고리에는 해당 정보자원을 소장하고 있는 기관, 혹은 해당 정



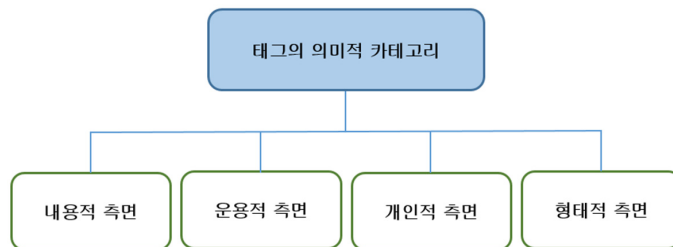
보자원이 위치하고 있는 지점 등에 대한 태그를 수록할 수 있다. '개인적 목적 측면' 카테고리는 정보자원을 이용하는 개인 혹은 태그를 부여하는 개인의 고유한 목적을 충족시키기 위한 태그를 수록할 수 있는 공간으로 설정하였다. 이들 태그의 의미적 카테고리 사이의 관계를 도식화하면 <그림 4>와 같다.

<그림 4>에 나타난 바와 같이, 태그의 의미적 카테고리는 기존의 인포메이션 아키텍처에서 제공하는 조직화시스템에 이용자들이 부여하는 태그를 적용할 수 있도록 조직화 한 것이다. 이러한 구조는 인포메이션 아키텍처의 조직화시스템을 이용해서 웹사이트의 콘텐츠를 조직화하는 구조와 유사한 형태를 지니고 있으며, 이를 통해서 웹사이트의 전체적인 구조에 태그의 의미적 카테고리를 결합할 수 있는 융통성을 부여해 줄 수 있다. 또한, 태그의 의미적 카테고리는 기술의 대상이 되는 정보자원의 서지적 컨텍스트를 제공해 줌으로써 이용자가 부여하는 태그의 의미를 사전에 조직화할 수 있는 기반으로서의 기능을 수행할 수 있다.

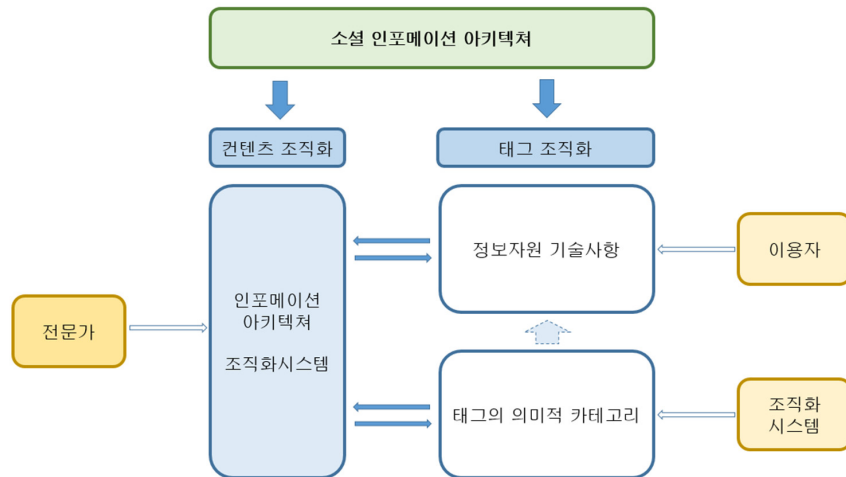
하지만, 이 구조가 기술의 대상이 되는 정보자원의 서지적 컨텍스트를 제공하는 것으로 그치게 되면, 이는 각각의 태그가 지닌 의미에 따라 태그를 카테고리화 한 것에 지나지 않게 된

다. 이 구조에서 제공하는 서지적 컨텍스트가 정보자원에 대한 기술사항으로 이어지기 위해서는, 태그의 의미적 카테고리들을 독립적인 조직화시스템으로 간주하는 것이 아니라 기존의 전문가에 의해서 생성되는 서지사항의 구조화된 체계와 상호운용할 수 있는 개념적인 프레임워크가 마련되어야 한다. 이는 소셜 인포메이션 아키텍처를 적용함으로써 이루어질 수 있으며, 이를 통해 특정 정보자원에 대한 풍부한 기술사항 및 다양한 접근점을 융통성있게 결합시킬 수 있게 된다. 전문가에 의해 생성되는 구조적인 서지사항과 이용자에 의해 생성되는 자유로운 기술사항이 결합될 수 있는 소셜 인포메이션 아키텍처를 개념화 한 구조는 다음 <그림 5>와 같다.

<그림 5>에 나타난 바와 같이, 이용자가 부여하는 태그는 기존의 인포메이션 아키텍처 구조와 결합하여 전체적인 소셜 인포메이션 아키텍처의 구조를 형성한다. 이용자가 부여하는 태그는 사전에 마련된 의미적 카테고리에 따라 생성되며, 생성된 태그는 독립적인 단위가 아닌 전체적인 태그의 집합으로써 정보자원에 대한 포괄적인 기술사항으로 인식될 수 있다. 이때, 태그의 의미적 카테고리는 인포메이션 아키텍처의 조직화시스템을 적용하여 태그의 부



<그림 4> 태그의 의미적 카테고리



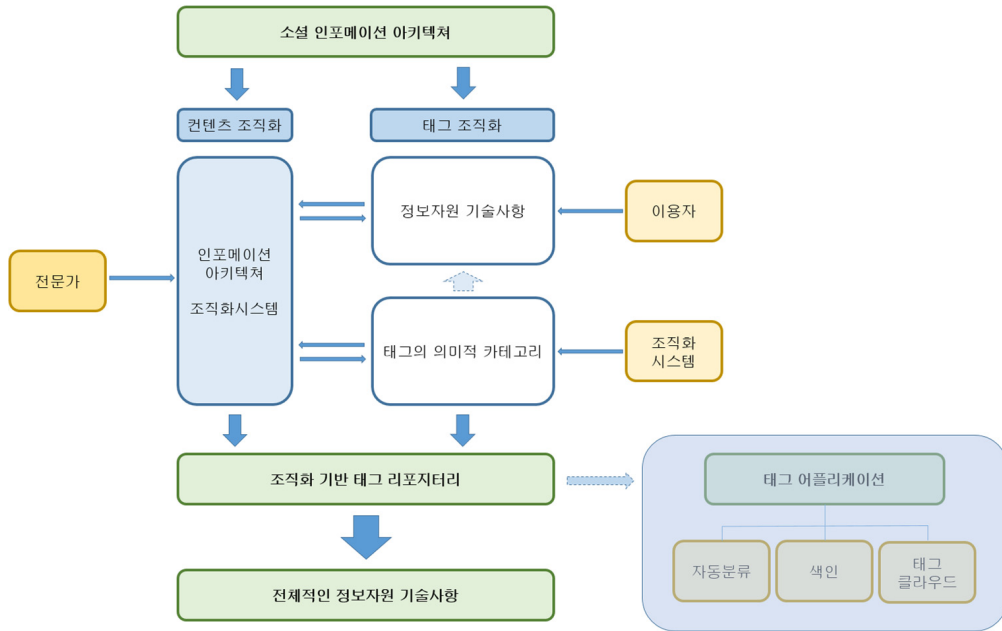
〈그림 5〉 정보자원 기술사항으로서의 소셜 인포메이션 아키텍처 구조

여 단계에서 조직화의 과정을 거치게 되며, 부여된 태그는 정보자원에 대한 전체적인 기술사항으로서의 기능을 하게 된다.

인포메이션 아키텍처는 전문가에 의해서 생성된 정보자원의 기술사항을 체계화 된 구조에 따라 제공해 주며, 이용자가 부여한 태그의 의미적 카테고리는 전체적인 소셜 인포메이션 아키텍처에서 제공하는 기술항목들과 연계됨으로써 이용자가 참여하는 풍부한 기술사항을 제공할 수 있게 된다. 결국, 정보자원에 대한 전체적인 기술사항은 기존의 전문가에 의한 기술사항과 이용자가 참여하여 생성한 기술사항이 결합되어 소셜 인포메이션 아키텍처의 구조 내에서 생성되는 것으로 볼 수 있다. 이는 태그를 생성하는 이용자의 자율성과 전문가에 의해 생성된 구조화 된 체계를 접목시킨 것으로, 이 두 가지 형태의 기술사항은 상호보완적인 특성을 지니며, 이를 통해 보다 풍부하고 실제적인 정보자원의 기술사항을 제공해 줄 수 있게 된다. 이와 함께, 정보자원에 대한 전체적인 기술사

항은 특정 정보자원에 대한 다양한 접근점을 항구적으로 마련해 주는 리포지터리 형태로 제공될 수 있다. 이러한 전체적인 소셜 메타데이터 프레임워크를 도식화하면 〈그림 6〉과 같다.

인포메이션 아키텍처는 웹사이트에 수록된 콘텐츠의 전체적인 구조를 형성하는데 주로 적용되고 있지만, 소셜 인포메이션 아키텍처를 적용한 소셜 메타데이터 프레임워크는 특정 정보자원에 대한 기술사항으로서의 태그를 부여하는 경우에 특화된 프레임워크라고 할 수 있다. 따라서, 소셜 인포메이션 아키텍처는 정보자원을 기술해 주는 모든 웹페이지 상에서 구조화되고 일관적인 방식으로 소셜 메타데이터를 부여할 수 있는 구조를 마련해 주어야 한다. 이를 통해, 기존의 메타데이터와 함께 이용자가 부여한 태그가 동시에 제공될 수 있으며, 특정 정보자원에 대한 전문가 측면과 이용자 측면에서의 기술사항을 통합하여 풍부한 기술사항 및 접근점을 마련해 주는데 보다 효과적으로 활용될 수 있다.



〈그림 6〉 소셜 인포메이션 아키텍처를 적용한 소셜 메타데이터 프레임워크

기존의 서지사항과 태그의 집합이 결합한 전체적인 기술사항은 정보자원과 관련한 서지적 컨텍스트를 기반으로 형성되기 때문에, 이는 기존의 서지사항에서 적용하는 기술항목들과 의미적으로 연결될 수 있다. 이러한 매개적인 조직화의 체계를 통해서, 소셜 인포메이션 아키텍처에 기반한 전체적인 기술사항은 태그 리포지터리를 형성할 수 있다. 태그 리포지터리는 동일한 서지적 컨텍스트에 해당되는 정보원에 대한 전체적인 기술사항을 항구적으로 제공할 수 있게 되며, 이를 통해 정보원의 검색의 접근점으로 활용될 수 있는 개념적인 구조를 갖출 수 있게 된다. 또한, 태그 리포지터리에 수록된 정보자원 기술사항은 기존의 태그 조직화 접근방안이 지니고 있는 태그의 의미적 모호성이라는 문제를 구조적인 체계를 적용함으로써 해

결해 줄 수 있다. 이를 통해, 태그의 자동분류, 태그의 색인화 시스템, 태그 클라우드 등 기존의 태그 조직화 접근방안에 태그의 의미적 구조를 제공해 줌으로써 자동화 된 태그의 조직화를 실제로 구현할 수 있는 의미적 기반을 제공할 수 있을 것으로 기대된다.

이러한 매개적인 형태의 프레임워크는 소셜 메타데이터로서의 태그를 완전하게 조직할 수 있는 방안이 아니라 태그의 체계화 된 조직을 지원할 수 있는 대안적인 기반을 마련하는데 중점을 두고 있다. 이러한 측면에서, 소셜 메타데이터 프레임워크는 정보원에 대한 전문가 측면 및 이용자 측면의 전체적인 기술사항을 효과적으로 제공함과 동시에 태그를 이용한 정보검색의 효율성을 확보하는데도 적용될 수 있을 것으로 기대된다.

## 5. 결론

웹 2.0으로의 진화는 이용자의 정보활동에 있어서 패러다임의 변화를 이끌어 내고 있다. 웹 2.0 환경에서 이용자는 정보활용을 위한 플랫폼을 공유함으로써 정보의 소비자로서의 역할뿐만 아니라 콘텐츠의 생산자로서의 역할 또한 수행하고 있다. 이러한 이용자 중심의 정보 활동에는 정보의 생산과 소비에 이용자가 직접 참여하는 소셜 메타데이터가 큰 역할을 담당하고 있다. 하지만, 소셜 메타데이터는 이용자의 자발적인 참여라는 특성으로 인해서, 모호한 용어의 사용, 비체계적인 구조, 개인화 된 정보 생성 등의 문제가 지속적으로 제기되고 있다. 이러한 문제를 해결하기 위한 방안으로 이용자가 부여한 소셜 메타데이터를 조직화하기 위한 연구가 계속해서 수행되어 왔다. 하지만, 이들 연구의 대부분은 소셜 메타데이터가 지닌 의미적 모호성을 해결하지 않은 상태에서 조직화를 수행하였기 때문에 큰 효과를 거두지 못하고 있다. 이에 본 연구에서는 소셜 메타데이터의 다양한 형식 가운데 이용자가 직접 부여하는 태그를 중심으로 소셜 메타데이터를 의미적으로 조직화하는 대안적인 방안을 제안하고 있다.

본 연구에서 제안하는 접근방법은 소셜 메타데이터의 본질적인 기능인 정보자원에 대한 종합적인 기술사항으로서의 측면에 중점을 두고 있으며, 부여된 태그를 대상으로 하는 것이 아닌 이용자가 태그를 부여하는 단계에서의 조직화를 위한 방안을 제안하고 있다. 이는 기술의 대상이 되는 정보자원이 지닌 서지적 컨텍스트를 제공함으로써 이용자가 태그를 부여하는 상황에서 태그의 의미적 조직화를 유도하고, 이

를 통해 이용자가 참여하는 풍부한 기술사항으로서의 소셜 메타데이터의 효용성을 높이기 위한 목적을 지니고 있다. 이를 위해 본 연구에서는 소셜 인포메이션 아키텍처를 적용하여 소셜 메타데이터를 의미적으로 조직화하는 소셜 메타데이터 프레임워크를 구축하였다.

소셜 메타데이터 프레임워크의 구축을 위해서는 정보자원과 관련된 서지적 컨텍스트의 구축이 선행되어야 한다. 이를 위해, 본 연구에서는 LibraryThing 웹사이트에서 활용하고 있는 태그를 대상으로 각각의 태그가 지닌 의미를 분석하고, 이에 인포메이션 아키텍처의 조직화 시스템을 적용하여 태그의 의미적 카테고리를 설정하였다. LibraryThing 웹사이트에 부여된 태그를 의미적으로 구분한 결과, 태그의 의미는 '내용적 측면', '형태적 측면', '운영적 측면', '개인 목적 측면' 등 네 가지 카테고리로 구분할 수 있으며, 이는 이용자들이 태그를 부여하는 단계에서 기술의 대상이 되는 정보자원과 관련된 서지적 컨텍스트를 제공해 주게 된다.

소셜 인포메이션 아키텍처는 태그의 의미적 분석을 통해 구축된 서지적 컨텍스트를 기존의 정보자원에 대한 기술사항으로 연결시킬 수 있는 개념적인 구조를 마련해 준다. 이를 통해 이용자가 부여하는 태그의 전체적인 집합과 전문가에 의해 생성되는 기술사항을 구조적으로 연결시킴으로써 해당 정보자원에 대한 전체적인 기술사항을 제공해 준다. 이는 태그를 생성하는 이용자의 자율성과 전문가에 의해 생성된 구조화 된 체계를 접목시킨 것으로, 이 두 가지 형태의 기술사항은 상호보완적인 특성을 지니며, 이를 통해 보다 풍부하고 실제적인 정보자원의 기술사항을 제공해 줄 수 있게 된다.

이렇게 생성된 정보자원의 전체적인 기술사항은 특정 정보자원에 대한 다양한 접근점을 항구적으로 마련해 주는 리포지터리 형태로 제공될 수 있다. 또한, 태그 리포지터리에 수록된 정보자원 기술사항은 기존의 태그 조직화 접근방안이 지니고 있는 태그의 의미적 모호성이라는 문제를 구조적인 체계를 통해 해결해 줄 수 있다. 이를 통해, 태그의 자동분류, 태그의 색인화 시스템, 태그 클라우드 등 기존의 태그 조직화 접근방안에 태그의 의미적 구조를 제공해 줌으로써 자동화 된 태그의 조직화를 실제적으로 구현할 수 있는 의미적 기반을 제공해 줄 수 있을 것으로 기대된다.

이러한 매개적인 형태의 소셜 메타데이터 프레임워크는 소셜 메타데이터로서의 태그를 완전하게 조직할 수 있는 방안을 제시하지는 못하고 있다. 하지만, 이용자가 부여한 소셜 메타데이터로서의 태그를 의미적으로 구분하고, 이를 통해 태그의 조직화를 이룰 수 있는 의미적 기반을 마련해 줌으로써 전문가 측면 및 이용

자 측면에서 생성한 전체적인 기술사항을 효과적으로 제공해 줄 수 있는 대안적인 방안으로서의 의미를 지니고 있다. 또한, 기존의 태그 조직화를 위한 접근방안이 지니고 있는 태그의 의미적 구분이라는 문제를 해결할 수 있는 의미적인 조직화의 기반을 마련함으로써, 향후 소셜 메타데이터의 어플리케이션과 연계되어 소셜 메타데이터를 이용한 정보검색의 효율성을 확보할 수 있을 것으로 기대된다.

하지만, 본 연구에서는 LibraryThing 웹사이트만을 대상으로 하였기 때문에, 태그를 적용한 다양한 웹사이트의 경우를 반영하지 못하고 있다는 한계를 보이고 있다. 이와 함께, 대상이 되는 정보자원을 단행본으로 한정하였으며, 주제 분야 역시 메타데이터 분야로 한정하였다는 한계를 지니고 있다. 향후의 연구에서는 소셜 메타데이터를 제공하는 다양한 웹사이트를 대상으로 여러 가지 유형의 정보자원에 대한 소셜 메타데이터를 분석하여 보다 일반화 된 방식의 태그 조직화 방안을 제안하는 것이 필요하다.

## 참 고 문 헌

- [1] 강신재 and Ding, Y. 2008. 태그 온톨로지와 기계학습을 이용한 추천시스템. 『한국산업정보학회논문지』, 13(5): 133-141.
- [2] Al-Khalifa, H. S. and Davis, H. C. 2007. "Exploring the Value of Folksonomies for Creating Semantic Metadata." *International Journal on Semantic Web and Information Systems*, 3(1): 13-39.
- [3] Alemu, G. B., Stevens, P. R. and Chandler, J. 2012. "The Social Space of Metadata: Perspectives of LIS Academics and Postgraduates on Standards-Based and Socially Constructed Metadata Approaches." *Journal of Library Metadata*, 12(4): 311-344.

- [4] Bischoff, K. et al. 2008. Can All Tags Be Used for Search? In *Proceedings of the 17th ACM Conference on Information and Knowledge Management*, Napa Valley, California, October 26-October 30, 2008, pp.193-202.
- [5] Brown, D. 2010. "Eight Principles of Information Architecture." *Bulletin of the American Society for Information Science and Technology*, 36(6): 30-34.
- [6] Environmental Protection Agency (EPA). 2014. *Social Impact Assessment*. [online] [cited 2014. 8. 27.] <<http://www.epa.gov/sustainability/analytics/social-impact.htm>>
- [7] Golder, S. and Huberman, B. 2006. "Usage Patterns of Collaborative Tagging Systems." *Journal of Information Science*, 32(2): 198-208.
- [8] Guy, M. and Tonkin, E. 2006. Folksonomies: Tidying up tags?. *D-Lib Magazine*, 12(1). [online] [cited 2014. 7. 25.] <<http://www.dlib.org/dlib/january06/guy/01guy.html>>
- [9] Kakali, C. and Papatheodorou, C. 2010. "Exploitation of Folksonomies in Subject Analysis." *Library & Information Science Research*, 32: 192-202.
- [10] Kipp, M. 2007. Tagging Practices on Research Oriented Social Bookmarking Sites. In *Proceedings of the 35th Conference of the Canadian Association for Information Science*, Montreal, QC, May 10-12, 2007. [online] [cited 2014. 6. 24.] <<http://www.dlist.sir.arizona.edu/1533/>>
- [11] Lin, X. et al. 2006. Exploring Characteristics of Social Classification. In *Proceedings of the 17th ASIS&T Classification Research Workshop*, Austin, Texas, November 4, 2006.
- [12] Mika, P. 2007. "Ontologies Are Us: A Unified Model of Social Networks and Semantics." *Journal of Web Semantics*, 5(1): 5-15.
- [13] Online Computer Library Center (OCLC). 2012. *Social Metadata for Libraries, Archives and Museums*, Part 1: Site Reviews. [online] [cited 2013. 5. 15.] <<http://www.oclc.org/research/publications/library/2012/2012-01r.html>>
- [14] Paul, P. K. 2014. Information Architecture and Social Information Architecture: Valuable Domain for Building Information Society - A Conceptual Overview. *Pinnacle Educational Research & Development*, 2014. [online] [cited 2014. 9. 2.] <[http://pjpeg.org/perd/perd\\_110.pdf](http://pjpeg.org/perd/perd_110.pdf)>
- [15] Saracevic, T. 2010. The Notion of Context in "Information Interaction in Context." In *Proceedings of the IliX 2010 Information Interaction in Context Symposium*, New Brunswick, NJ, USA, August 18-21, 2010.
- [16] Smith, G. 2007. *Social Information Architecture*. Presentation in Information Architecture Summit 2007, Las Vegas, Nevada, US, March 22, 2007. [online] [cited 2014. 8. 25.] <<http://www.slideshare.net/gsmith/social-information-architecture-workshop>>

- [17] Smith, Gene. 2008. "Tagging: Emerging Trends." *Bulletin of the American Society for Information Science and Technology*, 34(6): 14-17.
- [18] Thomas, M., Caudle, D. and Schmitz, C. 2009. "To Tag or Not to Tag?" *Library Hi Tech*, 27(3): 411-434.
- [19] Tonkin, E. 2006. Searching the Long Tail: Hidden Structure in Social Tagging. In J. Furner & J. T. Tennis, ed., In *Proceedings of the 18th Workshop of the American Society for Information Science and Technology Special Interest Group in Classification Research*, Austin, Texas, November 4, 2006.
- [20] Voss, J. 2006. *Collaborative Thesaurus Tagging the Wikipedia Way*. [online] [cited 2014. 7. 25.] <<http://www.arxiv.org/abs/cs.IR/0604036>>
- [21] Warterna, C. and Brussee, R. 2008. Instanced-Based Mapping Between Thesauri and Folksonomies. In *Proceedings of the 7th International Semantic Web Conference*, Karlsruhe, Germany, October 26-30, 2008, pp.356-370.
- [22] Wra, T. and Eklund, P. W. 2010. Social Tagging for Digital Libraries Using Formal Concept Analysis. In Kryszkiewicz M. and Obiedkov, S. ed., *CLA-2010: Concept Lattice and Their Applications*. [online] [cited 2014. 9. 10.] <<http://works.bepress.com/peklund/1>>

• 국문 참고자료의 영어 표기

(English translation / romanization of references originally written in Korean)

- [1] Kang, S. and Ying Ding. 2008. "Recommendation System Based on Tag Ontology and Machine Learning." *Journal of the Korea Industrial Information Systems Research*, 13(5): 133-141.