

# 대학도서관 폭소노미 태그의 형태적 특성에 관한 연구\*

## A Study on Form of Folksonomy Tags in University Libraries

이 성 숙(Sung-Sook Lee)\*\*

### 목 차

- |                         |                       |
|-------------------------|-----------------------|
| 1. 서론                   | 3. 폭소노미 태그의 형태적 특성 분석 |
| 1.1 연구의 필요성 및 목적        | 3.1 용어의 형태            |
| 1.2 연구방법 및 범위           | 3.2 용어의 유형            |
| 2. 폭소노미 태그의 이론적 배경      | 3.3 용어의 선택            |
| 2.1 폭소노미 태그의 기본 개념      | 3.4 용어의 표기            |
| 2.2 폭소노미의 특징과 장단점       | 3.5 동형이의어             |
| 2.3 정보서비스 영역에서 폭소노미의 적용 | 3.6 분석결과              |
|                         | 4. 결론 및 제언            |

### 초 록

이 연구는 국내 대학도서관 폭소노미 태그의 형태적 특성을 분석하여, 통제어휘 개발 가이드라인과 비교해볼 때 어떠한 특징이 있는지 검토하기 위한 것이다. 연구대상으로 선정된 대학도서관의 6개월간 태그를 토대로, 폭소노미 태그의 구조와 형태를 조사하였다. 분석 대상 태그들은 시소러스 개발 가이드라인을 토대로 조사 분석하였다. 이 연구 결과는 디지털도서관에 폭소노미 태그의 활용 방안에 관한 기초자료를 제공할 것이다.

### ABSTRACT

This study was to review the possible characteristics and patterns that occur when comparing control language constructing guidelines, by analyzing the formal characteristics of folksonomy tags in university libraries. Based on subjected tags at university libraries for a period of 6 months the structure and form of folksonomy was examined. The object tags were analyzed based on the thesaurus development guidelines. The results for this research will provide baseline data for the use of folksonomy tag applications in digital libraries.

키워드: 폭소노미, 태그, 협력태깅, 소셜태깅, 시소러스, 대학도서관, 온라인목록

Folksonomy, Tag, Collaborative Tagging, Social Tagging, Thesaurus, Form

\* 이 논문은 2006년도 충남대학교 학술연구비의 지원에 의하여 연구되었음.

\*\* 충남대학교 사회과학대학 문헌정보학과 조교수(infolee@cnu.ac.kr)

논문접수일자: 2008년 11월 27일 최초심사일자: 2008년 11월 27일 게재확정일자: 2008년 12월 10일

## 1. 서론

### 1.1 연구의 필요성 및 목적

웹 1.0에서 진화한 웹 2.0은 도서관 정보서비스에 영향을 주고 있으며, 이에 따라 정보조직 환경에도 변화가 일어나고 있다. 웹 2.0은 오라 일리미디어사(O'reilly Media, Inc.)의 대표인 Tim O'reilly에 의해 도입된 개념으로, 2004년 Dale Dougherty가 닷컴 버블 붕괴 이후에도 지속적인 성장세를 보이고 있는 아마존, 이베이, 구글 등의 인터넷 서비스에서 발견되는 차별적 특징들을 다른 서비스들과 비교하기 위해 사용하였다.

웹 2.0의 특징으로는 개방과 참여, 공유를 들 수 있으며, 웹 2.0 환경에서는 정보와 정보 간, 사람과 사람간의 연결이 강조되고, 모든 정보를 쉽게 공유하고 활용할 수 있으며, 참여와 상호작용의 결과물인 '집단지성'의 중요성이 부각된다. 특히 주목할 점은 웹 2.0 참여자들의 협업에 의해 기존 전문가를 능가하는 결과물을 낳기도 한다는 점이다. 대표적인 예가 온라인 백과사전인 위키피디아(wikipedia)이다. 수많은 비전문가들이 온라인 참여자들이 만들고 있는 위키피디아는 이미 양적인 면에서나 질적인 면에서 브리태니커 백과사전을 넘어섰다는 평가를 받고 있다(구본홍 2007).

웹 2.0 환경에서 효율적인 정보검색을 제공하기 위한 방안으로 폭소노미(folksonomy)가 주목받고 있다. 폭소노미는 이용자들이 태그를 통해 공동으로 형성한 새로운 분류체계로서 정보들을 유기적으로 연결할 수 있으며 웹 정보의 변화를 빠르게 수용할 수 있다. 또한 폭소노

미를 통한 검색은 인지구조에서 도출된 다양한 접근점을 제공할 수 있을 뿐만 아니라 분산된 정보를 연결하여 통합적으로 이용할 수 있도록 하여 현재 정보서비스 영역에서 이용되고 있다(유시내 2007).

웹 2.0에서 언급되는 태그(tag)란 웹콘텐츠의 내용을 대표할 수 있는 짧은 단어로써 데이터에 대한 정보를 기술하는 메타데이터이다. 태그는 이용자의 참여를 통해 새로운 가치가 형성되는 과정을 가장 잘 보여준다.

전통적으로 도서관에서는 주제명 표목표나 시소러스를 이용해 내용의 주제를 색인해왔다. 그런 색인은 사서 혹은 전문 색인자들에 의해 이뤄지거나, 문헌 저자에 의해 파생될 수도 있다. 이용자 중심의 정보환경변화로 현재 도서관에 도입된 이용자 맞춤형서비스인 My Library에서도, 이용자는 검색된 결과의 자료를 자유롭게 조직하고 범주화할 수는 없다. Demspy(2003)와 Ketchell(2000)에 의하면, 도서관은 이용자가 관심을 갖는 자원에 대한 어노테이션(annotation)이 가능하도록 해야 하고, 그런 어노테이션을 다른 이용자와 공유할 수 있는 환경을 제공해야 한다.

최근 국내 도서관에서도 폭소노미를 적용한 소셜 북마킹, 소셜 온라인 목록, 참고문헌 공유 등의 서비스가 제공되고 있다. 폭소노미는 이용자 자신의 개인 정보 공간을 조직함과 함께 도서관 목록에 중요하게 기여할 수 있는 가능성을 가지고 있다. 예를 들어 이용자는 도서관 목록에서 개별자료, 외부 DB, 외부 웹페이지 등에 태그를 달고, 태그를 통해 개별자료를 검색하고, 저장하고, 조직할 수 있다. 이용자들은 태깅시스템의 간단한 인터페이스를 통해 메타

데이터 생성에 자발적으로 참여하며, 태깅에 사용된 이런 메타데이터가 검색에 재사용되어 검색의 효율성을 높이기도 한다.

그러나 폭소노미에는 인터넷상의 자원을 어떻게 분류해야 한다는 구체적인 정책이 없다. 이 때문에 무질서하고 부정확한 태그와 방대한 태그의 양, 태그가 어휘로서 본질적으로 가지는 한계 등이 문제점으로 나타난다.

현재 국내 폭소노미에 관한 연구는 폭소노미의 개념상의 이해와 의의에 관한 연구(이재윤, 황혜경 2006; 이정미 2007), 폭소노미를 어떻게 적용할 것인가에 대한 실용적인 측면을 강조한 연구(김수현 2007; 장석중 2007; 유시내 2007)가 있다. 또한 폭소노미를 색인어의 관점에서 통제어휘와 순수한 비통제어휘의 중간에 해당하는 특성을 보인다고 고찰한 연구(이재윤, 정도현 2008)가 있다. 그 외에 한글 웹 환경에서 태그 클라우드 표현(안희영, 이현주 2007)과 사용자간 상호작용의 개선방안(박찬표, 반영환, 정지홍 2007) 등 디자인적 측면을 다룬 논문이 있다.

최근 도서관에서는 점차 폭소노미를 도입하여 서비스하고 있지만, 폭소노미가 도서관 목록의 상호작용을 증진시킬 수 있는 도구로서 보다 효과적으로 활용되기 위해서는 기존의 통제어휘 체계와의 관계 정립이 필요하다. 이를 위해서는 폭소노미 태그와 전통적인 통제어휘 색인과의 부합 정도를 실증적으로 파악하는 것이 필요하다. 이 연구의 목적은 폭소노미 태그의 특징을 형태적인 측면에서 분석하여 디지털 도서관에서의 활용가능성을 살펴보기 위한 것이다. 이 연구에서는 국내 A 대학도서관에서 서비스되고 있는 폭소노미 태그의 형태적 특성

을 분석하여, 통제어휘 개발 가이드라인과 비교해볼 때 어떠한 특징이 있는지 검토하였다.

## 1.2 연구방법 및 범위

연구방법은 문헌연구와 조사대상 도서관의 태그 분석을 병행하였다. 현재 국내 대학도서관 중 태그 클라우드(cloud) 서비스를 제공하고 있는 A 대학도서관을 선정하였다. 도서관의 협조아래 2008년 5월부터 10월까지 6개월간의 태그(한글태그 271건, 영어태그 153건)를 수집하였다. 태그 분석에서는 여러 번 태깅된 태그는 하나의 태그로 간주하였고, 동일 태그의 변형은 두 개의 태그로 간주하였다.

이 연구에서는 폭소노미 태그의 형태적 특성을 파악하기 위한 기준으로 시소러스 개발 가이드라인을 사용하였다. 통제어휘색인에서 사용되는 시소러스는 색인시에는 색인어의 선정에, 검색시에는 검색어와 색인어간의 관계를 연결해주는 역할을 한다. 시소러스에서 중요한 것은 용어 적용의 일관성이며, 일관성을 위해서는 일반적인 기준이 필요하고, 이를 위해 시소러스 개발 가이드라인을 활용한다. 시소러스 개발 가이드라인은 통제어휘를 위한 가이드라인이지만, 폭소노미 태그의 특징을 살펴보는 데 유용할 것이다.

이 연구에서 사용한 시소러스 개발지침은 한글태그에 대해서는 최석두(2000, 2007), 조순영(2002), 이상현(2002) 등의 시소러스 개발 가이드라인을 분석한 결과, 국제표준을 준수하고 한글의 특성을 고려한 최석두(2000)의 가이드라인이 다른 가이드라인들을 포괄한다고 판단되어 이 가이드라인을 사용하였다. 영어태그에

대해서는 NISO(2005)의 '단일언어 통제어휘의 구축, 형식 및 관리 가이드라인'을 사용하였다.

이 연구의 범위는 폭소노미 태그의 형태(form)상의 관점에서 용어 적용의 일관성을 중심으로 한 특성으로 한정한다. 구체적으로 최석두(2000)의 7장 색인어와 NISO(2005)의 6장 용어의 선택과 범위, 형태에 해당된다. 이 연구에서는 폭소노미 태그의 형태적 특성을 용어의 문법적인 형태와 개념 유형, 용어의 선택과 표기, 동형이의어를 중심으로 분석하였다.

또한 태그의 품사와 의미, 외래어, 절차법, 신조어 등을 확인하기 위해서 국립국어원의 표준국어대사전, 신어사전, 동아프라이밍 영어사전, 웹스터 사전(Merriam-Webster Online)을 참고하였다. 이외에 개인명과 단체명, 사건명 등 개별실체(unique entites)의 의미를 파악하기 위해 사전이외의 위키피디아, 네이버, 구글도 활용하였다.

이 연구를 통해 폭소노미 태그에 어떤 특징이 존재하는지의 여부와 통제어휘 개발을 위한 가이드라인을 어느 정도 반영하고 있는지 살펴볼 수 있을 것이다. 이 연구의 결과는 디지털도서관에 폭소노미 태그의 활용 방안에 관한 기초자료를 제공할 수 있을 것이다.

## 2. 폭소노미 태그의 이론적 배경

### 2.1 폭소노미 태그의 기본 개념

폭소노미는 협력태깅(collaborative tagging), 사회적 분류(social classification), 사회적 색인(social indexing), 사회적 태깅(social tagging)

등으로 불리기도 한다. 폭소노미는 이용자 생성 메타데이터로서 '사람들에 의한 분류법'(folk + order + nomos)이라는 뜻의 합성어로 원래는 Thomas Vander Wal이 메일링리스트 토론 중에 대중을 뜻하는 'folk'와 분류(학)을 의미하는 'taxonomy'를 결합하여 만든 표현이다.

폭소노미는 일반 대중이 직접 정보를 분류하고 레이블링하여 만든 정보 분류 시스템 혹은 네트워크 구조로서 전통적인 분류 기준인 '디렉토리' 대신 '태그'에 따라 나누는 새로운 분류 체계를 말한다. 또한 위키피디아(2008)에서는 콘텐츠를 설명하고 분류하려는 태그를 협력적으로 만들고 관리하는 방식을 의미하는 신조어로 설명하고 있다.

폭소노미 정의에서 언급되는 '협력'이라는 용어는 폭소노미의 공통된 특성 중 하나이며, 이용자들의 관심과 가치부여에 따라 정보의 분류 체계가 달라질 수 있다는 것을 의미한다. 폭소노미는 다른 사람에 의해 할당된 태그에 접근하고 이용할 수 있는 환경에서 만들어지며, 폭소노미는 공유된 태그의 사용을 가능하게 한다. 이와 같은 기법으로 웹 2.0에서 활용되고 있는 것이 소셜 북마크(social bookmark) 도구이다.

폭소노미에서는 이용자가 정보에 의미를 부여하기 위해서 태깅(tagging) 기술이 사용된다. 태깅은 이용자가 태그를 붙이는 행위를 말한다. 꼬리표라고도 불리는 태그란 이미지, 지도, 블로그, 북마크 등 정보에 부여하는 키워드나 용어이다. 태그는 카테고리보다는 키워드에 더 가까운 의미이지만, 엄밀히 말하자면 키워드란 개념만으로는 사회적 참여와 공유를 표현하기에는 어려움이 있다.

현재 일반적으로 사용되고 있는 태깅방법은

‘태그’라는 이름이 붙은 빈칸이나 작성하는 글에 이용자가 직접 달고 싶은 태그를 입력하는 것이다. 빈칸에는 하나 이상의 태그를 입력할 수 있고, 태그 사이의 구분은 일반적으로 띄어쓰기를 이용하며, 쉼표를 사용하는 경우도 있다. 태그는 보통 알파벳순으로 30-150개로 목록화되어 태그 클라우드를 통해 보여진다. 태그 클라우드 속의 각각의 태그는 가중치가 부여되며, 사용 빈도가 높은 태그는 글자의 크기나 색상 등 시각적으로 강조된다.

## 2.2 폭소노미의 특징과 장단점

기존에 사용하는 카테고리 분류방식이 웹 사이트 관리자가 사전에 정해 놓은 분류 카테고리에 의거하여 정보를 분류하고, 분류한 정보를 이용자들이 검색하는 일방향적 분류방식이 라면, 폭소노미는 인터넷의 이용자들이 자발적으로 웹상에 존재하는 정보에 의미를 부여하고, 부여된 의미를 기준으로 정보가 분류되는 양방향적인 분류방식이다(이경상 2008).

이러한 폭소노미의 장점으로는 우선 이용자가 원하는 대로 색인이 된다는 점이다. 정보검색에서 오랜 골칫거리로 논의되고 있는 것이 색인어와 탐색어의 불일치 문제임을 감안하면 이것의 의미는 가볍지 않다(Mathes 2004; 이재윤, 황혜경 2006).

다음으로 ‘우연한 발견’이다. 이용자들은 브라우징 과정에서 상호연결된 관련 태그를 통해서 의도하지 않은 유용한 정보를 발견할 수 있다. 일반적인 키워드 검색이나 브라우징은 이용자의 배경지식에 의존하기 때문에, 우연한 발견의 기회가 줄어들게 된다. 폭소노미에서는 다양

한 이용자에 의해 집적된 정보가 태그로 묶여 있기 때문에, 다른 이용자의 도움을 받아서 연구자의 좁은 시야를 보완할 수 있다(Mathes 2004; 이재윤, 황혜경 2006).

또 다른 장점으로 폭소노미는 각 태그 자체가 그 시기 이용자의 인지기구조에 의해 도출된 것이기 때문에 이용자 행태를 알 수 있는 통찰력을 제공한다. 폭소노미를 통해 이용자들의 이용행태와 선호관점, 흥미를 알 수 있다(Quintarelli 2005; Fichter 2006). 정보서비스 제공자는 폭소노미를 통해 이용자들의 선택 키워드, 선호 주제 등의 다양한 정보를 파악할 수 있으며, 이런 정보는 이용자의 만족도를 향상시키기 위하여 다양하게 활용될 수 있다.

또한 양질의 통제어휘를 개발하고, 유지하고 강화하는 것은 개발 시간과 이용자에게 필요한 선행 지식의 관점에서 보면 상당한 비용이 들지만, 폭소노미는 그렇지 않다. 폭소노미의 발전은 텍소노미의 웹 적용상의 어려움과 비용의 문제를 반영한다(Shirky 2004).

그 밖의 장점으로는 이용자의 관점, 배경, 편견 등과 무관하게 이용자들의 모든 어휘를 반영하는 포괄성, 새로운 개념을 쉽게 수용할 수 있는 신축성, 하나의 콘텐츠에 부여된 다양한 개념의 태그에 따라서 해당 콘텐츠가 여러 개의 주제에 분류될 수 있는 융통성 등을 들 수 있다.

그러나 폭소노미의 단점은 여러 이용자가 다양한 방식으로 콘텐츠에 태그를 추가하기 때문에 모호성과 비일관성이 내재되어 있고, 수평적 구조에 따른 계층정보 및 각 태그간의 연관 정보가 부족하다는 것이다. 이러한 단점은 모호성(ambiguity), 다의성(polysemy), 동의성(synonymy), 어휘변형(variation) 등 비통제

어휘가 가지는 문제와 거의 동일하다(Fichter 2006; Mathes 2004; Golder and Huberman 2006). 또한 악의적이거나 상업적인 태그도 통제되지 않기 때문에 폭소노미가 가진 무질서하고 부정확한 태그와 방대함과 같은 문제를 가중시킨다.

이러한 문제에도 불구하고, 폭소노미는 도서관을 비롯한 다양한 분야에서 새로운 서비스의 핵심개념으로 사용되고 있다. 폭소노미는 이용자의 협력과 공유로 만들어진 태그를 통해 정보를 유기적으로 연결할 수 있어 새로운 정보검색의 방법으로 간주된다. 폭소노미는 분류의 측면에서 이용자가 주도하는 접근방식의 변화로 이해할 수 있다. 이런 변화를 통해 도서관에서는 이용자 관심의 파악이 용이하고, 이용자의 정보요구가 충족될 수 있는 기반이 마련될 것이다.

### 2.3 정보서비스 영역에서 폭소노미의 적용

폭소노미는 정보 서비스 영역에 있어서 크게 웹 정보의 공유를 위한 소셜 북마킹 형태와 이용자에 의해 생산된 사진, 동영상과 같은 콘텐츠의 공유를 위한 서비스에 널리 적용되고 있다. 소셜 북마킹에 적용된 대표적인 사례는 Delicio.us, Furl, Simpy 등을 들 수 있고, 멀티미디어 콘텐츠 공유 서비스에 있어서의 적용 사례로는 Flickr가 있다(유시내 2007). 도서관계에서도 소셜 북마킹, 소셜 온라인 목록, 참고 문헌 공유 등 다양한 도서관 서비스에 폭소노미가 적용되기 시작하고 있다. 대표적인 사례로는 Pennsylvania의 PennTags와 Ann Arbor District Library의 SOPAC, Darien Public Library의 SOPAC 2.0, LibraryThing, CiteUlike,

Connotea가 있다.

국내에서도 폭소노미를 적용한 서비스를 제공하는 도서관이 있으며, 대표적으로 서울대학교, 성균관대학교, 한국과학기술정보연구원, 포항공과대학교, 국회도서관 등을 들 수 있다. 이 연구에서 분석된 A 대학도서관은 폭소노미를 도서관목록에 적용한 사례로서, 2007년 '이용자 참여형 웹서비스 기반 시스템'을 개발하였는데, 여기에는 태깅시스템, 서평 및 비평입력 시스템, 이용자 서지 콘텐츠 작성 및 편집 시스템이 포함된다.

A 대학도서관에서 2007년 9월부터 서비스되고 있는 태깅시스템은 이용자의 편의를 향상 시키고 동시에 지식의 공유를 통한 집단지성을 활용하기 위한 기반을 조성하기 위한 것으로, 이용자가 직접 참여하여 이용자에게 익숙한 형태의 분류가 가능하도록 하기 위해 도입되었다.

이용자들은 태그 입력 시스템의 입력화면을 통해 태그를 입력할 수 있고, 이용자가 입력한 태그는 '나의 태그'를 통해 볼 수 있다. 또한 이용자들이 입력한 각각의 태그는 '최근 태그'와 '전체 태그' 클라우드를 통해 브라우징이 가능하다. 이용자에게 태그 입력시 참고할 수 있는 가이드라인은 제공되지 않으며, 도서관에서도 이용자가 입력한 태그를 클라우드 상에 그대로 반영한다.

### 3. 폭소노미 태그의 형태적 특성 분석

이 장에서는 대학도서관 폭소노미 태그의 형태적 특성을 통제어휘의 개발 가이드라인에 근

거하여 분석하고 그 특성을 파악하였다. 주요 분석내용은 ‘용어의 형태’와 태그에 의해 표현된 ‘용어의 유형,’ 고유명사, 약어형, 외래어에서의 ‘용어의 선택’과 철자법, 특수문자에서의 ‘용어의 표기,’ ‘동형이의어’ 등이다.

### 3.1 용어의 형태

#### 1) 용어의 문법 형태

시소러스 개발 가이드라인에 의하면(7.2 용어의 형태), 색인어로는 명사 혹은 명사구가 사용될 수 있다. 동사나 형용사는 특정 주제분야에 따라서 색인어로 사용될 수 있으나 일반적인 시소러스에서는 거의 색인어로 사용하지 않는다.

한글태그의 문법적인 형태는 복합명사 50.18% (136건), 단일명사 31.37%(85건), 명사구 9.59% (26건), 동사와 형용사, 부사 각각 0.37%(1건) 순으로 조사되었다(표 1 참고). 조사대상 한글태그 분석 결과를 개발 가이드라인에서 제시한 문법적 형태의 비율로 나타내 보면, 단·복수 명사와 명사구 등 명사류가 전체 태그의 91.14% (247건)를 차지한다.

기타 7.75%(21건)에 해당하는 태그는 특정

한 문법적인 형태를 판단하기 어려운 경우이다. 21건 중 11건은 단순한 숫자 태그이고, 4건은 ‘위대한 기업에 투자하라’와 같은 문장 형태의 태그이고, 나머지 6건은 ‘陶? 明’, ‘이것도’, ‘자동차값’ 등과 같은 이용자만이 알 수 있는 태그와 오타로 인한 오류 태그이다.

NISO(6.4 용어의 문법적 형태)에서 제시하는 색인어의 문법적 형태는 명사나 명사구가 바람직하고, 부정사와 분사는 단독 용어로 사용해서는 안 되며, 행위를 나타낼 때는 반드시 명사나 동명사를 사용해야 한다. 명사구로는 형용사적 수식형(전위수식형) 명사구, 전치사적 수식형(후위수식형) 명사구가 있다. 형용사와 형용사구, 부사는 특별한 경우에만 색인어로 사용하며, 그 이외의 경우에는 단독 색인어로 사용하지 않는다.

〈표 2〉는 조사대상 영어태그의 문법적 형태의 분포이다. 복합명사와 단일명사 각각 34.64% (53건), 명사구-후위수식형 6.54%(10건), 명사구-전위수식형 5.23%(8건), 형용사 1.31%(2건), 전치사 0.65%(1건) 순으로 조사되었다. 이 분포는 사전의 품사를 기준으로 분석하되, 명사, 동사, 형용사 등 여러 품사로 사용될 수 있는 태

〈표 1〉 한글태그의 문법적인 형태

| 구 분  | 건(%)        |
|------|-------------|
| 복합명사 | 136(50.18)  |
| 단일명사 | 85(31.37)   |
| 명사구  | 26( 9.59)   |
| 동사   | 1( 0.37)    |
| 형용사  | 1( 0.37)    |
| 부사   | 1( 0.37)    |
| 기타   | 21( 7.75)   |
| 계    | 271(100.00) |

〈표 2〉 영어태그의 문법적인 형태

| 구 분       | 건(%)        |
|-----------|-------------|
| 복합명사      | 53(34.64)   |
| 단일명사      | 53(34.64)   |
| 명사구-후위수식형 | 10( 6.54)   |
| 명사구-전위수식형 | 8( 5.23)    |
| 동명사       | 0( 0.00)    |
| 동사        | 0( 0.00)    |
| 형용사       | 2( 1.31)    |
| 부사        | 0( 0.00)    |
| 전치사       | 1( 0.65)    |
| 기타        | 26(16.99)   |
| 계         | 153(100.00) |

그의 경우, 한 태그가 단독 사용되었을 때는 명사일 가능성이 높다고 판단하여 명사로 간주한 결과이다.

조사대상 영어태그의 문법적 형태를 NISO 가이드라인에서 제시된 문법적 형태의 비율로 나타내 보면, 명사와 명사구 등 명사류가 81.05% (124건)로 조사되었다. 영어태그에서 기타에는 한글태그와 같이 이용자가 알 수 있는 태그와 오타 태그 등 문법적인 형태를 알 수 없는 태그가 포함된다. 예를 들어 'ndwjrk,' 'wkfk,' 'toughkingca' 등이다. 영어태그에서 기타의 비율(16.99%)이 한글태그의 기타(7.75%)보다 높은 것은 '001.642 L497 v.1635' 등 숫자와 알파벳의 조합에 의한 태그(13.07%, 20건)가 다수 존재하기 때문이다.

## 2) 명사의 단·복수 형태

분석대상 한글태그 명사류의 총 247건 중 98.38%에 해당하는 243건이 단수형이고, 1.62% (4건)는 복수형이다. 가이드라인(7.2.1 명사와 명사구)에 의하면, 명사의 복수형은 색인으로 사용되어서는 안 된다. 조사 결과 복수형에 해

당하는 4건의 태그는 표제에 포함된 용어들이며, 예를 들어 '사람들'이 해당된다.

NISO(6.5 명사)에 의하면, 명사의 단·복수형의 채택여부는 시소러스를 사용할 색인의 형식에 영향을 받으며, 일반명사는 대부분 복수형을 취한다. 단·복수형과 관련하여, '몇개나?'라는 질문을 할 수 있는 가산명사는 복수로 표현해야 하고, '얼마나?'라는 질문을 할 수 있는 불가산명사는 단수형으로 표현해야 한다. 또한 NISO에서는 믿음, 활동, 감정, 특성, 학문영역, 개별실체는 단수형으로 표현되어야 한다고 제시하고 있다. 개별실체는 일반적으로 고유명사로 표현되는 개인과 단체, 지명을 의미한다.

영어태그의 명사류 총 124건 중 가산명사는 20.16%(25건)로, 불가산명사는 79.84%(99건)로 조사되었다. 가이드라인에 의하면 복수형으로 표현되어야 하는 가산명사 25건 중 24.00%(6건)만이 복수형으로 표현되었다. 단수형으로 표현되어야 하는 불가산명사인 질량명사(9건), 믿음(1건), 활동(15건), 감정(4건), 특성(13건), 학문영역(3건), 개별실체(54건) 모두 단수형으로 표현되었다.



### 3.2 용어의 유형

한글 시소러스 개발 가이드라인(7.1 색인어 일반)에서는 통제어휘에 의해 표현되는 개념을 크게 구체적 실체와 추상적 개념으로 구분한다. 구체적 실체에는 사물과 물리적인 부분, 재료가 해당되고, 추상적 개념에는 행위나 사건, 추상적 실체와 사물, 재료 및 행위의 특성, 학문분야 및 과학, 측정의 단위가 해당된다. 또한 인명, 단체명, 작품명 등 개별실체나 고유명사로 표현되는 특정 종류의 계층도 있다.

조사 도서관의 한글태그가 표현하고 있는 개념은 <표 3>과 같다. 높은 비율 순으로 나타내면, 추상적 실체와 사물, 재료 및 행위의 특성 32.47%(88건), 개별실체(고유명사) 30.63%(83건), 행위나 사건 12.18%(33건), 사물과 물리적인 부분 8.85%(24건), 학문분야 및 과학 8.49%(23건), 재료 0.74%(2건), 측정의 단위 0.37%(1건)이다.

한글태그에서 구체적 실체의 비율은 9.59%(26건), 추상적인 개념은 53.51%(145건), 개별실체(고유명사)는 30.63%(83건)로 시소러스 가이드라인에서 제시한 개념 유형이 전체의 93.73%

(254건)를 차지한다. 기타에 해당하는 6.27%(17건)는 <표 1>의 기타 21건 중 단순숫자 11건과 오타태그 6건이 포함된 것이다. 문장형태의 태그 4건은 작품명에 해당되어 개별실체로 처리되었다. 또한 가이드라인에 의하면, 각 색인어는 단일개념을 나타내야 한다. 조사대상 한글태그 271건 중 기타 17건을 제외한 254건의 단일개념 비율은 95.28%(242건)로, 복수개념 비율은 4.72%(12건)로 조사되었다.

NISO에서 제시하는 개념의 유형은 한글 시소러스 가이드라인과 유사하지만, 한글 가이드라인에서는 하나로 묶은 행위나 사건을 NISO에서는 구분하고 있다(6.3.2 개념의 유형). 영어태그의 개념 유형 분포는 <표 4>와 같다. 높은 비율 순으로 나타내 보면, 개별실체(고유명사) 35.30%(54건), 사물과 물리적인 부분 18.30%(28건), 행위나 과정 14.38%(22건), 특성과 상태 11.76%(18건), 학문분야와 주제분야 1.96%(3건), 재료 1.31%(2건) 순으로 나타났으며, 사건과 측정단위를 제외한 6가지의 유형이 분포하였다.

영어태그에서 구체적 실체의 비율은 19.61%(30건), 추상적인 개념은 28.10%(43건)이었고, 개별실체(고유명사)는 35.30%(54건)로, 시소

<표 3> 한글태그의 개념 유형

| 구 분                     | 건(%)        |
|-------------------------|-------------|
| 사물과 물리적인 부분             | 24( 8.85)   |
| 재료                      | 2( 0.74)    |
| 행위나 사건                  | 33(12.18)   |
| 추상적 실체와 사물, 재료 및 행위의 특성 | 88(32.47)   |
| 학문분야 및 과학               | 23( 8.49)   |
| 측정의 단위                  | 1( 0.37)    |
| 개별실체(고유명사)              | 83(30.63)   |
| 기타                      | 17( 6.27)   |
| 계                       | 271(100.00) |

〈표 4〉 영어태그의 개념 유형

| 구 분         | 건(%)        |
|-------------|-------------|
| 사물과 물리적인 부분 | 28(18.30)   |
| 재료          | 2( 1.31)    |
| 행위나 과정      | 22(14.38)   |
| 사건          | 0( 0.00)    |
| 특성과 상태      | 18(11.76)   |
| 학문분야와 주제분야  | 3( 1.96)    |
| 측정단위        | 0( 0.00)    |
| 개별실체(고유명사)  | 54(35.30)   |
| 기타          | 26(16.99)   |
| 계           | 153(100.00) |

리스 가이드라인에서 제시한 개념 유형이 전체의 83.01%(127건)를 차지한다. 영어태그의 기타는 16.99%(26건)로 조사되었다. 이와 같이 기타의 비율이 높은 것은 개념의 유형을 파악할 수 없는 무의미한 태그, 숫자와 영어의 단순한 조합 태그 등의 비율이 높기 때문이다. 기타의 예로는 '??,' 'gpdf,' 'toughkingca' 등을 들 수 있다.

또한 단일개념의 비율을 보면, 영어태그는 기타 26건을 제외한 총 127건 중 92.91%에 해당하는 118건이 단일개념이었고, 7.09%에 해당하는 9건이 복수개념이었다.

### 3.3 용어의 선택

시소러스 가이드라인에서는 동일한 개념이 둘 혹은 그 이상의 동의어로 표현될 경우에는 일관성 있는 색인 작성을 위하여 이들 용어 중 하나를 우선어로 선택하도록 하며, 이용자가 어느 비우선어에서도 우선어를 참조할 수 있도록 참조를 작성해야 한다.

#### 1) 고유명사

시소러스 가이드라인에 의하면(7.4.6 기관단체명, 개인명, 지명 등), 시소러스의 목적이나 특정 데이터베이스에 종속된 시소러스의 경우에는 인명, 회사명, 단체명, 화합물명, 지명(국명, 행정구역명, 도시명, 섬, 산, 사찰 등), 작품명, 사건명, 법률명, 행사 경기대회명, 행정기구명, 건축물명 등의 고유명사가 색인어로 사용될 수 있다.

한글태그와 영어태그의 고유명사 사용비율은 〈표 5〉와 같다. 한글태그의 고유명사는 〈표 3〉의 개별실체(83건)와 사건명(2)을 합한 85건이다. 한글태그 총 85건 중 표제를 포함한 작품명 69.41%(59건), 개인명 18.82% (16건), 단체명과 지명 각각 4.71%(4건), 사건명 2.35% (2건) 순으로 조사되었다. 영어태그에서도 작품명 77.78%(42건), 개인명 16.67%(9건), 행사 경기대회명 3.70%(2건), 단체명 1.85%(1건) 순으로 조사되었다.

고유명사에 대한 가이드라인에 의하면(한글 시소러스 가이드라인 7.4.6과 NISO 6.6.7), 고유명사는 이용자에게 가장 친근한 형식을 용어

〈표 5〉 한글태그와 영어태그에서의 고유명사

| 구 분      | 한글태그       | 영어태그       |
|----------|------------|------------|
| 작품명      | 59(69.41)  | 42(77.78)  |
| 개인명      | 16(18.82)  | 9(16.67)   |
| 단체명      | 4( 4.71)   | 1( 1.85)   |
| 지명       | 4( 4.71)   | 0( 0.00)   |
| 사건명      | 2( 2.35)   | 0( 0.00)   |
| 법률명      | 0( 0.00)   | 0( 0.00)   |
| 회사명      | 0( 0.00)   | 0( 0.00)   |
| 화합물명     | 0( 0.00)   | 0( 0.00)   |
| 행사 경기대회명 | 0( 0.00)   | 2( 3.70)   |
| 건축물명     | 0( 0.00)   | 0( 0.00)   |
| 계        | 85(100.00) | 54(100.00) |

로 지정하여야 하고, 친근한 정도가 동일하다면 속명보다는 공식명칭을 선택하는 것이 바람직하다. 개인명의 선택은 원명을 사용하거나 일반적으로 널리 알려진 명칭을 우선어로 선택하도록 한다. 또한 고유명사가 시소러스에 포함되면 그 이름의 표기는 목록규칙과 같은 규칙을 따라야 한다.

이 연구에서는 작품명은 문헌상의 사례(literature warrant)에 기반하고 있기 때문에 보편적이라고 보았고, 작품명을 제외한 태그(한글태그 26건, 영어태그 12건)에 대한 보편성 여부는 단체명과 지명, 행사명, 사건명은 공식명칭 사용여부로 확인하였고, 원어명이 아닌 개인명은 사전의 등재여부로 확인하였다. 한글태그 26건 중 공식명칭을 사용한 태그는 6건으로, 개인명의 경우 원명을 사용하거나 원명이 아닌 경우 사전에 등재된 태그는 10건으로 조사되었다. 한글태그에서 가이드라인과 일치하는 고유명사 선택은 총 85건 중 88.24%(75건)로 조사되었다. 영어태그 12건 중 공식명칭을 사용한 경우는 1건, 개인명의 경우 원명을 사용하거나

원명이 아닌 경우 사전에 등재된 태그는 4건으로 조사되었다. 가이드라인과 일치하는 영어태그의 고유명사 선택은 총 54건 중 87.04%(47건)로 조사되었다.

개인명의 표기를 보면, 성과 이름을 순서대로 기술한 개인명 한글태그는 16건 중 8건(50.00%), 영어태그는 9건 중 0건(0.00%)이다. 조사 대상 태그에서는 가이드라인에서 제시하고 있는 것과 같은 선택된 태그와 선택되지 않은 태그간의 참조는 작성되지 않는다.

2) 약어와 두문자어

분석대상 한글태그의 경우 준말의 사용 비율이 1.85%(5건), 영어태그에서는 19.61%(30건)로 조사되었다. 약어 혹은 두문자어에 대한 가이드라인(한글 시소러스 가이드라인 7.2.5와 NISO 6.6.3)에서는 약어형은 완전형이 거의 사용되지 않을 정도로 보편적인 경우에만 선택하여 사용할 것을 권고한다.

이 연구에서는 한글과 영어태그의 보편성을 사전의 등재여부로 확인하였다. 한글태그 1.85%

(5건), 영어태그 12.42%(19건)가 사전에 등재되지 않은 것으로 조사되었다. 가이드라인에 의하면, 사전에 수록되지 않은 이 태그들은 완전형으로 선택될 것이며, 사용된 약어형과 완전형간의 상호참조가 제시될 것이다.

한글태그와 영어태그의 약어형에서는 이용자 임의로 축약한 약어형 태그가 발생하였다. 한글태그에서는 새로운 한국음악사를 줄인 '새한음'과 포스트모더니즘을 줄인 '포스트모던', 컴퓨터를 줄인 '컴', 레포트를 줄인 '레포트' 등이, 영어태그에서는 journal을 축약한 'jorn,' letters를 축약한 'Lett,' Microengineering을 축약한 'microeng' 등이 해당된다.

또한 의미가 모호한 태그의 사례도 발견되었다. 예를 들어 'Lett'를 사전에서 찾아보면, 그 의미는 '발트 해 연안의 주민,' '레트 말'이다. 조사 대상 태그 'Lett'는 'letters'를 축약한 것이다. 또한 태그 'INF'는 중거리 핵전력(intermediate range nuclear forces)의 약어이지만, 조사 대상 태그 'INF'는 웹서비스에 관한 자료에 부여된 약어로서 그 의미가 모호하다.

### 3) 외래어와 신조어

한글 시소러스 가이드라인에 의하면(7.4.1 외래어와 번자, 신조어), 외래어의 경우 외래어가 더 넓게 쓰인다면 이것이 우선어로, 번역어가 잘 제정되었으면 그것이 우선어가 된다. 한글태그 총 271건 중 외래어의 사용은 18.08%(49건)이다.

외래어의 보편성 여부 역시 사전의 등재 여부로 조사하였다. 조사 결과 49건의 태그 중 사전등재 태그는 미니멀리즘(minimalism), 페르소나(persona) 등 40건이었고, 미등재 태그는 예를

들어 모더니티(modernity) 등 9건으로 전체 태그의 3.32% 이었다.

또한 널리 알려진 다른 용어가 없을 때는 신조어와 속어도 색인어로 채택될 수 있다(7.4.3 속어와 특수전문용어). 국립국어원의 신어사전에 의해 한글태그를 조사한 결과 신조어 사례로는 레시피 1건(0.37%)이 조사되었다.

## 3.4 용어의 표기

### 1) 철자법

한글태그의 경우 일상적인 철자법에서 벗어나는 경우는 주로 오타 태그와 준말 태그로서 분석대상 271건 중 4.06%(11건)로 조사되었는데, 예를 들어 '레포트,' 'ㅇ,' '자동차점,' '짱짱장!!!,' '이것도,' '티비' 등이 해당된다.

NISO 가이드라인(6.6.2 철자법)에 의하면, 용어의 철자법은 문헌상의 사례와 전거를 따르도록 하고 있다. 즉 문헌상의 사례에 의해 가장 널리 사용되고 있는 단어의 철자법 혹은 전거에 의해 정착된 사전이나 용어사전의 관례를 따라야 한다. 영어태그에서 철자법에서 벗어나는 태그는 총 153건 중 29.41%(45건)로 조사되었다. 여기에는 한글태그와 비슷하게 오타태그, 약어 태그로서 예를 들어 'jorn,' 'ndwirk,' 'wkfk,' 'toughkingca,' 'zz' 등이 해당된다.

### 2) 특수문자

특수문자에 대한 가이드라인에 의하면(한글 시소러스 가이드라인 7.5.5와 NISO 6.7.1), 표목형식의 일관성과 배열을 단순화하기 위하여 용어 내에서 특수기호의 사용을 최소화하며, 그런 문자가 필수적일 때는 정착된 철자법을

따른다. 원괄호(( ))는 괄호한정어이거나 등록상표 지시기호이거나 그 일부일 때만 사용한다. 불임표(-)는 약어, 등록상표, 화학제품명, 고유명사 등의 일부일 때는 용어에 포함된다.

한글태그의 특수문자 사용여부를 보면, 총 271건 중 5.90%(16건)에서 특수문자가 사용되었다. 한글태그에서 원괄호(( )) 4건, 빗금(/) 3건, 각괄호([ ]) 2건, 의문부호(?) 2건, 온점(.) 1건, 느낌부호(!) 1건, 밑줄문자(\_) 1건, 쌍점(:) 1건, 불임표(-) 1건이 사용되었다.

영어태그의 특수문자는 총 153건 중 26.80%(41건)에서 사용된 것으로 조사되었다. 온점 25건, 불임표 5건, 쌍점 4건, 빗금 3건, 앰퍼센드(&) 3건, 의문부호 1건이 사용되었다. 온점의 비율이 높은 것은 'url'과 '청구기호' 태그가 다수 포함되었기 때문이다. 또한 특수기호의 사용에서도 비일관성이 발견되었다. 태그에서 불임표는 청구기호를 연결하기 위해, 단어를 연결하기 위해, 주제의 연도를 제한하기 위해 사용되었다.

### 3.5 동형이의어

동형이의어는 표기는 같으나 어원과 의미가 상이한 단어를 말한다. 시소러스 가이드라인(7.3 동형이의어)에 따르면, 동형이의어를 구별하기 위해 한글 한정어의 사용을 권고한다. 한글태그에서 동형이의어의 비율은 43.17%(117건)였다. 예를 들어 태그 '수영'은 ① 스포츠나 놀이로서 물속을 헤엄치는 일, ② 변방을 지키는 수자리의 진영, ③ 나무의 그림자, ④ 승부, ⑤ 건물이나 구조물 따위를 고치고 세움의 의미를 가진다.

NISO 가이드라인(6.2.1 동형이의어) 역시 동형이의어는 가능하면 색인어로 선정하지 말고, 애매함을 구별하기 위해 한정어나 범위주기를 사용할 것을 권고한다. 영어태그 153건 중 동형이의어의 사용 비율은 13.07%(20건)이고, 이 중 8건은 약어형 태그이며, 예를 들어 ABC, ODA, ACT 등이 해당된다. 로마자의 첫 세글자인 ABC는 다의성을 가진 태그로 미국방송사(American Broadcasting Company), 신문잡지의 발행부수감사기구(Audit Bureau of Circulations), 오스트레일리아 방송위원회(Australian Broadcasting Commission), 화생방(atomic, biological and chemical), 활동기준원가계산(Activity-Based Costing), 일본 아사히 방송(Asahi Broadcasting Corporation)을 의미하기도 한다.

또한 예를 들어 태그 'eng'는 4회 태깅되었는데, 영국(England)을 의미하는 사례가 2건, 영어(English)를 의미하는 사례가 2건이었다. 가이드라인에 의하면, 이런 태그의 의미를 명확히 하기 위해 한정어나 범위주기가 사용된다.

### 3.6 분석결과

분석결과 나타난 폭소노미 태그의 형태적 특성을 종합적으로 살펴보면 다음과 같다.

첫째 태그의 문법적인 형태를 보면, 시소러스 가이드라인에서 제시하고 있는 문법 형태를 한글태그에서는 91.14%, 영어태그에서는 81.05% 사용하였다(표 1, 2 참고). 명사류 이외의 비율은 한글태그 8.86%, 영어태그 18.95% 이다. 이 중 문법적인 형태를 알 수 없는 태그는 한글태그 7.75%, 영어태그 16.99% 이다. 여기에는 단

순한 숫자 태그, 문장 형태의 태그, 오타 태그, 이용자만이 알 수 있는 태그 등이 포함된다. 특히 영어태그에서는 청구기호로 판단되는 태그가 13.07% 존재하였다.

명사의 단·복수 형태를 보면, 한글태그에서는 명사의 단수형을 98.38% 사용하였으며, 복수형을 사용한 1.62%의 태그는 표제에 포함된 용어를 사용한 경우이다. 영어태그에서는 NISO 가이드라인에 의하면 복수형으로 표현해야 하는 가산명사 중 24.00%만이 복수형으로 표현되었고, 나머지 76.00%는 단수형으로 표현되었다. 또한 단수형으로 표현해야 할 불가산 명사는 모두 단수형으로 표현되었다.

둘째 시소러스 가이드라인에서 하나의 색인어는 하나의 개념을 나타내야 한다고 제시하고 있는데, 한글태그에서 단일개념의 사용 비율은 95.28%, 영어태그에서는 92.91%로 조사되었다. 또한 시소러스 가이드라인에서 제시하고 있는 개념 유형을 한글태그에서는 93.73%, 영어태그에서는 83.01% 사용하였다. 한글태그에서는 추상적 개념의 비율이, 영어태그에서는 개별실체의 비율이 높았다. 무의미한 태그는 한글태그 6.27%, 영어태그 16.99% 존재하였다(표 3, 4 참고).

셋째 용어의 선택에서, 우선 고유명사를 보면 작품명의 비율이 한글태그(69.41%), 영어태그(77.78%)에서 높았으며(표 5 참고), 이는 도서관 폭소노미 태그의 특징으로 볼 수 있다. 고유명사의 선택에서 가이드라인과 일치하는 한글태그는 88.24%, 영어태그는 87.04%로 조사되었고, 개인명의 표기에서 가이드라인과 일치하는 한글태그는 50.00%, 영어태그는 0.00%로 조사되었다. 개인명의 경우 성만을 포함한

태그, 성과 이름을 띄어쓰기 없이 붙인 태그, 이름 성의 순서로 나열한 태그 등 비일관성이 존재하였다.

약어와 두문자어의 사용비율을 보면, 한글태그에서는 준말이 1.85%, 영어태그에서는 약어형이 19.61%가 사용되었다. 가이드라인에 의하면 이 중 보편적이지 않다고 보여지는 한글태그 1.85%와 영어태그의 12.42%는 완전형으로 선택될 것이다. 특히 이용자 임의로 축약한 약어형 태그, 원래 의미와는 다른 의미를 갖는 약어형 태그가 존재하였다. 이러한 약어형은 폭소노미 태그의 모호성을 증가시키는 역할을 한다.

또한 한글태그를 중심으로 조사한 외래어의 사용비율을 보면 18.08%이며, 사전에 등재되지 않을 정도로 보편적이지 않다고 판단되는 외래어의 비율은 3.32%이다. 조사 대상 한글태그의 신조어 비율은 0.37%로 조사되었으며, 이는 통상적으로 폭소노미에서 언급되는 이용자들의 신조어 유입이 빠르다는 것과는 일치되지 않는 결과이다. 시소러스 가이드라인에서는 일관성 있는 색인 작성을 위하여 우선어와 비우선어간의 참조를 작성하도록 하고 있는데, 폭소노미 태그에서는 이런 참조는 작성되지 않고 있다.

넷째 용어의 표기를 보면, 일상적인 철자법에서 벗어나는 경우로는 한글태그에서 4.06%, 영어태그에서 29.41%로 조사되었고, 이런 태그들의 대부분은 무의미한 태그나 오타에 의한 태그가 대부분이다. 시소러스 가이드라인에서는 표현형식의 일관성과 배열, 검색을 위해 특수문자의 사용을 최소화할 것을 제시하고 있는데, 특수문자는 한글태그 5.90%, 영어태그 26.80%에서 사용되었다. 또한 하나의 특수기호가 다양

한 의미로 사용되기도 하였다. 예를 들어 붙임 표는 청구기호를 연결하기 위해, 단어를 연결하기 위해, 주제의 연도를 제한하기 위해 사용되었다.

다섯째 동형이의어를 보면, 한글태그에서의 비율은 43.17%로 조사되었고, 영어태그에서는 13.07%로 조사되었다. 동형이의어의 분포 비율은 한글태그가 영어태그보다 높았다. 동형이의어를 명확하게 구별하지 않으면 용어의 의미가 다르게 되어 색인과 검색시 잡음률이 높아지게 된다.

조사대상 도서관의 폭소노미 태그 분석 결과, 시소러스 개발 가이드라인과의 부합정도를 백분율로 살펴보면 다음과 같다. 태그의 문법적인 형태는 한글태그의 91.14%, 영어태그의 81.05%, 명사 태그의 단·복수형에서는 한글태그의 단수형 98.38%, 영어태그의 단수형 100%, 복수형 24.00% 부합되었다. 태그의 단일개념은 한글태그 95.28%, 영어태그 92.91%, 태그의 개념 유형은 한글태그 93.73%, 영어태그 83.01% 부합되었다. 용어의 선택에서 고유 명사는 한글태그 88.24%, 영어태그 87.04% 부합되었고, 완전형 사용은 한글태그 98.15%, 영어태그 87.58%, 보편적인 외래어의 사용은 한글태그 96.68% 부합되었다. 용어의 표기에서는 인정되고 있는 철자법의 사용은 한글태그는 95.94%, 영어태그 70.59%, 특수문자의 사용에서 한글태그는 94.10%, 영어태그는 73.20% 부합되었고, 개인명의 표기는 한글태그 50.00%, 영어태그 0.00% 부합되었다. 태그의 동형이의어에서는 한글태그 56.83%, 영어태그 86.93% 부합되었다.

#### 4. 결론 및 제언

이 연구에서는 디지털도서관에서 폭소노미 태그의 활용가능성을 살펴보기 위하여, 국내 대학도서관에서 서비스되고 있는 폭소노미 태그의 형태적 특성을 시소러스 개발 가이드라인과 비교하여 분석하였다.

조사대상 도서관의 폭소노미 태그 분석 결과, 대학도서관 폭소노미 태그는 이용자들이 어떠한 제약 없이 자유롭게 등록하는 것이지만, 시소러스 가이드라인의 여러 형태적 분석 기준에 90% 이상 부합하는 것을 알 수 있었다. 예를 들어 한글태그의 문법적인 형태, 한글·영어태그의 명사 단수형 사용, 한글·영어태그의 단일개념의 사용, 한글태그의 개념 유형 사용, 한글태그의 완전형 사용, 한글태그의 외래어 사용, 한글태그에서 보편적인 철자법의 사용, 한글태그의 특수문자 미사용 등이 해당된다. 이러한 분석결과는 이용자들이 협력하여 만든 폭소노미 태그가 대학도서관 목록에서의 활용가능성이 높다는 것을 나타낸다.

한편 분석결과 시소러스의 가이드라인에 부합되지 않는 부분도 발견되었다. 이러한 부합되지 않은 부분은 폭소노미 태그의 단점인 모호성·다의성과 관련이 있고, 태그의 공유와 재사용의 관점에서 고려할 때, 태그 검색의 정확률과 재현율에 증대한 영향을 미치게 된다. 예를 들어 태그의 다의성에 영향을 주는 동형이의어가 한글태그에서 43.17% 사용된 것으로 조사되었다. 검색결과에 영향을 주는 가산명사의 단수형이 영어태그에서 76.00% 사용되었고, 태그의 신뢰성과 관련있는 무의미하거나 모호한 오류태그가 영어태그에서 18.95% 사용되었

다. 특히 용어의 표기에서 보편적인 철자법에서 벗어나는 태그가 영어태그에서 29.41%, 모호하고 일관성 없는 특수문자가 영어태그에서 26.80% 사용되었다. 가이드라인과 일치하지 않는 개인명의 표기는 한글태그 50.00%, 영어태그 100%로 나타났다.

폭소노미는 개방과 참여, 공유라는 특징을 갖는 Library 2.0의 핵심 개념으로 이미 많은 기관에서 사용되고 있다. 이용자들의 참여와 협력으로 만들어진 폭소노미는 인지구소에서 도출된 다양한 접근점을 통해 정보를 유기적으로 연결할 수 있어 효율적인 정보검색을 제공할 수 있는 방안이다. 또한 정보조직의 관점에서 보면, 폭소노미는 사서 주도적인 접근 방식에서 이용자 주도적인 접근방식의 변화로 이해할 수 있으며, 이를 통해 이용자의 정보요구가 충족될 수 있는 기반이 마련될 것이다. 이 연구의 분석결과, 폭소노미 태그의 형태적 특성은 시소러스 가이드라인의 기준에 부합하여 도서관 목록에의 활용가능성은 높으나, 태그의 검색과 공유를 어렵게 하는 부합하지 않는 몇몇 부분이 있다는 것을 알 수 있었다.

폭소노미가 대학도서관 목록의 상호작용을 증진시킬 수 있는 도구로 활용되기 위해서는 폭소노미의 의미를 훼손하지 않는 범위내에서 태그를 최적화하는 방안이 필요하다. 이는 도서관에서 폭소노미 태그의 가치를 높이는 것과 더불어 개인적인 메타데이터의 가치를 높이는

방안이 될 것이다.

이를 위해 태그 부여와 관련된 최소한의 가이드라인을 고려할 수 있다. 가이드라인에는 특히 동형이의어, 철자법, 특수문자의 사용, 무의미한 태그 생성, 태그의 선택, 표기 등에 관한 내용이 포함되어야 한다. 무엇보다 동형이의어의 다의성을 해결하기 위한 방안으로 시소러스 개발 가이드라인에서 제시하고 있는 한정어와 범위주기를 고려해 볼 필요가 있다. 권위 있는 온라인 사전과의 연계 서비스를 제공하여, 동형이의어 태그의 의미를 명확히 하고, 약어형에 우선하는 완전형, 적절한 철자법, 외래어 선택과 표기 등에 도움을 줄 수 있다. 또한 태깅시스템 하단에 통제어휘를 활용한 추천태그, 인기태그, 오류태그 검출 프로그램 등 기술적으로 가능한 도구들을 활용할 수 있다.

이 연구는 폭소노미 태그와 전통적인 통제어휘 색인과의 부합 정도를 실증적으로 제시함으로써 폭소노미를 이해하고 디지털도서관에서 폭소노미 태그의 활용 가능성을 검토하였다는데 의의가 있다. 그러나 이 연구가 특정 대학도서관의 6개월간의 폭소노미 태그를 형태적 측면에서 분석했다는 점에서 연구의 결과를 전체 도서관의 폭소노미 태그의 특성으로 일반화하기는 어려울 것이다. 향후 다양한 도서관의 폭소노미 태그를 대상으로 연구가 진행될 필요가 있고, 형태적 측면이외의 다양한 관점에서 접근한 연구가 필요하다.



## 참 고 문 헌

- 구본홍. 2007. 21세기 기업문화 차원을 높이자(하). 『웹2.0시대의 조직문화』, [인용 2008.9.15]. <<http://www.naeil.com/>>.
- 김수현. 2007. 이용자 참여형 웹서비스 - 태깅 (Tagging) 시스템 『KORSA 학술 Workshop 자료집』, [cited 2008.9.1]. <<http://www.korsa.or.kr/>>.
- 박찬표, 반영환, 정지홍. 2007. 웹 이미지의 협력적 폭소노미에 관한 연구. 『한국디자인학회 봄 학술발표대회 논문집』, 5: 58-59.
- 안희영, 이현주. 2007. 웹 환경에서 텍스트 표현 방식에 따른 가독성에 관한 연구. 『한국디자인학회 봄 학술발표대회 논문집』, 5: 86-87.
- 유시내. 2007. 『기록 검색 서비스 개선을 위한 폭소노미 도입에 관한 연구』. 석사학위논문, 명지대학교 기록관리대학원, 기록관리학과.
- 이경상. 2008. 군중에 의한 정보분류체계. [인용 2008.9.10]. <<http://kslee7.tistory.com/11>>.
- 이상현. 2002. 『국립중앙도서관 주제명표목표 개발』. 서울 : 국립중앙도서관.
- 이재운, 정도현. 2008. 폭소노미 태그 사용 패턴 분석: 통제어휘 및 비통제어휘와의 비교 『한국정보관리학회 학술논문발표집』, 15: 21-26.
- 이재운, 황혜경. 2006. 이용자 생성 메타데이터에 관한 연구. 『정보관리연구』, 37(3): 1-24.
- 이정미. 2007. 폭소노미의 개념적 접근과 웹 정보 서비스에의 적용. 『한국비블리아학회지』, 18(2): 141-159.
- 장석중. 2007. 성균관대학교 학술정보관의 Library2.0 적용 사례. 『KORSA 학술 Workshop 자료집』, [cited 2008.9.1]. <<http://www.korsa.or.kr/>>.
- 조순영. 2002. 『교육학술정보 시소러스 개발』. 한국교육학술정보원.
- 최석두. 2000. 『시소러스 개발 지침』. 서울: 한국데이터베이스진흥센터, 이화여자대학교.
- 최석두. 2007. 『과학기술 분야 시소러스 구축 연구』. 대전: 한국과학기술정보연구원.
- Dempsey, L. 2003. "The recombinant library: portals and people." *Journal of Library Administration* 39(4): 103-136.
- Fichter, D. 2006. "Intranet applications for tagging and folksonomies." *Online*, 30(3): 43-45.
- Golder, S. A., and B. A. Huberman. 2006. "Usage patterns of collaborative tagging systems." *Journal of Information Science* 32(2): 198-208.
- Ketchell, D. S. 2000. "Too many channels: making sense out of portals and personalization." *Information Technology and Libraries* 19(4): 175-179.
- Mathes, Adam. 2004. Folksonomies-Cooperative Classification and Communication Through Shared Metadata. *Computer Mediated Communication*. [cited 2008.

- 9,11]. <<http://www.adammathes.com/academic/computer-mediated-communication/folksonomies.html>>.
- NISO. 2005. Guidelines for the construction, format, and management of monolingual controlled vocabularies. ANSI/NISO Z39.19-2005. [cited 2008.9.11]. <[www.niso.org/standards/resources/Z39-19-2005.pdf](http://www.niso.org/standards/resources/Z39-19-2005.pdf)>.
- Quintarelli, E. 2005. Folksonomies: Power to people. [cited 2008.9.12]. <<http://www.iskoi.org/doc/folksonomies.htm>>.
- Shirky, C. 2004. Folksonomy. [cited 2008.9.11]. <<http://www.corante.com/many/archives/2004/08/25/folksonomy.php>>.
- Spiteri, Louise F. 2007. "The Structure and Form of Folksonomy Tags: The Road to the Public Library Catalog. Information Technology and Library." *Information Technology and Libraries* 26(3): 13-25.
- Vander wal, Thomas. 2007. Folksonomy. [cited 2008.10.15]. <<http://vanderwal.net/folksonomy.html>>.
- Wikipedia. 2008. [cited 2008.9.20]. <<http://www.wikipedia.org/>>.