

# 학술정보서비스의 폭소노미 분석 연구\*

## An Analysis of the Foxonomy Constructed at Research Information Service and Future Perspectives

조재인(Jane Cho)\*\*

### 목 차

1. 서론	3.3 폭소노미 태그의 어휘상의 문제와 부분적 태그 통제 노력
2. 폭소노미의 의의	4. 학술정보서비스의 폭소노미 분석
2.1 폭소노미의 의미와 기능	4.1 분석 방법의 설계
2.2 폭소노미와 텍소노미	4.2 폭소노미 분석 결과
2.3 폭소노미의 한계와 문제점	4.3 학술연구정보서비스 폭소노미의 특징
3. 폭소노미 구축 사례	5. 학술연구정보서비스의 폭소노미 운영을 위한 제언
3.1 포털형 폭소노미	6. 결론
3.2 OPAC형 폭소노미	

### 초 록

폭소노미는 전통적인 분류체계인 텍소노미에 대비되는 용어로 사람들에 의한 분류법이란 의미로 해석된다. 인터넷 컨텐츠를 대상으로 개방적으로 조직된 레이블로, 소셜 태깅의 결과로 만들어진 협력형 색인이다. 폭소노미는 다양한 이용자가 다양한 관점에서 생성한 태그를 통해 우연한 접근의 기회를 제공하며, 동일한 관심사를 가진 타인과 소셜 네트워크 형성을 위한 연결 고리가 된다. 본 연구에서는 폭소노미의 의의를 짚어보고, 전통적인 텍소노미 시스템을 어떻게 보완할 수 있는지, 그리고 그 문제점과 한계는 무엇인지 고찰해 본다. 더불어 학술정보서비스에 부여되고 있는 폭소노미 태그를 분석하여 그 특징을 살펴보고 폭소노미 시스템의 운영 방향을 제시한다.

### ABSTRACT

In contrast to traditional taxonomy, folksonomy is generated not only by experts but also by creators and consumers of the content. Folksonomy is the practice and method of collaboratively creating and managing tags to annotate and categorize content. It is also known as collaborative tagging or social indexing. Folksonomy is also used to link to create social network that connect people to people who share same interest. Folksonomy users can generally discover the contents by which the tag sets of another user who tends to interpret contents in a way that makes sense to them. Firstly, this study consider the significance and some critical issues about folksonomy. Secondly, analyze special features of Korean academic site's folksonomy, which is managed by academic information site. Accordingly consider the directions of development about folksonomy system.

키워드: 폭소노미, 텍소노미, 소셜 태그, OPAC, 학술정보서비스

Folksonomy, Taxonomy, Social Tag, OPAC, Research Information Service

\* 본 연구는 2008년도 인천전문대학 교내연구비 지원에 의하여 수행되었음.

\*\* 시립인천전문대학 문헌정보학과 전임강사(chojane@icc.ac.kr)

논문접수일자: 2008년 11월 3일 최초심사일자: 2008년 11월 27일 게재화정일자: 2008년 12월 4일

## 1. 서 론

OCLC가 6개국 정보 소비자를 대상으로 한 조사에서 응답자의 84%는 정보 탐색을 시작할 때 검색 엔진을 이용한다고 하였다(OCLC 2005). 이러한 현상은 도서관 목록 시장을 위축시키고 있으며, 그로 인해 목록은 이미 제품의 라이프 사이클로 본다면 쇠퇴기에 접어들었다고 할 수 있다(Calhoun 2006). 미의회도서관 마кам(Marcum 2005)은 고비용의 집약적 노력을 요구하는 목록 작성 작업의 경제성에 대해 지적하고, 더불어 기술 목록에 치중하여, 자원의 식별과 평가 그리고 주제 분석을 등한시하는 도서관 목록의 위기와 장래에 대한 우려를 표현하였다. 이러한 우려는 카페 목록, 자동 처리 강화 등 도서관 목록의 비용 효과를 높이는 다각적 방법에 대한 고민과 더불어 이용자들의 자발적 태그 부여 행위로 형성될 수 있는 자연적 질서에 대해서도 논하게 하였다. 그러한 논의 중에 폭소노미가 관심을 받기 시작하였는데, 사람들에 의한 분류법(Folk+order+nomos)이라는 의미로 해석되는 폭소노미는 도서관 자원이 목록에 의해 체계화되듯이 군중의 힘에 의해 인터넷 자원에 자연적 질서를 형성시킬 수 있다.

폭소노미는 전통적인 분류체계인 ‘텍소노미’에 대치되는 용어이다. 웹페이지, 이미지 등 인터넷 컨텐츠를 대상으로 개방적으로 조직된 레이블로, 소셜 태깅의 결과로 만들어진 협력형 색인이다. 태깅 작업은 기본적으로 분류와 검색을 용이하게 하기 위한 개인적인 행위이지만, 하나의 컨텐츠에 대해 다수의 사용자가 다양한 태그를 부여함으로써 ‘소셜’ 기능이 구현된다. 소셜 태깅은 엄청난 규모로 성장하고 있는 인

터넷 컨텐츠를 다양한 관점으로 군집시킬 수 있으며, 세맨틱을 공유한 사람들 간에 소셜 네트워크 형성을 위한 연결 고리가 된다.

폭소노미는 일반적으로 딜리셔스(<http://delicio.us>)와 같은 포털형 소셜 북마크 사이트에서 형성되지만, Connotea([www.connotea.org](http://www.connotea.org)), CiteULike([www.citeulike.org](http://www.citeulike.org))와 같은 학술연구자 전용의 북마크 사이트에서 형성되기도 하며, 도서관 OPAC에 적용되어 주제명 표목을 대신한 검색 도구로 이용되고 있다.

폭소노미는 계층적이지도 배타적이지도 않고 학습이 불필요하여 이용자 누구나가 부여할 수 있지만, 유의어, 다의어, 개인적 태그, 스팸 태그와 같이 검색 시스템에 치명적 오류를 야기할 수 있는 단점도 가지고 있다. 통제되지 않은 태그는 검색의 정확성을 저해하고 동일 관심사를 가진 이용자들을 어긋나게 만든다. 비록 폭소노미를 둘러싼 논의의 쟁점은 군중의 협업에 의한 집단 지성의 창출이지만, 규모의 힘을 기대하기까지 감내해야만 하는 부작용의 문제와 부분적 통제를 통한 효과의 극대화 방안도 동시에 관심의 대상이 되고 있다.

본 연구는 이러한 맥락에서 폭소노미의 기능과 의미를 고찰해 보고 국내 학술정보사이트에 실제 구축된 폭소노미 태그에 대한 실증 분석을 수행해 보고자 한다. 폭소노미 태그 분석을 통해 학술연구자가 부여한 태그의 유형과 특성을 파악해 보면, 특히 사회적으로 기능할 수 없는 개인적 태그의 유형을 집중적으로 살펴본다. 본 연구에서는 구체적으로 다음의 세 가지를 목표로 연구를 진행한다. 첫 번째는 폭소노미의 의미와 기능, 그리고 그 한계에 대하여 고찰해 보는 것이다. 전통적인 텍소노미 시스템을

어떻게 보완할 수 있는지, 그리고 문제점과 그 한계는 무엇인지 고찰해 본다. 두 번째는 폭소노미 구축 사례를 통하여 폭소노미 태그의 유형과 특징을 살펴보고 통제되지 않은 태그의 부작용을 최소화하기 위하여 어떠한 노력이 이루어지고 있는지 살펴본다. 세 번째는 한국의 대표적인 학술정보서비스 RISS(Research Information Service System) 내서재에 부여되고 있는 폭소노미 태그를 분석하여 그 유형과 특징을 살펴보며, 이를 기반으로 학술정보서비스를 위한 효과적인 폭소노미 운영 방안을 제안한다.

## 2. 폭소노미의 의의

### 2.1 폭소노미의 의미와 기능

폭소노미는 디렉토리를 기반으로 하는 전통적 분류 기준인 텍소노미 대신 ‘태그’에 따라 나누는 새로운 분류 체계를 의미한다. 위키피디아에 따르면, ‘자유롭게 선택된 키워드를 사용해 구성원이 함께 정보를 체계화하는 방식을 의미하는 신조어’라고 풀이하고 있다. 폭소노미는 웹페이지, 온라인 사진 등의 컨텐츠를 이용자가 개방적으로 조직하여 부여한 레이블이 모여져서 비 계층적으로 형성된 용어의 그룹이라고 정의할 수 있겠다. 이는 ‘ethnoclassification’, ‘distributed classification’, ‘social classification’으로도 명명되지만, 폭소노미라는 용어가 구글 검색에서 다른 용어들보다 월등하게 많은 결과를 나타내고 있다는 점에서 현재까지 가장 압도적으로 사용되고 있는 용어로 보여 진다(Speller 2007).

많은 연구자가 합의한 폭소노미의 의의는 다음과 같다. 첫 번째, 체계적이지 않지만 복잡하지 않고 학습이 불필요하다. 두 번째, 동의어, 유의어 등의 통제가 불가능하지만, 분류 작업의 경제성을 도모하고 색인 시스템의 관료성이 배제된다. 세 번째, 다양한 관점에서 부여된 태그를 통해 이용자는 우연한 발견의 기회를 만날 수 있다. 여기에 더해, Weinberger(2005)는 다양한 의도를 가진 이용자가 다양한 관점에서 부여한 태그가 저자나 사서가 부여한 주제어보다 정보 검색에 더 유용할 수 있다고 언급하였다.

덧붙여 연구자들은 소셜 시스템으로서 폭소노미의 의의에 대하여 다음과 같이 논평하고 있다. Vander Wal(2005)은 폭소노미는 개인에 의해 조직되지만, 사회적 환경에서 이루어지고 많은 수의 대중에 의해 발전된다고 말하였다. Shirky(2005)는 폭소노미에 의해 다수의 이용자들은 특정 용어에 대한 의미적 컨센서스를 이를 것이며, 더불어 태그 간에 나타난 통계적 해석을 통해 의미는 결국 선호되는 하나의 용어로 정리될 수 있을 것이라고 말하였다. ‘무비’와 ‘시네마’, ‘퀴어’와 ‘호모섹슈얼’과 같이 미세한 차이가 존재하는 용어는 혼자 생각해서는 쉽게 그 차이를 정리할 수 없지만, 다수 대중이 참여한다면 그 의미의 차이는 확연하게 드러나게 된다. 그는 이것이 바로 폭소노미가 가진 힘이라고 정리하였다. 더불어 이재윤(2008)은 집단 지성의 힘은 폭소노미를 어느 정도 통제하는 효과를 나타내고 있음을 입증하였다. 그는 CiteULike를 대상으로 폭소노미 빈도 분포의 실증적 분석을 수행하였는데, 여기에서 그는 폭소노미가 준통제어휘와 유사한 수준의 통제 효과를 보이고 있음을 밝혀냈다. 그리고 이와

같은 통제 효과는 집단 지성의 힘 때문이라고 유추하였다.

소셜 시스템으로서 폭소노미는 다양한 견해의 수렴을 통해 힘을 분산시키는 동시에 대중에 의한 강력한 의사 결정 체계를 이룬다. 연구자들은 이와 같이 폭소노미가 대중의 군집에 의해 제임스 슈로위키가 언급한 집단 지성<sup>1)</sup>을 이끌어 낼 수 있는 도구라는 것에 공감하고 있다.

한편, 객관적 진리는 없으며, 관찰자의 시각에 의해 대상은 달라질 수 있다는 전제를 바탕으로 몇몇 연구자들은 폭소노미 이면에 있는 상대적 특성에 대해서도 논평하고 있다. Peterson(2006)은 폭소노미를 상대성 이론에 입각하여 논평하였다. 그는 모든 태그는 다양한 관점을 가진 개별 이용자에 의해 생성되었기 때문에 전통적 분류 체계가 천명하는 ‘A is B’ 또는 ‘A is not B’라는 논리가 적용되지 않는다고 말한다. 그는 이러한 특성 때문에 한 가지 관점에 의존한 전통적인 분류 체계와 폭소노미는 일치되지 않을 수 밖에 없다고 지적하였다. Weber(2007)는 Shirky(2005)의 말을 인용하면서, 폭소노미 시스템에서의 객체는 부분적으로만 ‘그것’으로 간주될 수 있다고 말한다. 하나의 객체는 보는 사람의 관점에 따라 특정 범주에 분류될 수 있지만, 또 다른 범주에도 분류될 수 있다. 기존의 분류 체계는 이러한 특성을 수용하지 못하였으나, 폭소노미에서는 이를 수용한다고 말하였다.

## 2.2 폭소노미와 텍소노미

목록자들은 일반적으로 아리스토텔레스의

학분 분류에 기반한 전통적 분류 체계에 익숙하다. 전통적 의미의 분류는 개념의 외연을 구분하여 하나의 체계를 조직하는 것으로 최저의 종개념에서부터 최고의 유개념에 이르기 위해 목적에 맞는 집합을 되풀이 하는 귀납적 체계를 가진다. 전통적 분류 체계는 확장성이 없어 신생 주제 삽입이 어렵다는 단점 이외에도 구분지가 배타적 성격을 가진다. 정해진 프레임 워크에 사물의 특정 위치를 부여하고자 하는 특성 때문에 하나 이상의 주제 분야에 머무를 수 있는 다양한 아이템에 배타적 위치가 부여된다. 그러나 물리적 위치가 불필요한 네트워크 상에서 디지털 아이템은 굳이 한 개의 주제 분야로 그 위치를 지정할 필요가 없게 된다. 폭소노미는 실제로 이용자가 도서관의 책을 브라우징하는 것처럼, 온라인상에서 이용자에게 우연한 발견의 기회를 제공한다. 이용자에 의해 생성된 색인은 또 다른 이용자가 접근하기 쉬운 통로를 제공할 수 있으며, 이러한 특징은 서명이나 주제명 표목에 기반한 검색 시스템보다 유용할 수 있다.

한편, Noruzi(2007)는 폭소노미와 텍소노미의 차이점을 다음과 같이 정리하고 있다.

첫 번째, 텍소노미는 상하위 개념이라는 계층적 구조를 가지며, 동위 계열간에 배타적 성격을 띤다. 반면, 폭소노미는 계층적이지도 배타적이지도 않다. 이러한 특성은 양자의 장점이 될 수도 있고 단점이 될 수도 있다고 덧붙이고 있다.

두 번째, 물리적 개체를 대상으로 고안된 전통적 분류 체계는 배타적 위치를 지정해야 한다.

1) 개인의 행동을 모으고 연결시켜 집단적인 가치를 만들어내는 구조.

도서가 복합적 주제를 담고 있어도 더 강조되는 측면으로 분류하여 한 곳에 위치를 지정한다. 그러나 폭소노미를 기반으로 한 분류 체계에는 물리적 제약이 없다. 인위적으로 한 곳에 꾀분류체를 위치시켜야 할 필요가 없다. 연관된 주제 분야는 링크를 통해 연계시킬 수 있다.

세 번째로 텍소노미는 최저의 종개념에서부터 최고의 유개념에 이르기 위해 목적에 맞는 집합을 되풀이하는 개념의 조직 체계이다. 반면, 폭소노미는 개체를 개인이 이해하는 바에 따라 또는 목적에 의해 정의한 개념의 집합체 일 뿐이다(표 1 참조).

### 2.3 폭소노미의 한계와 문제점

폭소노미는 텍소노미와 달리 태그 선택에 제한이 없고 자유로우며, 실시간으로 이용자의 관심사가 반영된다는 장점을 가지지만, 다른 한편으로는 다음과 같이 검색 시스템에 치명적인 문제점을 가지고 있다고 지적받고 있다.

첫 번째, 유의어, 동음이의어에 대한 통제가 부재하다. 이러한 특징은 검색 시스템의 정확율을 떨어뜨린다. ‘Apple’이라는 검색어는 과일일 수도 있으며, 컴퓨터 회사일 수도 있다. 이

러한 동음이의어는 이용자 요구에 적합하지 않은 결과를 나타낼 수 있다(Weinberge 2006).

두 번째, 어형적으로 조악하거나 의미상의 변형이 있는 태그를 통제하지 못한다. 교육받지 않은 대규모 이용자에 의한 색인은 일관성이 결여될 수 밖에 없다. 어형적으로는 단복수 구분과 같은 문제가 생길 수 있으며, 어의적으로는 과일 ‘apple’과, 컴퓨터회사 ‘Apple’과 같이 대소문자의 차이에 따라 의미가 바뀌어 버리는 경우도 발생한다. 더불어, 오타와 조합어, 다국어의 문제도 여기에 한 끝하고 있다.

세 번째는 태깅의 레벨에 관한 문제이다. 동일한 새 사진에 어떤 사람은 ‘새’라고 태깅하는 반면, 어떤 사람은 ‘참새’라고 태깅 할 수도 있다. 두 이용자가 개체를 인식하는 수준이 다른 것이다. 이러한 문제점 때문에 연구자들은 이용자 태그의 장점을 유지하면서도 계층 구조를 도입할 필요가 있다고 주장하기도 한다(Golder and Huberman 2007).

한편, 일각에서는 전술한 몇 가지의 문제점들은 다수의 태그가 군집된다면, 자연스럽게 희석될 수 있다고 바라본다. 따라서 무엇보다 중요한 것은 다수 군중의 참여를 통해 태그의 규모를 증가시키는 것이라고 논평한다. 그러나

〈표 1〉 텍소노미와 폭소노미의 특징

	텍소노미	폭소노미
개념구조	위계구조	비 계층적
배타성	배타적, 인위적 분류	비 배타적
신생주제 처리	신생 주제 삽입이 어려움	신생 주제 삽입 용이
분류 주체 및 특성	사서 또는 운영자가 분류 (특정 체계에 대한 지식 필요) 기계적 인위적	이용자가 부여 (학습이 불필요) 색인 시스템의 관료성 부재
검색 기능	검색의 정확률 제고	검색의 재현율 제고 유의어 동음이의어 등 검색 효율에 문제 발생 가능

운영적 측면에서 볼 때, 폭소노미 사이트들은 무작정 규모의 힘만을 기대하면서 기다리고 있을 수만은 없다. 따라서 그들은 태그 생성 가이드를 제시하고, 관련어 태그를 제안하는 등 오류를 최소화하기 위한 부분적 통제를 시도하고 있다. 더불어 통계 기법을 응용하거나, 부분적으로 시소러스를 도입하자는 주장도 나오고 있다. 시소러스 접목을 주장하는 연구자들은 계층 구조를 도입하고 관련어와 동의어를 제공해 줌으로써 검색의 효율성을 제고할 수 있다고 이야기한다(Spiteri 2007).

### 3. 폭소노미 구축 사례

폭소노미는 일반적으로 소셜 북마크 사이트에서 이용자가 부여하는 태그를 통해 형성되고 있다. 소셜 태깅은 대상 컨텐츠나 이용자에 대한 두지 않고 운영되는 포털식 북마크 사이트를 통하여 가장 많이 형성되고 있다. 이를 소셜 북마크는 인터넷 자원의 효율적 접근과 발견, 집단 지성의 창출에 초점을 맞추고 있으며, 딜리셔스(<http://del.icio.us>), Flickr(<http://www.flickr.com>), Frassle(<http://www.frassle.org>)과 같은 사이트가 대표적이다. 한편, Connotea, CiteUlrike와 같이 학술정보 데이터베이스나 전자저널 등의 컨텐츠를 대상으로 운영되고 있는 소셜 북마크 사이트도 존재한다. 이들은 대상 컨텐츠를 학술연구정보로 제한하고 있어, 대부분의 이용자가 학술연구자이다. 이들 사이트는 학술정보 제공 기관과 협력하여 컨텐츠에서 메타데이터를 자동 추출하거나, endnote와 같은 참고문헌관리 기능을 제공하기도 하며

동일 태그를 생성한 이용자간에 학술 커뮤니케이션이 가능하도록 소셜 네트워킹 기능을 제공하고 있다.

한편, 소셜 태깅을 도입해 시스템의 주제 접근성을 향상시키고 있는 경우도 있다. OPAC에 도입된 소셜 태그 시스템은 도서관 소장 정보원에 대한 서지정보를 자신의 컬렉션으로 축적하여 분류, 관리하기 위한 목적으로 활용되고 있다. 이용자는 컨텐츠를 북마크하면서 동시에 태그를 부여하는데, 태그를 통해 같은 태그를 생성한 타인의 정보를 열람하거나 공유할 수 있다. 더불어 이렇게 생성된 태그는 OPAC 시스템의 검색 효율을 제고하는데 활용될 수 있다. 이는 기존의 분류나 주제명 표목을 대신하여 이용자의 주제 접근성을 향상시킬 수 있다.

이 장에서는 포털형 폭소노미와 OPAC형 폭소노미의 구축 및 운영 사례를 살펴본다. 구축된 폭소노미의 유형과 특징, 그리고 통제되지 않은 태그가 야기할 수 있는 오류를 최소화하기 위해 어떠한 노력이 이루어지고 있는지를 살펴보도록 한다.

#### 3.1 포털형 폭소노미

포털형 폭소노미 사이트에서는 북마크링이나 가젯을 사용하여 이용자가 현재 방문한 웹사이트를 북마크하고 태그를 생성할 수 있도록 지원한다. 대표적인 폭소노미 사이트인 딜리셔스와 한국형 딜리셔스를 표방하고 있는 마가린의 사례를 살펴본다.

##### 3.1.1 딜리셔스(<http://del.icio.us>)

2003년 9월 Joshua Schachter가 시작한 딜

리셔스는 소셜 북마크의 대명사로 자리잡았다. 현재까지 가장 활발한 사용자 층을 확보한 소셜 북마크로 Flickr와 함께 집단지성을 이용한 웹2.0 서비스의 대표 주자로 불리 우고 있다. 딜리셔스는 소셜 네트워크에 기반한 인터넷 북마크 서비스이다. 관심이 있거나 중요하다고 판단되는 인터넷 페이지를 등록해두고 그 페이지에 대한 주석을 달면 그 내용이 인덱싱 처리되어 본인의 계정에 저장된다. 이렇게 저장된 정보는 딜리셔스 사용자 누구에게도 공개되어 공유될 수 있으며, 이 데이터를 바탕으로 가장 북마크가 많이 된 페이지를 보여주기도 하고 특정인이 관심을 갖는 북마크 정보의 내역을 출력해 주기도 한다. 또한 내가 관심을 갖는 태그와 특정 페이지에 동일한 관심을 가지는 타인과도 연계될 수 있다. 많은 사용자들이 인터넷 페이지를 스스로 분류하고 저장한 결과, 그 어려운 로봇이 할 수 없었던 또 하나의 거대한 데이터베이스가 구축되고 있는 셈이다. 딜리셔스 성공의 여파로 영어권에는 수십 종의 소셜 북마크 사이트가 서비스 중이다.

한편, 딜리셔스의 태그는 관련 연구자들에 의해 다각도의 분석 작업이 수행되고 있다. 이러한 연구의 대부분은 폭소노미가 기준의 텍소노미를 대체할 수 있는지, 폭소노미 시스템의 문제점은 무엇인지를 실사하는 연구이다.

Spiteri(2007)는 딜리셔스의 태그를 샘플링하여 태그의 개념 유형을 분석하였다. 딜리셔스에 생성된 태그의 개념 유형으로는 Things가 76%로 가장 많았으나, 그밖에 activities(12%), properties(8%), disciplines(4%)에 관한 것도 존재한다고 밝혔다. 그밖에 단복수, 동음이의어와, 유의어, 오탈자 및 용어의 변형 등의 문제를

지적하면서, 폭소노미는 그 나름의 장점을 충분히 가지고 있으나 열거한 문제점들은 검색 시스템에 큰 오류를 야기할 수도 있다고 지적하였다.

### 3.1.2 마가린(<http://mar.gar.in>)

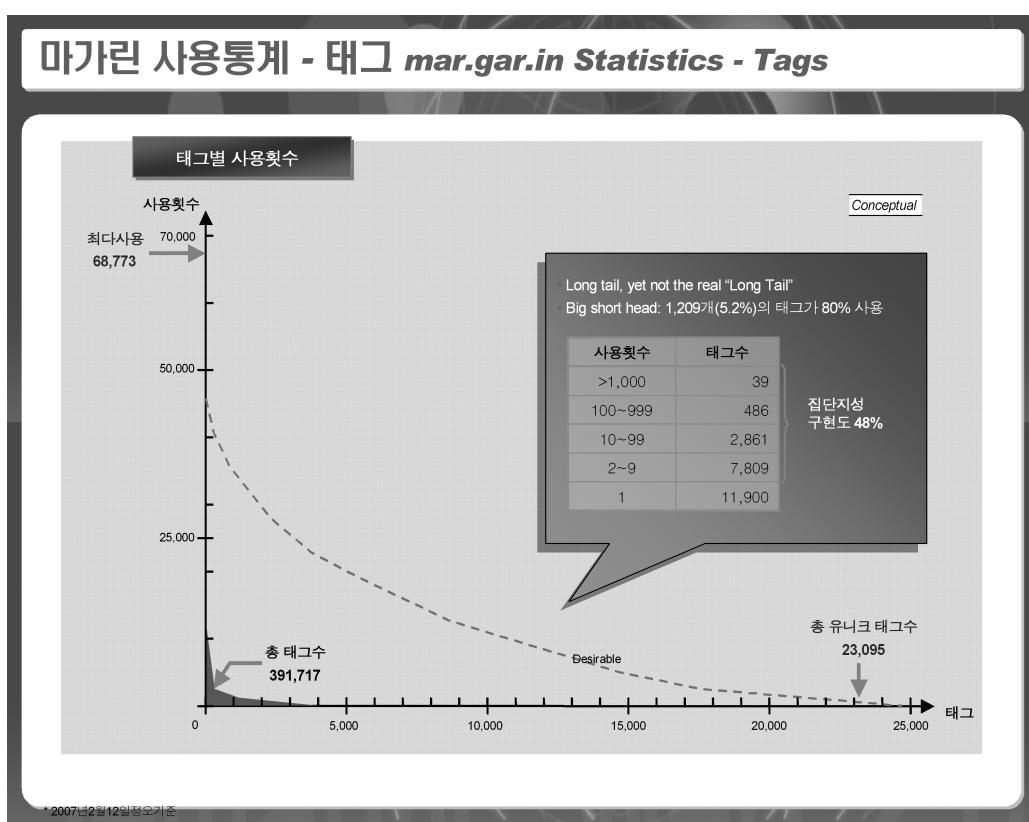
한국에서 소셜 북마크가 그동안 활성화되지 못한 이유는 프레임의 과다 사용 및 태이틀 태그를 잘 사용하지 않는 한국 웹사이트의 제작 관행에 있었다고 한다. 게다가 포털의 지배적 환경 때문에 굳이 북마크가 필요 없었으며, 짧은 웹페이지의 수명과 잘 깨지는 링크 때문이라는 지적도 있다. 이용자 역시 즐겨찾기와 북마크의 개념 구분이 부재하고 태깅에 대한 이해도 부족하다. 마가린은 딜리셔스의 한국판을 표방하고 나선 사이트이다. 딜리셔스의 한국판이라기보다 한국형 딜리셔스를 지향하는 사이트이다. 딜리셔스에는 영어를 모국어로 하는 사용자가 대다수이기 때문에 대상 페이지는 대부분 영문 페이지들이다. 한국인 사용자가 얼마 없기 때문에 수십 명의 한국 사람이 북마크한 웹페이지라도 영문 페이지에 묻혀 탐색하기 어렵다. 마가린에서는 한국 사람끼리 한국의 웹 페이지와 태그를 모을 수 있고, 동시에 의미있는 해외의 웹 페이지와 태그도 모아 한국의 집단지성을 구현할 수 있다(마가린[mar.gar.in](http://mar.gar.in) 2007).

마가린은 앞서 언급한 바 있는 이용자 태그의 근본적 문제점을 보완하기 위하여 태깅을 위한 이용자 가이드를 마련하고 있다. 여기에 가장 강조하고 있는 것은 '소셜' 개념이다. 마가린은 소셜 북마크 서비스이므로 집단 지성의 실현을 위해 같은 의미에 대해서는 가능하면 같은 태그를 붙일 수 있도록 가이드하고 있다.

태깅을 할 때 다른 사람의 태그나 인기 태그를 참고하여 동일 단어를 선택하거나, 연관태그, 인기태그, 추천태그, 권장태그, 내가 썼던 태그를 적절히 이용해 선택하도록 한다. 게다가 태그는 공백을 허용하지 않는데, 이것은 여러가지 다른 표현 방식이 존재할 수 있는 동일 개념의 태그를 최소화하고 실수로 공백이 두 개 들어갈 수 있는 오류를 줄이기 위해서이다. 또한 태그 구름에서의 가독성 저하 문제, 검색시 정확성이 떨어지는 문제 등을 감소하기 위함이라고 한다.

한편 마가린은 단순한 어미변화만으로도 같

은 단어가 여러 개의 다른 개념이 되어 버릴 수 있는 한국어의 고유한 특성을 지적하고 이러한 문제 등에 의해 마가린의 태그 사용은 롱테일을 보이고 있다고 지적한 바 있다. 2007년 통계를 보면, <그림 1>과 같이 5.2%의 태그가 82%의 사용율을 나타내고 있으며, 52%태그가 유니크하게 사용되고 있다고 하였다. 마가린은 유니크한 태그를 최소화하기 위해서 공백비허용, 태그 추천, 인기태그, 태그 뮤음 등의 노력을 하고 있으며, 향후에는 오토 태깅 및 구조화된 태그, 태그 위키 등을 적용할 계획이라고 밝혔다(마가린mar.gar.in 2007).



<그림 1> 마가린의 태그 사용 통계(마가린mar.gar.in 2007)

### 3.2 OPAC형 폭소노미

폭소노미가 구축된 대표적인 OPAC은 SOPAC (<http://www.aadl.org/catalog/>)과 LibraryThing ([www.librarything.com](http://www.librarything.com))이다. SOPAC은 미시간주 어바나 지역 도서관의 도서관 소장 자원만을 대상으로 한 태깅 시스템이다. 따라서 컨텐츠와 사용자 집단이 한정되어 있다. 이러한 특징은 소셜 네트워크를 통한 집단 지성 창출에 제약이 된다. 반면 미국 국회도서관 목록, 아마존의 목록 뿐 아니라, Z39.50 프로토콜을 지원하는 어떠한 OPAC 사이트도 자유롭게 추가하여 태깅할 수 있는 라이브러리링은 컨텐츠의 제한은 물론 이용자의 제한도 없다. 따라서 방대한 태그를 생성할 수 있다. 이러한 특징은 포털형 소셜 북마크 사이트와 같이 집단 지성의 실현에 한발자국 더 가까이 다가갈 수 있다. 여기에서는 라이브러리링의 폭소노미 구축 사례에 대하여 조금 더 자세히 살펴본다.

라이브러리링은 팀 스펠딩이라는 개인에 의해 개발된 웹 2.0 기반 도서정보 공유 시스템이다. 이 사이트에서 이용자는 서지 정보를 태깅하고 참조 컬렉션을 축적할 수 있으며 다양한 소셜 기능을 통해 타인과 교류할 수 있다. 리뷰, 평점과 같은 소셜 기능 이외에도 관심사를 중심으로 이용자들이 자유롭게 대화하거나 그룹을 형성하여 활동할 수 있도록 지원하고 있다. 서지상세정보에는 해당 서지를 북마크한 이용자수, 생성된 리뷰와 평점, 또한 해당 서지와 관련된 대화 정보가 표시되며, 이를 종합한 ‘대중성(Popularity)지수’가 표시된다. 그리고 서지 상세정보를 중심으로 ‘멤버’, ‘대화’, ‘그룹’ 페이지가 유기적으로 연계된다(조재인 2008).

라이브러리링의 큰 특징 중에 하나는 FRBR (Functional Requirements For Bibliographic Records)의 work(저작) 개념 단위로 서지 정보를 묶어서 태그가 생성될 수 있도록 하고 있다는 점이다. 이러한 특징은 유관한 저작에 동일 태그가 자동으로 생성되어, 다수의 군중이 참여하는 것과 동일한 효과를 나타내게 한다. 라이브러리링의 ‘무기여 잘있거라’에 생성된 태그의 숫자가 아마존 사이트의 태그보다 수십배나 많은 것은 이러한 운영 방식 때문이라고 한다. 한편, 라이브러리링에 3,900개의 태그가 생성된 ‘Jared Diamond’s Guns, Germs and Steel’이라는 서지를 보면, “Ancient History”, “anthropology”와 같은 태그가 다수 생성되어 있음을 확인할 수 있다. 그러나 생성된 태그 중에는 “apple”, “rice”와 같은 엉뚱한 태그들도 있어, 이 태그만 보면 마치 이는 요리와 관련된 주제 분야의 서지로 오해할 수 있을 것이다. 그러나 3,900개의 방대한 태그가 생성된 이 서지의 태그 구름을 보면, 여기에서 한두개의 이러한 노이즈는 자연스럽게 희석되고 있음을 알 수 있다.

한편, 라이브러리링에서 실시간으로 제공하고 있는 폭소노미 통계와 관련 블로그에서 발표하고 있는 몇 가지 내용을 살펴보면, 여기에 형성된 태그에는 다음과 같은 특징을 보이고 있음을 알 수 있다(LibraryThing Zeitgeist 2008).

첫 번째, 라이브러리링에 생성된 이용자 태그에는 사회화되지 않고 개인적인 의미가 담긴 태그(signed, living room self)가 다수 포함되어 있으며, 이용자 개인의 상황이나 활동들과 관련된 태그도 다수 존재한다. ‘read’, ‘unread’가 가장 많은 수를 차지하는 태그에 포함될 정도로 사회화되지 못한 다수의 태그가 존재한다.

두 번째, 라이브러리링은 서지 정보에 기반한 태깅 시스템이므로 여기에는 픽션, 시리즈, 문현, 페이퍼백과 같은 서지 용어들이 많은 빈도를 보인다.

세 번째, 'queer', 'gayfiction'과 같이 주제명 표목으로 많이 쓰이지 않는 구어적 일상 표현이 직설적으로 표현되어 있다(표 2 참조).

### 3.3 폭소노미 태그의 어휘상의 문제와 부분적 태그 통제 노력

앞에서 언급한 폭소노미 사이트들은 대상 컨텐츠, 언어, 이용자, 운영 방식에 있어서 상이한 특징을 가지고 있지만, 폭소노미가 가진 어휘상의 문제는 모든 사이트가 공통적으로 안고 있다. 비통제어인 폭소노미에는 동의어, 유의어 등의 문제를 완전히 배제하기 어려우며, 이러한 문제 이외에도 이용자의 어휘 선택에는 'funny', 'let's study'와 같은 사회화되지 못한 단어가 포함되어 있었다. 이러한 어휘 선택의 문제는 우선적으로 폭소노미 시스템의 겸색 성능에 장애를 가져올 수 있는데, 앞에서 살펴본 폭소노미 사례에서 발견된 문제점들을 아래와 같이 다시 몇 가지로 정리해 본다.

첫 번째, 딜리셔스 폭소노미 태그는 주제를 가장 많이 반영하고 있었지만, 주제 이외에도

문서의 유형이나 문서 소유자 또는 정보원을 반영하고 있는 경우가 종종 있었다. 더불어 개인의 느낌이나 특성, 행동과 관련된 사회화되지 못한 태그도 다수 사용되고 있었다.

두 번째, 공통적으로 단복수, 동음이의어, 유의어, 오토자나 용어의 변형이 태그의 일관된 채택을 방해하는 요인이 되고 있으며, 특히 한국어의 경우 어미가 발달되어 있는 특징으로 어미에 따른 의미의 변화가 나타나고 있었다.

세 번째, 라이브러리링은 서지정보를 대상으로 폭소노미 태그가 생성되고 있으므로 서지적 표현을 반영한 태그가 다수 존재하고 있었다. 또한 주제명 표목표에 존재하지 않는 구어적 표현이 다수 존재하였다.

이러한 이용자 태그의 문제를 경감하기 위하여 폭소노미 운영 사이트에서는 태그의 부분적 통제 노력을 병행하고 있었다. 첫 번째, 이용자에게 태그 생성 가이드를 제공하거나, 인기태그, 추천태그, 연관태그 등을 제시하고 있었으며, 태그 구름을 통해 이용자에게 태그에 대한 가독성을 높이고 있었다. 두 번째, 오토 태깅이나 구조화된 태그 또는 태그 위키 등을 도입해 태그 시스템을 좀 더 효과적으로 운영하고자 하는 노력도 진행 중이다(마가린mar.gar.in. 2007). 더불어 FRBR의 work(저작) 단위로 서지를 묶어서 연관된 서지에 동일 태그가 자동 생성되도록

〈표 2〉 Librarything top tag(2008년 9월 19일 기준: Tags added 40,376,000)

태그	사용 회수(회)	태그	사용 회수(회)
fiction	2,139,044	mystery	389,724
fantasy	719,963	read	374,591
history	595,729	nonfiction	342,689
non-fiction	416,552	unread	276,023
science fiction	391,900	poetry	273,143

함으로써, 태그의 규모를 자연적으로 증가시키고, 그에 따라 스팸 태그의 가시성을 희석시키는 운영 노하우는 주목할 만하다.

#### 4. 학술정보서비스의 폭소노미 분석

앞에서는 폭소노미 사이트의 운영 사례와 구축된 태그의 어휘상의 특징에 대하여 살펴보았다. 본 장에서는 한국의 대표적인 학술정보서비스인 RISS에서 구축되고 있는 폭소노미의 특징에 대하여 살펴보고 효율적 운영을 위한 몇 가지 사항을 제안하도록 한다.

##### 4.1 분석 방법의 설계

RISS(<http://www.riss4u.net>)는 한국의 교육인적자원부 산하기관인 한국교육학술정보원에서 운영하고 있는 학술연구정보 포털 서비스이다. RISS에는 학술정보 컨텐츠를 이용자가 태깅하고 북마크할 수 있도록 하는 내 서재라는 개인화 서비스가 있다. 이용자가 컨텐츠를 북마크할 때, RISS가 제공하는 지식맵을 참조하여 학문 분야를 먼저 지정하도록 하는 특징이 있다. 여기에서 이용자는 컨텐츠의 분야를 먼저 지정하고 지정된 분야 내에서 자유롭게 태그를 생성할 수 있으며 태그는 내서재의 책장명으로 등록된다. RISS의 내서재는 “소셜 태그”, “폭소 노미”라는 용어를 사용하

고 있지 않으며, 개인화 기능이 일반 소셜 태그 사이트보다 강조되고 있으나, 이용자 태그를 통하여 검색을 수행할 수 있고, 타인의 관심사를 공유할 수 있다는 측면에서 기능적으로 동일하다.

##### 4.1.1 샘플링 데이터 추출 방법과 기준

태그 분석을 위하여 샘플링 데이터를 추출하는 방법과 기준은 다음과 같다. 2008년 8월까지 RISS 내서재의 문헌정보학 분야, 교육공학 분야, 컴퓨터 교육 분야, 교육심리분야, 교육학 일반분야에 등록되어 있는 전체 태그의 숫자는 <표 3>과 같이 각각 134건, 154건, 29건, 477건, 400건이다. 본 연구에서는 해당 분야에 등록된 전체 태그 총 1,000여건을 추출하여 분석한다.

##### 4.1.2 태그 분석 방법

RISS의 내서재에서 샘플링된 태그는 다음과 같은 방법으로 분석한다.

첫 번째, 사회적 태그로써 기능할 수 있는 태그의 빈도와 그 특징을 살펴본다. 여기에서 사회적 태그라고 함은 컨텐츠를 지칭하고 있는 개념에 대해 컨센서스가 가능한 단어로 이루어진 태그로 정의한다.

- ① 먼저 전체 분석 대상 태그 중에 사회적 태그의 사용 빈도를 파악하고
- ② 그중에서 국립중앙도서관의 주제명 표목 표에 존재하는 태그와 그렇지 않은 태그의 비율을 파악해 본다.

<표 3> 분석 대상 태그

분야	문헌정보	교육공학	컴퓨터교육	교육심리	교육학일반	계
분석대상 태그(건)	134	154	29	477	400	1,194

③ 더불어 국립중앙도서관의 주제명 표목표에 존재하지 않는 태그를 대상으로 그 유형과 특성을 분석해 본다.

두 번째로 개인적 태그의 사용 빈도를 살펴보고 그 유형과 특징을 분석해 본다. 개인적 태그는 태그 부여자의 목적, 의지, 감정, 행동 등이 반영된 태그로, 검색 시스템에 오류를 야기할 수 있으며, 동일 관심사를 가진 타인을 연계시킬 수 없는 비 사회적 태그로 정의한다.

세 번째로 전체 태그에 존재하는 약어, 유의어, 다의어, 복합어 및 오타의 빈도를 파악하고 그 특성을 분석해 본다.

마지막으로 분석의 대상이 된 태그 중, 복수개 존재하는 태그와 유니크한 태그의 빈도를 알아본다. 복수개 존재하는 태그는 집단 지성의 실현 정도를 의미할 수 있는 태그이며, 유니크한 태그는 컨텐츠의 개인적 분류, 관리에만 사용될 수 있는 태그로 분류할 수 있겠다.

#### 4.2 폭소노미 분석 결과

분석 대상이 되는 태그를 가장 먼저 사회적으로 기능할 수 있는 태그와 그렇지 않은 태그로 구분하였다. 사회적으로 의미가 있는 태그 중 ‘~관련’, ‘~관한 학술지’와 같이 기계적으로 처리가 가능한 접미어가 포함된 태그는 사회적 태그의 범주에 포함시켰다. 분석 대상 태그 중에 사회적으로 의미가 있는 태그의 빈도는 전체의 81.7%(978개)이며, 개인적 태그는

전체의 18%(216개)에 이른다(표 4 참조).

##### 4.2.1 사회적 태그에 대한 분석

사회적 태그로 구분된 978건을 대상으로 국립중앙도서관의 주제명표목표와 대비하는 과정을 거쳤다. 978개의 태그 중에, 주제명표목표에 존재하는 경우는 526건으로 53.8%에 불과하였다. 주제명표목표에 존재하지 않는 452건(26.2%) 태그의 대부분은 ‘문서클러스터링’, ‘의학사서’와 같이 복합어로 구성된 전문 용어이거나 ‘GBS’와 같은 약어로된 전문 용어, ‘문정(문현정보)’, ‘교공(교육공학)’과 같이 해당 학문분야에서는 구어적으로 사용되고 있는 약어가 포함되어 있었다. ‘전자기록 학술지’, ‘교육공학 관련’과 같이 특정 주제를 의미하는 명사나 복합 명사 뒤에 서지적 표현을 추가하거나, ‘~관련’, ‘~대하여’ 같은 첨가어가 따라 오는 경우도 상당수 존재하고 있었다(표 5 참조).

##### 4.2.2 개인적 태그에 대한 분석

한편, 사회적으로 의미가 없다고 분석한 태그는 216건으로 전체의 18%였는데, 그중 58% 이상이 ‘과제중심의 컴퓨터 교육’, ‘논문 준비’, ‘구하지못한해외자료’, ‘공부하자’, ‘XX꺼’ 등 개인의 연구 주제, 개인의 목적 및 활동, 개인의 감정과 의지, 개인 이름이나 소유와 같은 종류이다. 그 밖에 ‘학위논문’, ‘해외논문’과 같은 서지적 표현의 태그와 의미를 알 수 없는 단어로 이루어진 태그도 존재하였다(표 6 참조).

〈표 4〉 사회적 태그와 개인적 태그의 빈도

구 분	사회적태그	개인적태그	계
태그(건/%)	978(81.7%)	216(18%)	1,194(100%)

〈표 5〉 사회적 태그의 분석 결과

사회적 태그					
구 분	주제명표목에 존재	주제명표목에 비존재			계
태그(건/%)	526 (53.8%)	452 (46.2%)			452 (100%)
	예)영재, 대안학교	복합어로 구성된 전문 용어	예)의학사서, 다문화의해교육	312 (59.3%)	
		구어적으로 사용되는 약어	예)문정 교공	14 (3.1%)	
		고유명사	예)퍼시픽예일대	33 (7.3%)	
	주제어+부가어	서지적 표현	예)블런디드러닝학술지	46 (8.7%)	
		접미어	예)실업계관련	47 (8.9%)	

〈표 6〉 개인적 태그의 분석 결과

개인적 태그					
구 분	개인의 주제, 목적, 활동, 감정, 의지, 소유			서지적태그	의미 없는 단어, 오타
태그(건/%)	127 (58.8%)			64 (29.6%)	25 (11.6%)
	개인적 주제	예)과제 중심의 컴퓨터 교육	8 (6.3%)	127 (100%)	216 (100%)
	개인의 목적, 활동	예)강의 준비, 구하지 못 한 해외 자료	87 (68.5%)		
	개인의 감정, 의지, 소유	예)공부하자, XX꺼	32 (25.2%)		

#### 4.2.3 약어, 유의어, 다의어, 오타 태그에 대한 분석

분석대상 태그 중에 '문정', 'MBTI'와 같은 약어는 14개가 존재하였으며, '이러닝 - e- 러닝', '미술치료 - 예술치료'와 같은 유의어는 26회 사용되었다. 반면 'Apple - apple'과 같은 다의어는 거의 부재하였는데, 그 이유는 학술연구정보 컨텐츠를 대상으로 하고 있어, 복합 주제로 이루어진 학문적 용어가 태그의 다수를 차지하고 있기 때문이라고 본다. 한편 오타는

전체의 2%(25회) 정도로 그렇게 많은 비중을 차지하고 있지 않았다.

#### 4.2.4 복수개 생성 태그 및 유니크한 태그에 대한 분석

학술정보 태깅이 집단 지성과 소셜 네트워크를 형성할 수 있는지 파악해 보기 위하여 분석 대상 태그를 대상으로 복수로 존재하는 태그와 유니크한 태그를 분석해 보았다. 그 결과, 분석 대상 태그 중 총 45개의 태그만이 중복 사용되

었으며, 중복 사용된 횟수를 고려하면 이는 전체의 12%에 불과한 것으로 나타났다. 대부분을 구성하는 88%는 유니크한 태그인 것으로 분석되었다.

#### 4.3 학술연구정보서비스 폭소노미의 특징

위의 분석 결과를 바탕으로 한국의 대표적인 학술정보포털 사이트인 RISS에 생성된 폭소노미 특징을 몇 가지로 정리해 보겠다.

첫 번째, 학술연구정보 폭소노미에는 사회적 태그의 사용이 80% 이상인 것으로 나타났다. 이러한 통계는 폭소노미가 시스템의 검색 효용, 다시 말해 주제 접근 기능 향상에 도움이 될 수 있음을 의미한다. 더욱이 주제명 표목표에 존재하지 않지만 학술적으로 의미있는 주제 개념이 다수 존재하고 있었다. 특히 '의학사서'와 같은 복합어나 '문정', '교공'과 같이 구어적으로 쓰이는 약어가 다수의 이용자에 의해 사용되고 있었다. 이러한 특징은 학술정보 폭소노미가 주제명 표목표를 보완해 컨텐츠에 대한 접근성을 높일 수 있고, 동일 관심사를 가진 연구자를 연계하는 고리로 기능할 수 있음을 시사한다. 대부분의 국내 도서관 메타데이터에 주제 접근을 위한 표목의 기술이 이루어지지 않고 있다는 점을 고려하면, 이용자의 태그 생성 자체는 의미 있는 도움이 될 것이라 보여진다.

두 번째, '블런디드러닝 관련 학술지'와 같이 특정 주제 개념에 서지적 표현을 비롯한 접미어가 부가된 태그가 다수 사용되고 있었다. 이러한 특징은 교육/연구 목적으로 정보를 수집하는 이용자 집단의 특성 때문으로 보여진다. '~관련', '~관련 논문', '~학술지'로 접미어가

특정 패턴을 보이고 있으므로 기계적으로 처리하여 활용할 수 있을 것이다.

세 번째, 18%정도의 태그가 "강의 준비", "공부하자"와 같이 사회적으로 써는 기능하기 어려운 개인적인 태그이거나 스팸 태그인 것으로 나타났다. 개인적 태그는 개인의 감정, 목적, 행위, 소유 등을 의미하고 있는 태그이다. 이용자 가이드를 통해 소셜 태그의 기능을 설명하고 바람직한 태그 선정 방법을 주지시킬 필요가 있을 것으로 보여진다.

네 번째, 포털형 소셜 태그 시스템과 달리 여기에는 다의어가 희소한 것으로 나타났다. 이는 학문적 주제 개념을 나타내고 있는 복합어가 널리 활용되고 있기 때문인 것으로 보여진다.

다섯 번째, 중복 태그의 생성 비율이 전체에 12%에 불과한 것으로 나타났다. 이는 집단 지성과 소셜 네트워크의 실현도를 그리 낙관적으로 볼 수만은 없는 척도이다. 집단의 크기가 결정적인 영향을 미치므로 좀 더 많은 이용자가 집중된다면 다른 양상을 보일 수 있겠지만, 일반 대중이 아닌, 한정된 이용자 집단이 이용하고 있는 학술정보서비스의 특수성을 주지할 필요가 있겠다. 더불어 이용자의 관심사가 다양한 세부 연구 주제 분야로 분산되어 있다는 점을 고려한다면 학술연구정보사이트에서의 집단 지성의 실현은 방대한 규모의 군중이 집중해야만 실현 가능하지 않을까 싶다.

#### 5. 학술연구정보서비스의 폭소노미 운영을 위한 제언

소셜 태그는 이용자가 정보를 탐색하고 탐색

한 정보에 대한 분류와 검색을 용이하게 하기 위한 개인적인 행위이다. 그러나 하나의 컨텐츠에 대해 다수의 사용자가 다양한 관점의 태그를 부여하거나, 하나의 태그에 다양한 컨텐트가 북마크되면서 컨텐츠는 소셜 오브젝트로 진화된다. 개인의 다양한 관점으로 부여된 태그를 끌어 모아 체계화하면 전체로써의 큰 가치 즉 집단 지성이 형성되며, 특정 태그라는 공통된 세マン틱을 공유한 사람들 간에 소셜 네트워크를 위한 연결 고리가 형성된다(조재인 2008). 더불어, 이는 주제 접근이 용이하지 않은 학술연구 정보의 효과적인 접근 도구로 활용될 수 있을 것이다.

그러나 위에서 분석한 바와 같이 태그에는 유의어, 다의어 등의 문제가 내재되어 있을 수 밖에 없으며, 태깅 행위 자체가 개인적인 목적으로 이루어지므로 이용자의 어휘 선택에 사회화되지 못한 단어가 다수 포함되어 있다. 특히 학술연구정보 폭소노미에는 서지적 표현이 반영된 태그가 다수 존재하였고 복합 명사와 구어적으로 쓰이는 약어도 존재하였다. 이러한 태그는 동일 관심사를 가지는 동료 연구자를 효과적으로 연계시키는데 장애가 되며, 검색 시스템에도 치명적인 오류로 작용한다. 학술정보서비스의 폭소노미를 효과적으로 운영하기 위한 방안으로 다음의 몇 가지를 제언한다.

첫 번째, 태그 생성 가이드를 마련하라는 것이다. 이용자에게 사회적 태그의 의미를 충분히 이해시키고, 사회적 태그가 기능할 수 있는 여러 가지 장점에 대해 충분히 설명하며, 태그를 생성하는 원칙과 방법을 안내해야 한다.

두 번째, 추천태그, 연관태그를 통해 이용자 가 태그 선택시 참조할 수 있도록 하며, 하나의

컨텐츠를 기준으로 가장 많이 쓰여지고 있는 태그를 통계적으로 추출하여 가시화할 필요가 있다.

세 번째, 부분적으로 시소러스를 도입하는 방법도 고려해 볼 필요가 있다. 통제 어휘 시스템의 장점과 이용자 태그 시스템의 장점이 시너지를 발휘할 수 있도록, 부분적으로 시소러스를 이용하는 것도 고려할 필요가 있다. 지리명, 시대명은 일차적으로 적용 가능할 것으로 보여진다. 더불어 오토태깅, 스트럭쳐드 태깅을 적용하여 태깅 시스템을 진화시킬 필요가 있다. 오토태깅이란 사용자가 직접 태그를 입력하지 않아도 자체 프로그램이 글의 내용 등을 분석해 자동으로 태그를 추출하는 기술이다. 한편, 스트럭쳐드 태깅은 특정 퍼셋에 따라 구조화해서 정의하는 태깅시스템이다.

네 번째, 학술정보서비스에 구축된 태그의 특징을 기반으로 패턴을 추출하여, 기계적인 처리를 병행할 필요도 있을 것이다. 언급한 바와 같이 학술연구정보 폭소노미에는 ‘~관련’, ‘~논문’ 등과 같이 일정 접미어가 덧붙여진 태그가 다수 존재하였다. 몇 가지 패턴을 찾아서 기계적 처리를 선행한다면 성능을 더욱 제고시킬 수 있을 것이다.

마지막으로 주지해야 할 것은 다수의 군중이 참여할 수 있는 곳에서 집단적 가치가 창출될 수 있다는 점이다. 도서관 및 정보서비스 기관은 1차적인 서비스 대상을 자체 소장 정보원으로 하고 있다. 제한된 컨텐츠를 대상으로는 방대한 이용자의 참여를 끌어들이지 못하며, 따라서 여기에서는 비통제어휘가 가진 오류를 회복시킬 수 있는 동력을 제공받기 어렵다.

## 6. 결 론

학술연구자는 도서관에 소장된 기존의 오프라인 자원보다 앞으로는 더욱 더 인터넷 컨텐츠를 선호하게 될 것이다. 연구도서관에서는 많은 비용을 소요하는 오프라인 자원 목록 작업의 경제성을 고민해야 할 것이다. 더불어 이용자 개인의 자발적인 행위로 부여된 태그를 끌어 모아 체계화하면 컨텐츠에 자동적으로 질서가 부여되며, 전체로써의 큰 가치가 형성될 수 있음을 주지할 필요가 있을 것이다.

본 연구에서는 한국의 대표적인 학술연구정보 포털 서비스인 RISS의 내서재에 생성된 폭소노미 태그 1,000여개를 샘플링하여 분석해 보았다. 분석 대상 태그 중 81.7%가 사회적 태그였으며, 18%는 개인적 태그로 나타났다. 사회적 태그 중에 46.2%가 ‘다문화이해교육’, ‘문정’과 같이 복합주제이거나 구어적으로 활용되고 있는 약어 등 주제명 표목표에는 존재하지 않는 태그였으며, 비사회적 태그는 ‘과제중심의 컴퓨터교육’, ‘강의준비’와 같은 개인적 주제와 목적, 활동, 감정 등을 반영하고 있는 태그로 나타났다. 한편, 여기에서 중복되어 사용된 태그의 비율은 전체의 12%에 불과하여, 이로써는

아직까지 집단 지성의 실현을 언급하기 어려운 정도로 보여진다.

폭소노미의 규모가 방대해지면 자연적인 개인적 태그의 희석 효과가 나타난다. 그러나 폭소노미 운영 사이트 입장에서는 ‘소셜’로써 제대로 기능할 수 있는 사회화된 태그를 부여하도록 이용자를 안내할 필요가 있을 것이다. 더불어 개인 태그, 스팸 태그를 최소화하기 위한 다양한 방법의 도입도 같이 고민해야 할 것이다. 그렇지만 앞서 말한 바와 같이, 무엇보다 중요한 것은 다수의 군중을 참여시키는 것이다. 더구나 세부 전문 분야에 관심사를 가지는 한정된 연구자들이 참여하는 폭소노미 사이트를 대중적 관심사가 반영된 포털형 폭소노미 사이트와 동일하게 생각할 수는 없을 것이다. 이 점을 고려한다면, 단위 도서관에서 운영하는 학술정보서비스에서는 쉽게 집단지성의 효과를 기대하기 어려울 것이다. 따라서 학술정보 포털이나 도서관 연합 사이트에 폭소노미 서비스를 도입하고 다수의 연구자들이 참여할 수 있도록 홍보할 필요가 있을 것이며, 더불어, 여기에 생성된 폭소노미 태그를 단위 도서관이 활용할 수 있는 방법을 고려하는 것이 효과적일 것이다.

## 참 고 문 헌

이재윤, 정도현. 2008. 폭소노미 태그 사용 패턴  
분석: 통제어휘 및 비통제어휘와의 비교  
『제15회 한국정보관리학회 학술대회 논  
문집』, 21-26.

이정미, 2007. 폭소노미의 개념적 접근과 웹 정보  
서비스에의 적용. 『한국비블리아학회  
지』, 18(2): 141-159.  
조재인. 2008. 도서관 정보 수요자를 위한 소셜

- 네트워크 서비스 도입에 관한 연구. 『한국 도서관정보관리학회지』, 39(2): 169-186.
- 마가린mar.gar.in. 2007. 소셜북마킹, 소셜태깅 그리고 마가린mar.gar.in. [cited 2008.9.9]. <[http://www.fortytwo.co.kr/tt/attach/ppt/SemanticWeb2.0\\_PRAK\\_SocialTagging\\_Final.ppt](http://www.fortytwo.co.kr/tt/attach/ppt/SemanticWeb2.0_PRAK_SocialTagging_Final.ppt)>.
- Arch, Xan 2007. "Creating the academic library folksonomy: Put social tagging to work at your institution." *C&RL News*, 68(2). [cited 2008.4.10]. <<http://www.ala.org/ala/acrl/acrlpubs/crlnews/backissues2007/february07/libraryfolksonomy.cfm>>.
- Calhoun, Karen. 2006. The Changing Nature of the Catalog and its Integration with Other Discovery Tools: Final Report. [cited 2008.10.27]. <<http://www.loc.gov/catdir/calhoun-report-final.pdf>>.
- Chan, Sebastian. 2007 "Tagging and Searching - Serendipity and museum collection databases." *Museum & web 2007: International Conference for Culture & Heritage Online*. April, 2007 San Francisco. [cited 2007.12.10]. <<http://www.archimuse.com/mw2007/papers/chan/chan.html>>.
- Golder, S. & Huberman, B. 2006. "Usage patterns of collaborative tagging systems." *Journal of Information Science*, no.32, vol.2 : 198-208. Quoted in Edith. Speller. Collaborative tagging, folksonomies, distributed classification or ethnoclassification: a literature review: *Library Student Journal* February, 2007. [cited 2008.3.9]. <<http://www.librarystudentjournal.org/index.php/lSJ/article/view/45/58>>.
- Marcum, Deanna B. 2005. "The Future of Cataloging." *Ebsco Leadership Seminar*, 2005 Boston. [cited 2008.10.10]. <<http://www.loc.gov/library/reports/CatalogingSpeech.pdf>>.
- Library Thing Zeitgeist. 2006. [cited 2008.9.10]. <<http://www.librarything.com/users.php>>.
- Noruzi, Alireza. 2007. "Folksonomies: Why do we need controlled vocabulary?" *Weboology*, 4(2). [cited 2008.3.9]. <<http://www.webology.ir/2007/v4n2/editorial12.html>>.
- OCLC. 2005. Perception of Libraries and Information Resources [cited 2008.10.4]. <<http://www.oclc.org/reports/2005perceptions.htm>>.
- Peterson, Elaine. 2006. "Beneath the Metadata Some Philosophical Problems with Folksonomy." *D-Lib Magazine*, 12(11), [cited 2007.11.3]. <<http://www.dlib.org/dlib/november06/peterson/11peterson.html>>.
- Shirky, C. 2005. Ontology is Overrated: Categories, Links, and Tags. [cited 2008.3.15].

- 〈[http://www.shirky.com/writings/on\\_tology\\_overrated.html

SOPAC AADL, \[cited 2008.3.10\].  
〈<http://www.blyberg.net/2007/01/21/aadlorg-goes-social/>〉.

Speller, Edith. 2007. "Collaborative tagging, folksonomies, distributed classification or ethnoclassification: a literature review." \*Library Student Journal\*, February. \[cited 2008.3.9\].  
〈<http://www.librarystudentjournal.org/index.php/lsj/article/view/45/58>〉.

Spiteri, Louise F. 2007. "Structure and form of folksonomy tags: The road to the public library catalogue." \*Webology\* 4\(2\). \[cited 2007.11.10\].  
〈<http://www.webology.ir/2007/v4n2/a41.html>〉.

Thingology\(LibraryThing's ideas blog\), 2007. Tagmash: Book tagging grows up, \[cited 2008.4.10\].  
〈<http://www.librarything.com/thingology/2007/07/tagmash-book-tagging-grows-up.php>〉.

Vander Wal, T. 2005. Folksonomy definition and Wikipedia. \[cited 2008.4.10\].  
〈<http://www.vanderwal.net/random/entrysel.php?blog=1750>〉.

Weber, Jonathan. 2007. "Folksonomy and Controlled Vocabulary in Library Thing." \*LIS2452: Indexing and Abstracting Final project\*, \[cited 2007.11.5\].  
〈<http://jonathanweber.info/samples/2452-Folksonomy.pdf>〉.

Weinberger, D. 2005. Tagging and Why It Matters. \[cited 2008.9.10\].  
〈<http://cyber.law.harvard.edu/sites/cyber.law.harvard.edu/files/07-WhyTaggingMatters.pdf>〉.

Weinberger, D. 2006. Taxonomies and Tags: From Trees to Piles of Leaves.  
〈\[http://www.hyperorg.com/blogger/misc/taxonomies\\\_and\\\_tags.html\]\(http://www.hyperorg.com/blogger/misc/taxonomies\_and\_tags.html\)〉.

Quoted in Edith, Speller. Collaborative tagging, folksonomies, distributed classification or ethnoclassification: a literature review: \*Library Student Journal\*, February. 2007. \[cited 2008.3.9\].  
〈<http://www.librarystudentjournal.org/index.php/lsj/article/view/45/58>〉.](http://www.shirky.com/writings/on_tology_overrated.html)