

협력적 북마킹의 태깅 행태 분석

Analysis of the usage Pattern of Tagging in Collaborative Bookmarking

최준연*, 김용수**

세종대학교 디지털콘텐츠학과*, 성결대학교 e-비즈니스IT학부**

Joon-Yeon Choeh(zoon@sejong.ac.kr)*, Yong-Soo Kim(yskim@sungkyul.ac.kr)**

요약

웹문서의 특성을 표현하기 위해 키워드 형태로 부여하는 태깅의 이용이 다양한 웹서비스에서 확산되고 있다. 온라인 즐겨찾기 서비스라고 말할 수 있는 협력적 북마킹 서비스에서도 태깅이 중요한 역할을 수행하는데, 사용자가 부여한 태그는 자신의 북마크를 손쉽게 검색하고 타사용자의 웹문서 검색을 정교하게 만들어준다. 본 연구에서는 사용자들의 태깅 데이터를 분석하여, 웹문서와 사용자의 태그 수에 영향을 주는 요인이 무엇인가를 탐색하였다. 웹문서의 태그 수에는 웹문서의 특성보다 사용자의 특성에 따라 더 큰 편차를 보였으며, 이것은 사용자의 성향이 태그의 다양성에 더 큰 영향을 미친다는 것을 의미한다. 또한 추종적 사용자보다는 가장 먼저 웹문서를 북마킹하는 선도적 사용자들이 더 많은 태그를 생성하는데 기여하는 것으로 나타났다. 풍부하고 다양한 태깅을 통해 서비스의 품질을 향상시키기 위해서는 선도적 사용자들에 대한 보상과 인센티브를 통해 더 많은 지식을 생성할 수 있도록 해야 한다는 시사점을 제공한다.

■ **중심어** : | 태깅 | 협력적북마크 | 집단지성 | 웹2.0 |

Abstract

The use of tagging to describes web documents in the form of keyword has experienced rising popularity among various web services. Tagging also plays an important role in collaborative bookmarking services which can be regarded as an online favorite bookmark service. Tags which are created by users make it easier to search other users' bookmarks as well as user's own bookmarks. In this paper we analyze usage patterns of collaborative tagging for exploring factors influencing the number of tags in web documents and users. We discovered that user's characteristics have more effect on the tags than the web documents' characteristics. Moreover, leading users contribute to make a variety of tag than following users. Our study implies that more knowledge can be created through the incentives for leading user in order to improve the service quality of tagging service.

■ **keyword** : | Tagging | Collaborative Bookmark | Collective Intelligence | Web 2.0 |

I. 서론

최근의 Web 2.0 서비스 환경은 사용자들의 적극적인

참여와 공유를 유도하는 형태를 지향하고 있으며, 이로 인해 UCC, 블로그, SNS(Social Network Service)에서 태깅(tagging) 기능이 다양하게 활용되고 있다[1]. 태깅

접수번호 : #090406-006

접수일자 : 2009년 04월 06일

심사완료일 : 2009년 06월 24일

교신저자 : 최준연 e-mail : zoon@sejong.ac.kr

은 웹문서, 사진, 동영상 등의 콘텐츠에 붙이는 꼬리표로서 자신이 찍은 사진이나, 블로그에 올린 글, 스크랩한 뉴스, 북마크한 웹사이트에 연관된 키워드를 붙여주는 행위이다. 사용자가 콘텐츠에 부여하는 태그는 콘텐츠에 대한 일종의 부연설명인 메타데이터 역할을 수행하는데, 이는 향후에 콘텐츠를 검색하거나 분류할 때 이용된다. 메타데이터는 전통적으로 정보전문가 또는 관리자만이 생산하고 사용자는 정보를 검색할 때 이용하기만 해왔다. 그러나 인터넷이 대중화 되고 인터넷상의 정보가 폭증하면서 정보 전문가의 역할은 한계에 다다르게 되었다[2]. 정보전문가는 엄격하게 정해진 규칙에 따라 정해진 범주 안에서 메타데이터를 부여하여 다양성과 자유도가 떨어졌다면, 일반 사용자들이 부여한 태그는 종류와 개수에 제약이 없어 다양한 정보를 만들어내고 있다. 최근의 태그의 확산은 사용자가 정보의 소비자인 동시에 메타데이터를 생산하는 역할까지 수행한다는 것을 의미한다. 일반 사용자들이 생산한 태그는 콘텐츠를 검색하고 연관 콘텐츠를 분류하는데 큰 역할을 수행하고 있다. 특히 메타데이터가 거의 전무하였던 사진이나 동영상 등의 멀티미디어 콘텐츠에서는 사용자가 부여한 태그가 멀티미디어 검색에서 절대적인 기여를 하고 있다[3].

태깅을 활용하는 서비스에는 여러 가지가 있으나, 그중에서 북마크 공유 서비스는 협업태깅 (collaborative tagging 또는 social tagging)을 활용한 서비스이다. 기존의 북마크는 자신의 브라우저에 저장하여 두었다가 다시 방문하고자 할 때 이를 찾아서 이용하게 된다. 이와 같은 북마크 방식은 다른 컴퓨터를 사용하게 되면 자신의 즐겨찾기를 이용할 수 없게 된다는 단점을 갖고 있다. 그러나 협업태깅을 지원하는 북마크 공유서비스는 자신이 방문했던 웹페이지를 PC와 같은 클라이언트에 저장하지 않고, 북마크 공유 서버에 저장하여, 어떤 컴퓨터를 이용하더라도 인터넷만 연결되어 있으면 언제든지 이를 조회하여 이용할 수 있게 된다. 또한 자신의 북마크뿐만 아니라 타 사용자의 북마크도 검색하고 공유할 수 있기 때문에, 타사용자들이 자주 이용하는 웹페이지가 무엇인지도 알 수 있게 된다. 이는 '집단지성'(collective intelligence)의 대표적인 서비스 형태로

서, 여러 사용자들의 경험과 지식을 축적함으로써 신뢰도 높은 웹페이지가 무엇인지를 알 수 있게 된다. 사용자는 자신에게 의미 있는 웹사이트 또는 웹문서를 북마크하기 위해 태깅하지만, 자신뿐 아니라 다른 사용자들에게도 이정보가 공유된다는 특징을 갖게 되는 것이다. 이러한 구조를 통해 다른 사람들이 즐겨 찾는 웹사이트를 알 수 있고, 동일한 웹사이트를 여러 사용자들이 태깅함으로써 보다 풍부한 메타데이터가 쌓이게 되고, 이로 인해 정교한 검색과 분류가 가능해진다. 일반 사용자가 이러한 지식 생산에 참여함으로써, 과거에 몇몇 일부 관리자의 지식에 한정되었던 단점을 극복하고, 여러 사용자들의 지식을 한곳에 모아 공유함으로써 훨씬 풍부한 메타정보를 획득할 수 있게 된 점이 협력적 북마크 서비스의 가장 큰 가치라고 할 수 있다[4].

사용자가 생성한 메타데이터들이 쌓이게 되면, 이들을 이용하여 분류체계를 만들 수 있는데 이를 폭소노미 (folksonomy)라고 한다[5]. 이는 기존의 전통적인 분류 방법인 택소노미 (taxonomy)와 상반되는 개념으로, 택소노미는 전문가가 사전에 정해진 카테고리에 따라 분류하는 것으로, 도서관의 서적분류 체계와 야후의 디렉토리 서비스[6] 등이 여기에 해당한다. 카테고리 시스템의 단점은 소수의 관리자만이 분류 권한을 갖기 때문에 시시각각 변하는 사회현상을 반영하기 힘들다는 점이다. 반면에 폭소노미는 사용자들이 부여하는 태그를 통해 생성되므로, 현재의 사회현상이 바로 반영될 수 있으며, 수많은 사람들이 생성한 태그에 기반하므로 택소노미보다 풍부한 분류체계를 구축할 수 있다. 택소노미는 통제된 언어만을 사용하는데 반해 폭소노미는 모든 어휘를 사용할 수 있으며 새로운 발견의 기회를 갖는 장점을 갖는다. 협력적 태깅 서비스에서 점점 더 많은 태그가 수집되어 누적될수록, 분류체계로서의 형태를 갖추게 된다[7].

협력적 북마크 서비스에 대한 많은 연구가 이루어져 왔는데, 사용자의 태깅 행태 특성을 찾으려는 실험적 연구도 점점 증가하고 있다. [8]에서는 사용자의 지식 수준이 태깅 행태에 영향을 미치는 점을 분석하였고, [9]에서는 기업에서 내부북마크와 외부북마크에 따라, 그리고 공적인 북마크와 사적인 북마크에 따라 태그의

수가 차이가 나는지 분석하였다. [10]에서는 태깅 행태를 시각화하기 위해 태그를 네트워크로 표현하고, 태그간의 관계를 분석하였다. 태그의 동시사용률에 대한 연구도 많이 이루어졌는데, [11]에서는 문서를 분류하는데 태그가 어떻게 기여하는지 알아보고자, 태그의 사용빈도와 동시사용률에 대한 정량적 분석을 하였고, [12]에서는 시간이 경과함에 따라 태그의 사용빈도가 안정화되는지에 대한 실험을 하였다. 태그의 개수에 대한 연구로는 [13]과 [14]에서 본격적으로 시도되었다. [13]에서는 협력적 북마킹서비스의 평가지표를 만들기 위해 태그의 성장률, 재사용률, 사용빈도를 분석하였고, [14]에서는 delicious의 데이터셋을 이용해 사용자의 북마크의 개수와 태그의 개수간의 관계와 시간이 경과함에 따라 사용자의 태그의 개수가 증가하는 사례를 발견하여 분석하였으나, 대규모 사용자에 대한 통계적 검증은 이루어지지 않았다.

본 연구에서는 폭소노미의 완성도에 직접적인 영향을 미치는 태그의 영향요인을 분석하고자 한다. 태그는 전문가가 아닌 일반 사용자들이 생성하는 산출물인 만큼, 사용자의 행태에 따라 태그의 품질과 양이 크게 달라질 수 있다. 기존 연구에서 다루어지지 않았던 사용자와 웹문서의 태그의 수에 영향을 미치는 변수를 검증하였고, 각각의 변수 중 어떤 변수가 더 큰 영향력을 갖는지 비교하였다. 본 논문의 구성은 다음과 같다. 2장에서는 협력적 북마킹의 구조에 대해서 설명하고 3장에서는 사용자들의 서비스 이용 행태를 분석하여 웹문서와 사용자의 태그의 개수에 영향을 미치는 요인을 알아보고, 4장에서 결론을 맺는다.

II. 협력적 북마킹의 구조

협력적 북마킹 서비스는 사용자(User), 웹문서(URL), 태그(tag)의 3가지 개체로 구성되어 있고, 기본적인 동작 원리는 사용자가 웹문서를 저장하면서 태그를 부여하는 과정을 거친다. 다음의 [그림 1]은 위의 3가지 개체간의 관계를 표현한 그림이다.

사용자는 어떤 웹문서라도 북마킹할 수 있고, 웹문서

는 어떤 사용자에게 의해서도 북마킹 될 수 있다. 사용자는 웹문서를 북마킹하면서 키워드 형태의 태그를 개수에 제한 없이 부여할 수 있다. 다른 사용자가 사용한 태그라도 누구나 사용할 수 있으며, 자신이 사용한 태그도 재사용 가능하다. 자신이 부여한 태그를 수정하거나 삭제할 수 있지만, 타사용자가 부여한 태그는 수정하거나 삭제할 수 없다. 타 사용자가 북마킹한 웹문서와 태그를 누구에게나 공개하기 때문에, 좋은 웹문서는 더 많은 사용자들에게 북마킹되고 다양한 태그가 누적된다.

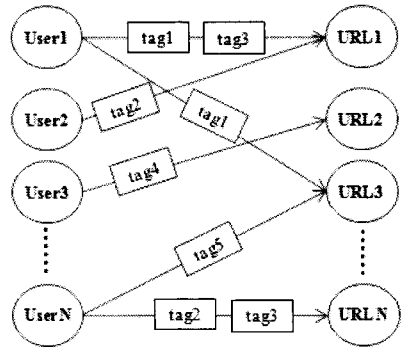


그림 1. 협력적 북마킹의 구조

사용자에게 서비스 이용 상 제약이 거의 없다는 점은 장점이 되기도 하지만, 근본적인 한계점을 갖고 있기도 하다[15]. 예를 들면 동의어나 유의어에 대한 관리가 불가능하기 때문에, 핸드폰, 휴대폰과 같은 태그가 의미상으로는 동일하지만, 태깅에서는 모두 다른 태그로 간주된다. 일부 사용자들은 자신의 웹사이트를 홍보하기 위해 비이상적으로 많은 북마킹과 태깅을 통해 서비스의 질을 저하시키는 행동을 하기도 한다. 또한 사용자가 부여하는 태그간의 의미적 상하종속관계가 무시되고, 모두 수평적인 관계로 설정되기 때문에 정교한 검색을 수행하는 데에는 한계점을 갖고 있다. 또한 태그입력에 있어서 아무런 단어의 제약이 없기 때문에, 맞춤법이 틀린 단어나 비속어 등이 여과 없이 등록될 수 있는 단점이 있다.

III. 태깅 행태 분석

본 연구에서는 협업태깅 서비스의 사용자 행태를 분석하기 위해 국내 대표적인 서비스인 M사의 소셜 북마킹 서비스를 크롤링하여 데이터를 취득하였다. 취득된 데이터에는 무작위로 추출된 1862명의 사용자가 지난 1년 2개월동안의 북마킹한 70089건의 웹문서와 78164건의 태그가 포함되어 있다.

협력적 태깅서비스의 사용자는 자신이 보관하고 싶은 웹문서를 북마킹하면서 태그를 부여하게 되는데, 사용자가 다양한 태그를 부여할수록 더 많은 지식이 축적된다. 지식이 많이 축적될수록 정교한 검색이 가능해져, 협력적 태깅서비스의 질이 향상된다. 그러나 어떤 웹문서에는 많은 종류의 태그가 부여되는 반면 어떤 웹문서에는 소수의 태그만이 부여되기도 하고, 어떤 웹문서에는 태그가 전혀 없는 상태로 북마킹 되기도 한다. 다음의 [그림 2]를 보면 0~3개의 태그가 부여되는 웹문서의 비율이 가장 높지만, 10개 이상의 태그가 부여되는 웹문서도 상당수 존재한다.

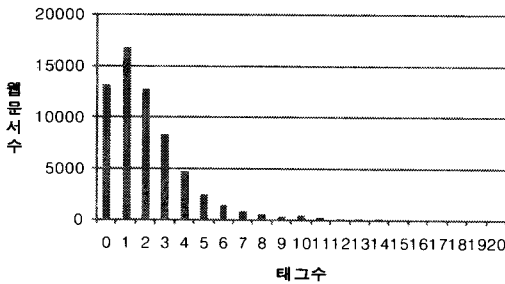


그림 2. 웹문서의 태그갯수 분포

웹문서에 따라 부여되는 태그의 수에 차이가 발생하는 영향요인은 여러 가지가 있을 수 있다. 본 장에서는 웹문서에 부여되는 태그수의 영향요인과 사용자가 부여하는 태그수의 영향요인으로 구분하여 분석해보고자 한다.

1. 웹문서의 태그수 영향요인

1.1 북마킹한 사용자수와 웹문서의 태그수와의 관계

협력적 북마킹 서비스에서는 동일한 웹문서를 여러 사람이 동시에 북마킹할 수 있기 때문에, 같은 웹문서에도 사용자마다 다른 태그를 다양하게 부여할 수 있다. 사용자들은 자신의 경험과 지식에 기반하여 웹문서와 연관된 태그를 부여하게 된다. 이것은 하나의 객체에 대해 다양한 사람들이 협력하고 경쟁하여 지식이 통합되고 공유되는 과정을 보여준다. 다음의 [그림 3]은 이와 같이 지식이 축적되는 과정을 분석한 것으로, 어떤 웹문서를 북마킹한 사용자수와 그 웹문서에 부여된 태그수와의 관계를 나타낸 그림이다.

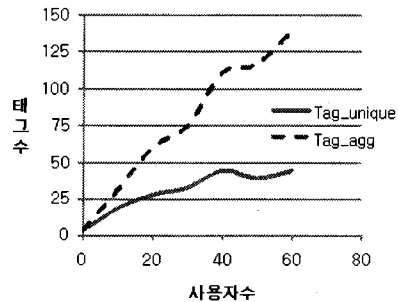


그림 3. 웹문서의 사용자수와 태그수간의 관계

복수의 사용자가 동일한 웹문서에 같은 태그를 부여할 수 있는데 tag_unique는 이것을 한 개의 태그로 간주한 개수이고, tag_agg는 이들을 모두 별도로 누적한 개수이다. tag_agg가 tag_unique보다 2배 이상 큰 값을 보이고 있는데, 이는 같은 웹문서에 대해 여러 사용자가 동일한 태그를 부여하기 때문에 발생하는 현상이다. 여러 사용자가 동시에 부여한 태그는 해당 웹문서에 대한 설명력의 신뢰도가 높은 태그로 볼 수 있다. 이와 같이 증가율의 차이는 있지만 두가지 변수 모두 북마킹한 사용자가 증가할수록 커지는 것을 알 수 있으며, 이것은 더 많은 사용자 집단이 참여할수록 지식의 양이 증가한다 라는 것을 보여주고 있다.

1.2 경과시간과 태그수와의 관계

협력적 북마킹 서비스에서 사용자는 웹문서를 북마킹한 후에도 자신이 북마킹한 웹문서에 대해 태그를 추가하거나 수정할 수 있다. 시간이 경과함에 따라 웹문

서에 부여된 태그의 수가 증가하는 것은 두 가지 측면으로 나눠볼 수 있는데, 첫 번째로 해당 웹문서를 북마킹하는 사용자가 증가함에 따라 자연스럽게 태그의 수도 증가하게 된다. 두 번째는 이미 북마킹을 한 사용자가 태그를 추가하는 경우이다. 첫 번째 측면은 앞에서 분석이 되었고, 두 번째 측면을 분석한 결과 총 70089건의 북마킹 중 최초의 북마킹 후 태그의 추가가 발생한 경우는 312건으로, 1%가 안 되는 미미한 수준으로 나타났다. 이는 대부분의 사용자들이 처음 북마킹할 때 태깅을 완료하며, 이후에는 거의 태그를 추가하거나 갱신하지 않는다는 것을 의미한다.

자신의 북마크 태그를 수정한 312건의 사례에 대해 분석해보면, [그림 4]에처럼 최초 북마킹 후 한달 이내에 추가로 태깅한 경우가 대부분으로 312건 중의 55%를 차지하고 있었다. 이 결과를 볼 때, 사용자는 한번 부여한 태그를 시간이 경과한 후 수정하거나 추가하는 경우가 거의 없음을 알 수 있다. 이는 시간이라는 변수는 웹문서의 태그수에 거의 영향을 주지 않는 것을 의미한다.

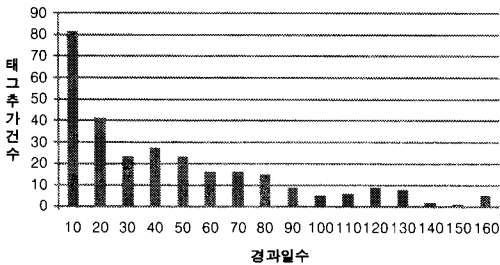


그림 4. 시간경과에 따른 태그추가수

1.3 웹문서의 종류와 태그수의 관계

협력적 태깅은 통제된 어휘 중 선택되는 것이 아니라 사용자가 자유롭게 태그용어를 개수에 제한 없이 부여할 수 있기 때문에, 웹문서의 종류나 특성은 북마킹할 때 사용자의 태깅 행위에 영향을 줄 수 있다. 여러 가지 태그를 부여하기 좋은 웹문서가 있을 수 있는 반면, 적절한 태그를 찾기 어려운 웹문서도 있을 수 있다.

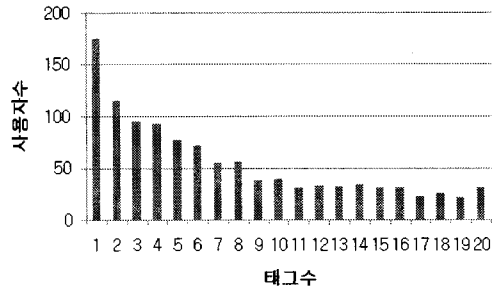


그림 5. 사용자의 태그수 분포

앞의 [그림 2]에서 웹문서들의 태그수 분포를 보면 대부분의 웹문서들이 5개 이하의 태그를 가지고 있으며, 이는 전체 62726개의 웹문서 중 약 92%를 차지하고 있다. 이와 비교하여 [그림 5]는 사용자별 태그수의 분포를 보여주는 그림이다. 5개 이하의 태그를 부여한 사용자가 상대적으로 많기는 하지만, 전체의 48%로 웹문서와 비교하여 훨씬 낮은 비율을 보여주고 있다.

[그림 2]와 [그림 5]를 종합하여 분석하면, 웹문서의 태그수는 대부분 5개 이하 그룹에 몰려있는 반면, 사용자는 사용자의 성향에 따라 보유하고 있는 태그수의 편차가 크다는 것을 보여준다. 분포의 편중도를 보다 자세히 살펴보기 위해 각각의 표준편차와 왜도(skewness)를 측정하였다.

표 1. 사용자와 웹문서의 태그수 분포 비교

	사용자	웹문서
표준편차	76.1	3.2
왜도	4.0	7.0

[표 1]에서 보면 사용자별 태그수 분포의 표준편차가 웹문서별 태그수 표준편차보다 월등히 크고, 왜도는 사용자의 분포가 웹문서의 것보다 작은 것으로 나타났다. 표준편차가 크다는 것은 사용자의 태그수가 웹문서의 태그수에 비해 다양하게 분포되어 있다는 것을 의미한다. 두 분포 모두 좌측으로 치우친 모양을 갖는데, 웹문서의 왜도가 사용자의 것보다 더 큰 것은 더 좌측으로 치우쳐진 모양을 갖는다는 것을 의미한다. 이것은 웹문서의 태그수는 대부분 1~5사이에 집중으로 몰려있는 반면, 사용자의 태그수는 상대적으로 덜 치우친 모양을

갖기 때문이다. 이 결과를 종합하면, 태그가 많이 부여된 웹문서는 웹문서의 특성 때문에 태그가 많이 부여되었다기보다, 그 웹문서를 북마킹한 사용자가 태그를 많이 부여하는 특징을 갖고 있기 때문에 태그가 많이 부여된 것이라고 볼 수 있다.

1.4 웹문서의 태그수에 대한 중회귀분석

북마킹한 사용자수와 경과시간이 웹문서의 태그수에 어떤 영향을 미치는지 알아보기 위해 중회귀분석을 수행한 결과, 두 변수 모두 웹문서의 태그수에 유의한 영향을 미치는 것으로 나타났으며, 경과시간보다는 북마킹한 사용자수가 더 큰 영향을 미치는 것으로 나타났다.

표 2. 중회귀분석 결과

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
constant	1.512	.012		122.6	.000
num_user	.866	.006	.583	148.6	.000
elapsedTime	.008	.000	.179	45.7	.000

다음절에서는 사용자의 태그수에 영향을 미치는 영향을 자세히 알아보려고 한다.

2. 사용자의 태그수 영향요인 분석

2.1 북마킹 문서수와 태그수의 관계

웹브라우저에서 북마킹할때는 태깅을 하지 않기 때문에, 이것에 익숙한 사용자들이 처음 협력적 북마킹 서비스를 이용할 때는 태그를 부여하지 않고, 웹문서의 제목과 주소만을 저장하는 경우가 많을 것으로 예상된다. 하지만, 북마킹한 웹문서가 많아질수록 자신이 북마킹한 웹문서도 검색을 해서 찾게 되는 경우가 발생하는데, 이때 정확한 태그를 부여한 웹문서는 손쉽게 빠르게 찾을 수 있게 된다. 사용자들이 이러한 경험을 하게 될수록 태깅의 중요성을 깨닫게 되고, 향후에 북마킹할 때 태깅에 더 많은 노력을 기울이게 된다.

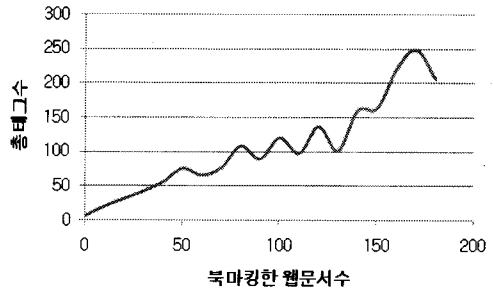


그림 6. 북마킹웹문서와 태그갯수와의 관계

[그림 6]은 사용자가 북마킹한 웹문서의 수가 증가할수록, 그 사용자가 등록한 총 태그의 개수간의 관계를 보여주고 있다. 북마킹한 웹문서가 많은 사용자일수록 그동안 부여한 누적태그수도 증가하는 것으로 나타났으며, 이는 북마킹 경험량이 지식생산량에 영향을 미친다는 것을 의미한다.

2.2 시간과 사용자 태그수와의 관계

시간이 경과함에 따라 북마킹 횟수가 증가하게 되면 [그림 6] 에서처럼 태그의 수도 증가하게 된다. [그림 7]은 사용기간에 증가함에 따라 사용자가 등록한 태그의 수가 증가하는가를 알아보기 위한 그림이다. 서비스를 사용한 시간이 경과할수록 누적태그의 수가 증가하는 하지만, 이것은 북마킹한 웹문서가 증가함에 따라 자연스럽게 증가한 수치일 수 있다. 그래서 표준화된 비교를 위해 웹문서당 태그의 수를 분석하였다. [그림 8]은 사용자들의 서비스 이용시간과 웹문서당 태그의 수의 관계를 보여주는 그림이다.

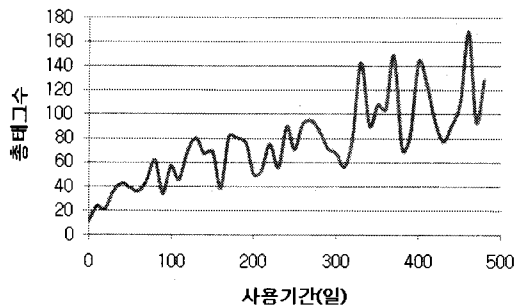


그림 7. 사용기간과 누적태그수와의 관계

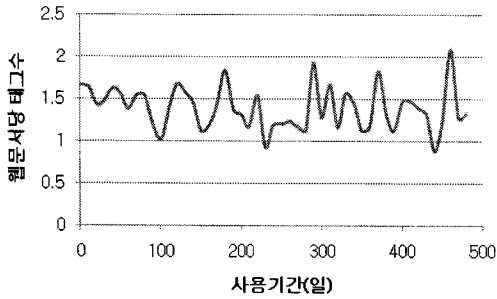


그림 8. 사용기간과 웹문서당 태그수의 관계

[그림 8]을 보면 사용기간과 상관없이 웹문서당 태그의 수는 1~2사이의 값을 갖는 것을 볼 수 있다. 서비스를 이용한지 1년이 넘는 사용자들도 초기 사용자들과 비슷하게 15개 정도의 태그를 웹문서에 부여하고 있다. 이는 시간이 경과해도 사용자의 웹문서당 태그 부여수는 차이가 없음을 보여준다. 즉, 시간이라는 변수는 사용자의 태그 생산능력에 큰 영향을 미치지 않는 것으로 볼 수 있다.

2.3 선도적/추종적사용자와 태그수와의 관계

협력적 북마킹 서비스에서는 다사용자의 북마킹행위를 엿볼 수 있고 인기 웹문서의 순위를 제공하기 때문에 다른 사용자들이 북마킹된 웹문서만을 서핑하는 것이 가능하다. 이러한 사용자는 추종적 사용자(following user)로 새로운 웹문서를 북마킹하기보다, 이미 검증되어 많이 알려진 웹문서들을 주로 북마킹한다. 반면 선도적 사용자(leading user)는 새로운 웹문서를 적극적으로 찾아서 처음으로 북마킹하는 경우가 많다. 이들은 자신이 웹서핑을 하면서 발견한 좋은 웹문서들을 적극적으로 북마킹하며, 이들이 북마킹한 웹문서들이 타 사용자들에게 널리 전파시키는 역할을 수행한다. 이와 같은 선도적 사용자와 추종적 사용자를 구분하는 방법은 웹문서를 북마킹한 시간적 순위를 이용한다. 어떤 웹문서를 북마킹한 사용자가 여러 명 있을 때 가장 먼저 북마킹한 사용자를 선도적 소비자라 볼 수 있으며, 이후에 북마킹한 사용자는 추종적 사용자로 분류할 수 있다.

선도적 사용자와 추종적 사용자를 구분하기 위해 사

용자별로 북마크한 웹문서 중 가장 먼저 북마크한 웹문서의 비율을 측정하였다. 이것은 자신이 북마크한 웹문서들 중 자신이 가장 먼저 북마크한 웹문서의 비율을 의미한다.

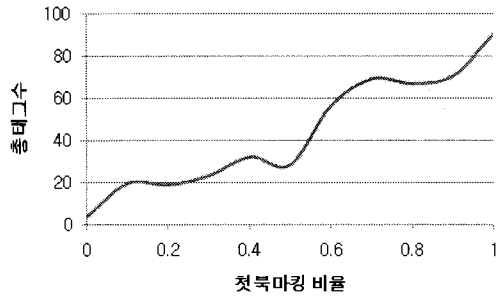


그림 9. 첫북마킹 웹문서 비율

[그림 9]는 사용자의 첫북마킹 웹문서 비율이 증가함에 따라 해당 사용자의 태그수가 어떻게 변화하는지는 보여주는 그림이다. 첫북마킹한 비율이 높을수록 해당 사용자가 등록한 태그의 개수가 증가하는 것을 볼 수 있다. 보다 정확한 검증을 위해 이 데이터를 이용해 회귀분석을 실시하였다. 독립변수는 첫북마킹 비율이고, 종속변수는 해당 사용자의 태그수로 단순회귀분석을 실행한 결과는 [표 3]과 같다.

표 3. 회귀분석 결과

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
constant	11.11	3.65		3.04	.002
First_URL	52.68	5.49	.21	9.59	.000

회귀분석의 t검정결과 유의도가 각각 0.002와 0.000으로 나타나 단순회귀모형의 유의성이 높은 것으로 나타났다. 첫북마킹의 비율이 태그의 개수에 비례한다고 볼 수 있다. 이것은 같은 수의 웹문서를 북마킹하더라도 선도적 사용자가 추종적 사용자보다 더 많은 태그를 생성하여 지식생산에 기여한다는 것을 의미한다.

2.4 사용자의 태그수에 대한 중회귀분석

북마킹한 웹문서수와 사용기간 및 첫북마킹 비율이 사용자의 태그수에 미치는 영향을 알아보려고 중회귀 분석을 수행하였다. [표 4]의 분석 결과를 보면, 3가지 변수 모두 유의한 것으로 나타났으며, 북마킹한 웹문서의 수가 가장 크게 영향을 미치는 것으로 나타났다.

표 4. 중회귀분석 결과

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
constant	-.927	1.749		-.530	.596
First_URL	.734	.152	.084	4.824	.000
elapsedTim	.102	.011	.181	9.386	.000
bookmark	.798	.028	.550	28.114	.000

IV. 결론

본 연구는 웹서비스에서 점차 확산되고 있는 태깅 기능이 지식의 축적에 기여한다는 점에 착안하여, 이를 더욱 활성화시키기 위한 방안을 찾아보고자 시작되었다. 사용자가 생성하는 태그는 웹문서의 내용과 특성을 잘 설명해주는 메타정보의 역할을 수행한다. 이는 사용자가 자신이 원하는 웹문서를 검색할 때 매우 유용한 매개자 역할을 수행한다. 이와 같은 태그들은 많이 생성될수록 더 정교한 검색을 이루어낼 수 있다. 협력적 북마크 서비스 운영자들은 이러한 태그들이 더 많이 생성되기를 기대하기 때문에, 본 연구에서는 이를 위한 실천적 방안을 찾고자 태깅 기능이 어떻게 이용되고 있으며, 웹문서와 사용자의 태그의 다양성에 영향을 미치는 변수들을 분석하였다. 본 연구는 기존 연구에서 분석되지 못했던 사용자의 태그수와 웹문서의 태그수를 구분하여 영향요인을 탐색하였고, 복수개의 영향 요인들 중 영향력이 큰 변수가 무엇인가에 대한 비교분석을 실시한 점이 기여점이라고 할 수 있다.

웹문서의 태그수에는 웹문서의 특성보다 사용자의 특성이 더 많은 영향을 미치는 것으로 나타났다. 웹문서의 종류에 따라 총 태그수의 차이웹문서별 태그수의

편차보다 사용자별 태그수의 편차가 더 크다는 것은 사용자가 태그를 부여할 때 웹문서의 종류나 내용보다는, 사용자의 성향에 따라 부여하는 태그의 개수가 결정된다는 것을 의미한다.

사용자의 태그수에 영향을 미치는 요인에 대한 분석에서는 첫 번째로 해당 사용자가 그동안 북마킹한 웹문서의 개수가 태그의 개수에 영향을 미치는 것으로 나타났다. 많은 웹문서를 등록한 사용자일수록 하나의 웹문서에 더 많은 태그를 부여하는 것으로 분석되었다. 두 번째 요인으로는 사용자의 능동성으로, 가장 먼저 등록된 웹문서가 많을수록 더 많은 태그를 부여하는 것으로 나타났다. 가장 먼저 등록된 웹문서가 많다는 것은 적극적으로 서비스를 이용하고 있다는 것을 의미하며, 이러한 사용자들이 다른 사용자들에 비해 더 많은 종류의 태그를 웹문서에 부여하고 있었다.

본 연구의 결과를 통해 실천적 방안을 찾는다면 협력적 북마킹 서비스에서 태그의 생성을 활성화시키기 위해서는 수동적 사용자들에게 많은 웹문서를 북마킹하도록 유도하는 것보다 능동적 사용자에게 보다 많은 인센티브를 부여하여 콘텐츠 생성활동을 지원하는 것이 더 좋은 효과를 볼 수 있을 것이다.

본 연구는 사용자의 인구통계학적 정보나 웹문서의 내용등과 같은 속성정보를 분석에 활용하지 않은 한계를 갖고 있다. 향후 연구로 추가적인 속성정보들을 분석대상에 포함시킨다면 본 연구에서 발견한 요인보다 더 큰 영향을 미치는 변수들을 찾을 수 있으리라 기대한다.

참고문헌

[1] S. Golder and B. Huberman, "Usage patterns of collaborative tagging systems," J. of Information Science, Vol.32, No.2, pp.198-208, 2006.
 [2] 이재운, 황혜경, "이용자 생성 메타데이터에 관한 연구", 정보관리연구, 제37권, 제3호, pp.1-24, 2006.
 [3] Q. Li and S. Lu, "Collaborative tagging

applications and approaches," IEEE Multimedia, Vol.15, No.3, pp.14-21, 2008.

[4] H. Wu, M. Zubair, and K. Maly, "Harvesting social knowledge from folksonomies," Proceedings of Hypertext06, 2006.

[5] 이정미, "폭소노미의 개념적 접근과 웹 정보 서비스에의 적용", 한국비블리아학회지, 제18권, 제2호, pp.141-159, 2007.

[6] <http://kr.dir.yahoo.com/index.html>

[7] C. Yeung, G. Ching, and N. Shadbolt, "Understanding the Semantics of Ambiguous Tags in Folksonomies," Proceedings of International Workshop on Emergent Semantics and Ontology Evolution, 2007.

[8] K. Dellschaft and S. Staab, "An Epistemic Dynamic Model for Tagging Systems," Proceedings of the 19th ACM Conference on Hypertext and Hypermedia, 2008.

[9] Y. Pan and D. Millen, "Information Sharing and Patterns of Social Interaction in an Enterprise Social Bookmarking Service," Proceedings of the 41st Hawaii International Conference on System Sciences, 2008.

[10] H. Kim, S. Scerri, J. Breslin, S. Decker, and H. Kim, "Analysis and Representation of Tagging Practices in Online Communities," Proceedings of the AAAI 2009 Spring Symposium on Social Semantic Web, 2009.

[11] M. Kipp and D. Campbell, "Patterns and Inconsistencies in Collaborative Tagging Systems: An Examination of Tagging Practices," Proceedings American Society for Information Science and Technology, 2006.

[12] H. Halpin, V. Robu, and H. Shepherd, "The Dynamics and Semantics of Collaborative Tagging," Proceedings of the 1st Semantic Authoring and Annotation Workshop, 2006.

[13] U. Farooq, T. Kannampallil, Y. Song, C. Ganoë,

J. Carroll and C. Giles, "Evaluating Tagging Behavior in Social Bookmarking Systems: Metrics and design heuristics," Proceedings of the international ACM conference on Supporting group work, 2007.

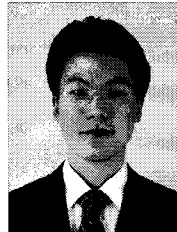
[14] S. Golder and B. Huberman, "The Structure of Collaborative Tagging Systems," HP Labs Technical Report, 2006.

[15] 이강표, 김두남, 김형주, "웹2.0환경에서의 태깅 기술 동향", 정보과학회지, 제25권, 제10호, pp.36-42. 2007.

저자 소개

최준연(Joon-Yeon Choeh)

정회원



- 1996년 2월 : 한국과학기술원 전 기및전자공학과(공학사)
- 1998년 2월 : 한국과학기술원 경영공학(공학석사)
- 2007년 2월 : 한국과학기술원 경영공학(공학박사)

• 2008년 3월 ~ 현재 : 세종대학교 디지털콘텐츠학과 교수

<관심분야> : 지능형 시스템, 집단지성, 데이터마이닝

김용수(Yong-Soo Kim)

정회원



- 1999년 2월 : 한국과학기술원 산업공학과(공학사)
- 2001년 2월 : 한국과학기술원 산업공학과(공학석사)
- 2006년 2월 : 한국과학기술원 산업공학과(공학박사)

• 2008년 9월 ~ 현재 : 성결대학교 e-비즈니스IT학부 교수

<관심분야> : 지능형 콘텐츠 시스템, 추천 시스템