

디지털화 문화유산 태그의 패턴 및 특성 분석

Analysis of Characteristics and Patterns of Tags in Digitized Cultural Heritage

김성희(Seong-Hee Kim)*, 이형미(Hyungmi Lee)**

【초 록】

본 연구에서는 디지털 문화유산에 이용자들이 부여한 태그들의 특성과 패턴을 분석하였다. 이뮤지엄에서 제공하고 있는 '유물정보검색서비스'에서 문화유산 20개를 무작위로 추출하여 100명으로 구성된 이용자 집단에게 추출한 유물 이미지로 작성한 조사지에 자유로이 태그를 달도록 하였다. 수집된 태그 데이터의 분석은 주로 형태적인 특성을 위주로 박물관 디지털 문화유산 태그가 갖는 일반적인 패턴, 박물관 디지털 문화유산 태그의 기능별 구분, 박물관 디지털 문화유산 검색에서의 태그의 활용가능성을 중심으로 이루어졌다. 이러한 연구결과는 박물관의 디지털 문화유산 검색에 태깅 시스템을 이용한 폭소노미 적용을 위한 기초자료로 활용될 수 있을 것이다.

【키워드】

태그, 문화유산, 폭소노미, 웹2.0

【Abstract】

This study analyzed the characteristics and patterns of tags in cultural heritage collections. Works of art were assembled to be tagged, 100 taggers were deployed, and tagging encouraged. The gathered tags were analyzed in terms of general patterns, tagging's functions, and usefulness for information access. The results can be used to improve the access of works of art.

【Keywords】

Tag, Cultural Heritage, Folksonomy, Web2.0

1. 서론

도서관, 기록관, 박물관과 같은 다양한 정보 서비스 영역에서 이용자 중심 서비스 제공을 위하여 웹 2.0 패러다임을 적용하는 움직임이 활발히 전개되고 있다. 일반 이용자 누구나가 능동적으로 콘텐츠를 생성하고 이를 서로 나누는 참여·공유·개방이라는 웹 2.0의 이념을 실현하기 위해 AJAX, OpenAPI, RSS, 블로그, 협력태깅 등과 같은 다양한 기술들이 활용되고 있다.

양질의 디지털 문화유산 콘텐츠 마련의 계기가 된 "국가지식정보자원 관리사업¹⁾"의 중점이 콘텐츠의 '축적 및 보존'에서 '활용'으로 옮겨지면서, 국내 대표 문화유산기관들은 그간 축적해 놓은 양질의 디지털 문화유산들을 활용하여 대국민 정보서비스를 제공하고 있다. 하지만 국가지식정보화사업 전체가 그동안 분야별 정보 콘텐츠의 양적 규모 확충에만 치중하여 체계적인 추진이 이루어지지 못했다(배정현 2007). 이로 인해 양질의 우수한 콘텐츠임에도 불구하고 그 제공 방법에서 서비스 자체의 편리성과 체계성의 부족으로 일반 대중의 활용이 저조한 실정이다. 이미 대다수의 일반 정보 서비스 영역에서 매우 능동적이고 개별화된 서비스를 지향하는데 반해 현재 국내 문화유산 포털에서 제공하고 있는 정보 서비스는 일반인들의 정보수요를 충분히 충족시키지

* 중앙대학교 문헌정보학과 교수(seonghee@cau.ac.kr) (교신저자)

** 중앙대학교 문헌정보학과 박사과정(recluse79@gmail.com) (공동저자)

논문접수일자 : 2009년 8월 29일 논문심사일자 : 2009년 9월 5일 게재확정일자 : 2009년 9월 19일

1) 2000년도부터 국책사업의 일환으로 추진되어온 지식정보자원관리 사업은 국가적으로 보존 및 이용가치가 높은 과학기술·역사·문화·교육 분야 등의 연구보고서·영상물·고문헌 등을 DB로 구축하고, 이 자료들의 상호 연계를 통해 사용자가 쉽게 검색·활용할 수 있도록 지원하는 사업으로 1단계 사업(2000~2004)이 종료되고, 한국정보문화진흥원(KADO)에서 제2단계 사업(2005~2009)을 추진하고 있음.

못하고 있다. 지금까지 주로 연구자들의 학술적 이용에 초점이 맞추어져 있는 이들 서비스의 대중적 활용을 위해서는 콘텐츠의 고품질화와 함께 이를 기반으로 하는 정보서비스의 제공, 합리적인 이용 시스템의 구축 및 서비스 방안이 고려되어야 할 것이다.

이에 본 연구에서는 문화유산의 디지털화와 그 활용에 있어서 일반 이용자들의 이용확대를 위한 방안으로 국내 대표 문화유산포털의 문화재검색서비스에 태깅시스템 도입에 대한 필요성과 그 효과성에 대해 검토해 보고자한다. 이를 위해, 먼저 소셜 태깅과 폭소노미의 개념적 이해와 함께 문화유산의 디지털화와 활용에 대한 문헌연구를 실시하였다. 다음으로, 박물관의 디지털 문화유산에 이용자들이 부여한 태그들의 특성과 패턴을 분석함으로써 디지털 문화유산 검색에 폭소노미와 태깅시스템의 적용 가능성과 적용방안을 제시하고자 한다. 이러한 연구결과는 이용자들이 문화유산에 접근하는데 있어 보다 쉽고 다양한 접근점을 제공해 줄 것으로 기대한다.

2. 이론적 배경

본 연구는 디지털 문화유산의 이용 활성화 방안의 일환으로 기존 박물관 유물검색서비스에 폭소노미 도입의 필요성과 적용가능성(유용성)을 알아보고자 한다. 이에 앞서 본 장에서는 문화유산의 디지털화와 활용, 소셜태깅에 의한 폭소노미의 개요와 적용 사례, 박물관의 디지털 문화유산 및 예술작품의 검색서비스에 폭소노미 도입을 통한 기대효과와 적용 가능성에 대해 문헌연구를 바탕으로 검토해 보았다.

2.1 문화유산의 디지털화와 활용

문화유산의 '디지털화' 및 '디지털 콘텐츠화'란 다음 세대에 상속할 만한 가치가 있는 물질적, 정신적 소산들을 보존하거나 이용에 효율을 높일 수 있도록 전자적인 형태로 변환하거나 혹은 재생산하여 디지털화된 형

식으로 존재할 수 있도록 하는 한 것이다. 디지털 문화유산은 인류의 지식과 표현의 고유 자원을 디지털 형식으로 재생산하거나 기존 아날로그 자료를 디지털화한 것이다.²⁾

2000년을 전후하여 역사와 문화유산 분야의 디지털 콘텐츠화는 CD-ROM, 영상물, 웹사이트 등의 다양한 형태로 시작하였지만, 그 중에서도 정부 주도의 '국가지식정보자원 관리사업(Knowledge Information Management Project)³⁾'은 역사 및 문화유산 자료의 본격적인 데이터베이스화의 직접적인 계기가 되었으며, 이를 통해 전문성과 신뢰성을 갖는 양질의 디지털 문화유산 콘텐츠가 마련되어 이를 기반으로 한 대국민 문화유산 정보서비스의 제공이 가능해졌다. 하지만, 그 동안의 사업이 전문분야별 콘텐츠의 양적 규모 확충에서 그쳤다는 지적과 함께 2005년부터 사업의 중점이 '보존가치'에서 '활용가치'로 옮겨지면서 그간 '축적'된 내용과 결과가 관련 분야에서 실질적으로 활용될 수 있도록 콘텐츠의 고품질화와 함께 합리적인 이용 시스템의 구축과 서비스 방안에 대한 고려가 요구되고 있는 실정이다(한문희 2005).

현재, 우리나라의 대표 문화유산 기관들로는 국립중앙도서관, 국립중앙박물관, 국가기록원, 국가편찬위원회, 한국학중앙연구원, 서울대 규장각, 민족문화추진회, 문화재청 등을 꼽을 수 있다. 이들의 역할은 한국학관련 주요자원들을 수집하고, 소장하고, 연구하며, 디지털화하는 업무를 주로 담당하고 있으며, 아울러 정부 주도의 문화유산 정보화 사업을 통해 같은 유형 및 주제 분야 업무를 담당하는 기관들은 서로 협력하고, 해당 기관의 특성화된 영역에 맞는 디지털 콘텐츠들은 서로 연계해 나간다. 인터넷의 보편화와 멀티미디어 처리 기술의 발전으로 국내 문화유산 기관에서는 축적된 결과물들을 바탕으로 웹을 통해 다양한 서비스를 제공하고 있다. 이러한 디지털 문화유산 콘텐츠의 활용성 증대라는 측면에서 소수의 전문연구자 이외에 보다 다양한 이용자 계층의 접근을 수용하기 위해 이용자 중심의 서비스 제공 방안이 요구되어지고 있다. 이러한 견해에 따라 본 연구에서는 국내 대표문화유산기관에서 제공하고 있는 문화재검색서비스에 폭소노미 기반 태깅 시스템 도입의 필

2) 배정현, 디지털 문화유산 콘텐츠 공동 활용을 위한 협력 방안 연구, 연세대학교 대학원.

3) '국가지식정보자원 관리사업(Knowledge Information Management Project)'은 2000년 지식정보관리법의 제정과 함께 국가차원에서 보존 및 이용가치가 높은 과학기술, 역사, 문화, 교육 분야 등의 연구보고서, 영상물, 고문헌 등을 데이터베이스로 구축하고 상호 연계하여 대국민서비스를 가능하게 함으로써 양질의 디지털 지식정보자원을 일반대중들이 쉽게 검색, 활용할 수 있도록 지원하는 국책사업이다. 현재 정보통신부산하의 한국정보문화진흥원에서 2단계 사업(2005-2009)을 추진 중에 있다.

요성과 그 효과성에 대해 검토해 보고자 한다.

2.2 소셜 태깅으로써 폭소노미

본 장에서는 웹 2.0 시대를 이끌고 있는 새로운 형태의 웹 메타데이터(metadata)인 소셜태깅으로써 폭소노미의 기본 개념과 특징을 살펴보고 폭소노미 기반의 정보 검색서비스 사례를 살펴보았다.

온라인 커뮤니티에서 서로 간의 정보와 지식의 공유를 향상시키기 위해 지난 몇 년간 AJAX, OpenAPI, RSS, 블로그, 협력태깅 등과 같은 각종 웹 2.0 기제가 등장하였다. 그 중에서 오늘날 엄청나게 많고 방대하게 분산되어 있는 웹 상의 정보를 효율적으로 연결하고, 집대성하고, 조직해 지식으로 추출하기위한 대표적인 방법으로 소셜 태깅(Social tagging)과 폭소노미(Folksonomy)가 주목받고 있다.

폭소노미는 전통적인 분류체계인 택소노미(taxonomy)와 대칭되는 용어로서 협력태깅(collaborative tagging), 사회적 분류(social classification), 사회적 색인(social indexing), 소셜태깅(social tagging) 등으로 불리우며 ‘사람들에 의한 분류법’(folk + order + nomos)이라는 뜻의 합성어이다(이정미 2007). 이는 인터넷 상에서 사용자가 사진, 웹페이지, 텍스트 등의 정보를 태그 등을 사용해 자유롭게 체계화하는 방식을 일컫는다. 이와 같이 개개의 사용자가 자유롭게 참여하여 모아진 정보는 여러 사람에 의해 수정, 첨가되는 과정을 거쳐 더욱 가치 있는 정보가 된다.⁴⁾ 폭소노미는 전통적 분류체계와 같이 소수의 전문가 집단에 의해 논리적으로, 모든 주제를 포괄하기위해 고안된 것은 아니지만, 오늘날의 웹 2.0 환경에서 더욱 동적이고 유연하게 대응할 수 있다. 또한, 폭소노미는 계층 구조가 아니라 사용자의 주관에 따라 태그라는 메타데이터를 붙여서 분류하는 방식으로 하나의 콘텐츠에 여러 개의 태그를 붙일 수 있으며 시간의 흐름에 따라 관련 콘텐츠 간의 연관 관계가 항상 변화될 수 있는 구조를 가진다.

협력태깅은 이러한 폭소노미를 대표하는 시스템이라고 할 수 있는데, 이는 이용자 스스로 탐색 가능한 키워드를 가지고 디지털 정보 각각을 태깅함으로써 이용자 생성메타데이터가 늘어가는 현상을 일컫는다. 태그는 특정 정보에 대해 이용자가 직접 의미를 부여하고 기술하

는 이용자 기반 메타데이터(user-generated metadata)이며 하나의 주제만을 포함하는 1:n 구조로서 다른 사람과 생각을 공유할 수 있는 키워드 개념과 매우 유사하다. 이미지나 해당하는 정보 등에 주제나 카테고리의 형태로 키워드 처리를 함으로써, 기존의 태그를 보면서 참여자들 간 연관된 정보를 부여하여 그 정보가 무엇인지를 알 수 있도록 하는 수단으로 활용된다.

폭소노미를 이용하는 태깅시스템이 가지는 장점은 다음과 같다. 첫째, 이용자 등록 태그는 이용자가 자신의 정신 상태나 배경지식에 바탕하여 스스로 자유롭게 선택한 주제어를 레이블링하며 정보를 분류하기 때문에 콘텐츠 공급자가 미처 생각하지 못했던 다양한 디렉토리가 생성되게 된다. 이렇게 생성된 태그들은 사회성을 반영하므로 사용자들이 원하는 정보를 보다 정확하게 제공할 수 있다.

둘째, 자동 검색 엔진으로 구분하기에는 너무나 방대하고 급속히 변화하는 오늘날의 인터넷 정보들에게 폭소노미 태깅으로 누구나가 쉽고, 재미있고 신속하게 메타데이터를 할당할 수 있다.

반면, 폭소노미를 이용하는 태깅시스템의 문제점은 첫째, 폭소노미는 수평적 분류이기 때문에 이용자가 선정한 용어들의 상위범주, 하위범주 등에 대한 개념과 구분이 불분명하다. 이에 따라 용어의 단수/복수, 대문자/소문자와 같이 형태에 따라 동일한 의미이지만 다르게 인식되는 문제가 생겨나며, 너무 일반적이거나 혹은 너무 구체적인 단어를 태깅에 사용하여 검색에서의 정확률을 떨어뜨리기도 한다.

둘째, 다양한 사용자 참여로 이루어진 태그들은 동적인 구조를 가짐으로써 소수의 전문가에 의한 체계적인 분류체계를 형성하는 것은 어렵다. 이 때문에 특정 태그에 대한 단편적인 관련 콘텐츠를 검색하기는 용이하지만 태그와 태그와의 연관성을 고려하여 정확한 결과를 검색하는 것이 어렵다(김운용, 박석규 2007).

폭소노미에 따른 웹 분류는 이와 같은 한계점에도 불구하고 오늘날의 웹 환경에서 기존시스템의 정보조직과 검색을 보조할 수 있는 유용한 메타데이터로써 웹상의 특정 주제, 특정 형태의 정보 접근에 부분적으로 효과적인 것으로 평가되고 있다(이정미 2007). 특히 사진과 같은 디지털 이미지나 음악, 동영상 등의 멀티미디어 정보에 대해 정확한 내용의 검색이 어렵기 때문에 입력된 태그를

4) Wikipedia <<http://en.wikipedia.org/wiki/folksonomy>> [cited 2009.7.15].

이용하여 이들의 분류를 보다 수월하게 해주며 검색 가능성을 향상시켜주는 기능을 한다(배주희, 이경원 2009).

‘대중에 의한 분류’로 이해되는 폭소노미가 주는 용이하면서도 즉각적인 사회적 피드백으로 인해 최근 태깅 시스템을 활용하고자 하는 노력들이 다수의 웹서비스 분야에서 활발하게 진행되고 있다. 이들의 가장 근본적인 목적은 좀 더 많은 다양한 사용자들의 활동을 끌어들이는 것이다. 폭소노미가 실제 웹 정보서비스에서 적용되는 사례들은 크게 웹 정보의 공유를 위한 소셜 북마킹 형태와 이용자에 의해 생산된 사진, 동영상과 같은 콘텐츠의 공유를 위한 서비스에 적용되는 형태로 나눌 수 있다(유시내 2007). 이들의 대표적인 적용사례로는 태깅을 통해 웹 정보를 공유하는 서비스인 Delicio.us⁵⁾와 Diigo⁶⁾가 있으며, 태깅을 기반으로 블로그와 포스트를 검색해 주는 블로그 검색엔진 Technorati⁷⁾가 있으며, 멀티미디어 콘텐츠의 공유 서비스 사례로는 Flickr⁸⁾가 있다. 그 밖에도 소셜태깅 방식을 학술분야에 도입하여 연구자 및 과학자들의 연구 활동 및 학술정보 공유를 지원하는 connotea⁹⁾와 CiteULike¹⁰⁾ 같은 온라인 참고문헌 관리 도구도 있다(황혜경, 이재윤 2006).

2.3 선행연구

본 연구와 관련하여 박물관의 디지털 문화유산 검색과 관련한 선행연구와 박물관 유관기관인 도서관, 미술관, 기록관의 검색 서비스에 폭소노미 도입에 관한 선행연구들을 살펴보았다.

박물관의 디지털 문화유산 검색에 관한 선행연구들은 크게 박물관과 메타데이터, 이미지 색인과 검색이라는 두 분야로 나누어 이루어지고 있다(박진철 외 2004). 박물관의 디지털 문화유산은 대부분 물리적인 박물관 컬렉션을 디지털 이미지화한 것으로 디지털 이미지 검색과 깊이 연관되어 있다고 할 수 있다. 디지털 문화유산의 이미지 데이터에 대한 기술과 접근의 문제는 텍스트 형태보다 복잡하다. 이미지는 개념적, 물리적, 시각적 수준 등으로 다양하게 기술될 수 있고, 이러한 특성이

내·외재적으로 담겨있으며, 시·공간 의존적이고, 색인자나 이용자에 따라 가변적이라는 점에서 색인과 검색 등에 어려움이 많다(박찬표 외 2007). 게다가 이미지는 텍스트에 비해 훨씬 더 중의적이고 함축적인 정보이기 때문에 디렉토리나 태그 등에 의해 완전하게 분류되기가 힘들며, 전문색인자의 내용기술정보를 기반으로 한 검색이 보다 효과적일 수 있다고 주장하는 견해도 있다(Holt & Hartwick 1994). 또한 대용량 이미지 데이터베이스에 대해 이미지 내용분석, 내용기술, 색인, 그리고 검색기법을 적용하기 위해서는 다양한 기술의 통합이 필요하다. 따라서 현재 대부분의 문화유산포털 문화재검색 서비스에서는 주로 박물관 컬렉션의 디지털 이미지에 대한 기본적인 박물관 서지정보에 기반 한 텍스트 검색만을 지원하고 있다. 이 같은 이유로 대부분의 유물정보 검색서비스에서 보다 정확한 검색결과를 도출하기 위해서는 찾고자 하는 유물의 서지사항에 해당되는 ‘유물의 명칭’, ‘작자/제작처’, ‘국적/시대’ 등에 관해 가능한 정확하게 숙지하고 있어야 한다.

박물관 디지털 문화유산 검색의 효율성을 높이기 위해서는 박물관에서 기존에 제공하고 있는 기초적 서지 기술정보와 함께 물리적 컬렉션의 유형별로 상이한 특성정보, 의미적 내용기술 정보를 추가적으로 활용해주어야 한다(김녹환, 김녹수 2002). 이와 같이 박물관 메타데이터가 미처 기술하지 못한 부분들을 이용자생성 태그에서 잘 반영할 수 있다면 박물관 디지털 문화유산 검색에서 유용한 역할을 할 수 있을 것으로 본다.

태그 이용 패턴을 분석한 연구가 다수 알려져 있지만, 텍스트 기반 웹 자원에 대한 태그 분석 연구가 그 중의 대부분을 차지한다는 것이 제한점이기도 하다(Angus et al. 2008). 이미지 혹은 멀티미디어 기반 자원이 우세한 오늘날의 웹 환경에서 이미지 혹은 멀티미디어 기반 자원에 대한 태그 분석 연구가 보다 중요할 것으로 판단된다.

일반적인 웹 정보 서비스 영역에서 뿐 아니라 도서관, 기록관, 박물관 소장 컬렉션의 디지털화가 증가하면서, 이들 기관에서도 양질의 디지털 문화유산에 대한 이용

5) <<http://www.delicious.com>>

6) <<http://www.diigo.com>>

7) <<http://technorati.com/>>

8) <<http://www.flickr.com>>

9) <<http://www.connotea.org>>

10) <<http://www.citeulike.com>>

자들의 보다 수월한 접근과 효율적인 사용을 도모하는 다양한 노력들이 활발히 시도되고 있다. 도서관 및 기록관, 박물관에서의 이용자 기반 태깅 시스템과 폭소노미 적용과 관련한 선행연구는 다음과 같다.

조재인(2008)의 도서관 정보 수요자를 위한 소셜 네트워크 서비스의 도입에 관한 선행연구에 따르면 폭소노미를 적용한 서비스 영역의 범위가 확장되면서 소셜 북마킹, 소셜 온라인 목록, 참고 문헌 공유 등의 폭소노미가 도서관 정보 서비스에 적용되는 사례가 늘고 있으며, 대표적인 사례로 미국 Pennsylvania 대학의 Penntag¹¹⁾와 Ann Arbor District Library의 SOPAC¹²⁾, 그리고 국내의 사례로는 성균관대 중앙학술정보관의 학술 온라인 소셜 북마킹 서비스인 SKKUBee¹³⁾ 등이 있다. 이성숙(2008, 2009)은 선행연구를 통해 실제 폭소노미를 도입하여 라이브러리 2.0을 구현하고 있는 대학도서관의 폭소노미 태그 데이터의 형태적 특성과 일반적 패턴을 분석하여 디지털 도서관에서 폭소노미 태그의 활용 방향에 관한 토대를 마련하였다.

미술관에서도 온라인 정보 서비스에 폭소노미를 도입하려는 시도가 이루어지고 있는데, 스티브 미술관 프로젝트(steve. museum project)¹⁴⁾가 그 대표적인 사례이다.

Trant(2006)의 연구에 따르면 스티브 미술관 프로젝트 참여 미술관들을 대상으로 폭소노미 적용 가능성 조사를 실시한 결과 미술관 컬렉션의 검색서비스에서 이용자생성 태그는 온라인 컬렉션에 대한 접근의 외연을 넓혀 새로운 관점의 접근점을 제공함으로써 검색을 향상시킬 수 있으며, 컬렉션에 대한 전문가 생성 메타데이터를 보완하고 보충해주는 역할도 할 수 있을 것으로 평가되었다. 아울러 박물관 측은 이용자생성 태그로부터 수집된 이용자 정보를 통해 이용자의 요구 및 이용 이력, 배경지식정도 등을 파악함으로써 이를 이후 박물관 행정이나 박물관에서의 이용자교육 등에 활용할 수 있게 될 것으로 분석하였다.

기록관 영역에 있어서도 폭소노미는 기록정보 서비스에서 제공되는 기존 기록물 검색 도구들이 가지는 한계점을 보완하기 위한 대안으로 제시되고 있다. 이미 국외 기록관에서는 폭소노미를 응용한 기록 검색서비스가 제

공되고 있으며, 국내의 경우 최근 국가기록원의 통합기록검색시스템인 '나라기록검색¹⁵⁾'에 태깅 시스템을 도입한 예가 있다. 유시내(2007)는 이용자 생성 태그를 통해 기록관에서는 방대한 기록의 다양한 측면에 대한 접근점을 제공함으로써 이용자의 검색 만족도를 향상시킬 수 있으며, 태그에 반영된 이용자의 관심과 흥미, 이용 이력 등을 분석하여 이용자 요구에 기반한 정보서비스의 기획과 제공이 가능해 질 것으로 평가하였다.

이상의 연구이외에도 Angus et al.,(2008)은 대학커뮤니티에서 이미지태깅의 유용성을 조사하였으며 Golder & Huberman(2006)의 연구에서는 협력태깅시스템의 구조를 분석하여 사용된 태그들의 패턴등을 조사한 결과 사용된 태그들에게 일정한 규칙성이 있다는 것을 보여주었다.

3. 연구설계

본 연구에서 디지털 문화유산 태그의 분석을 통해 논의해 보고자 하는 연구문제들은 다음과 같다.

- 연구문제 1. 기존의 태그 분석 연구들에서 제시된 태그의 전형적인 패턴들이 박물관 디지털 문화유산 태그에도 나타나는가? 또한 박물관 디지털 문화유산 태그는 일반적으로 어떠한 양상을 갖고 있는가?
- 연구문제 2. 이용자생성 태그를 통해 이용자들이 유물을 바라보는 관점, 혹은 어떠한 관점으로 태그를 부여하고 있는가?
- 연구문제 3. 유물의 유형에 따라 디지털 문화유산 태그의 패턴과 특성에는 차이가 있는가?
- 연구문제 4. 디지털 문화유산에 대한 이용자생성 태그는 박물관 메타데이터의 내용과 다른가? 만약 다르다면 박물관 메타데이터의 내용과 상이한 이용자생성 태그는 박물관 디지털 문화유산 검색에서 활용될 가능성이 있는가?

11) <<http://tags.library.upenn.edu>>

12) <<http://www.thesocialopac.net>>

13) <<http://bee.skku.ac.kr>>

14) <<http://www.steve.museum>>

15) 나라기록검색(<<http://search.archives.go.kr/myhome.htm>>

아직까지 국내 대표 문화유산포털의 문화재검색서비스에서는 폭소노미를 도입·활용한 사례가 없었으므로, 본 분석을 위해 사용될 태그 데이터의 수집은 다음과 같은 방식으로 이루어졌다. 먼저, 국내 대표 문화유산기관이라고 할 수 있는 국립중앙박물관이 관리하고 있는 문화유산포털인 '이뮤지엄'을 선정하였다. 다음으로, 이뮤지엄에서 제공하고 있는 '유물정보검색서비스'에서 문화유산의 디지털 이미지와 서지정보¹⁶⁾를 무작위로 추출하였다. 이를 유물 유형별로 회화(일반회화, 불교회화, 기타종교회화), 조각(일반조각, 불교조각, 기타조각) 공예(생활공예, 불교공예, 기타종교공예), 과학기술¹⁷⁾, 기록유산(전적류, 문서류, 서간류 등)으로 구분하였다. 이 중 분석에 사용된 20개의 유물 이미지는 회화류 7개, 조각류 2개, 공예류 2개, 기록유산류 3개, 과학기술류 6개로 구성되어 있다(표 1 참조). 이들 유물에 대한 이미지를 인쇄하여 태그 수집을 위한 조사지를 작성하였다(표 2 참조). 다음으로, 문헌정보학과 대학원생 및 학부생 100명으로 구성된 임의의 이용자 집단에게 추출한 유물 이미지로 작성한 조사지에 자유로이 태그를 달도록 하였다. 수집된 데이터를 가지고 박물관의 디지털 문화유산 태그의 일반적 패턴 및 특성을 분석하기 위해 관련 문헌 연구에서 제시된 태그 분석기준들을 종합하여 사용하였다. 수집된 태그 데이터의 분석은 주로 형태적인 특성을 위주로 박물관 디지털 문화유산 태그가 갖는 일반적인 패턴, 박물관 디지털 문화유산 태그의 기능별 구분, 박물관 디지털 문화유산 검색에서의 태그의 활용가능성을 중심으로 이루어졌다.


〈표 1〉 유물의 유형별 분류

ID	유물의 유형
#1	기록유산(서간류)
#2	회화(불교회화)
#3	조각(불교조각)
#4	회화(일반회화)
#5	회화(일반회화)
#6	기록유산(서간류)
#7	회화(일반회화)
#8	과학기술(인장)

ID	유물의 유형
#9	과학기술(약장)
#10	과학기술(평교자)
#11	회화(일반회화)
#12	기록유산(서간류)
#13	과학기술(화폐)
#14	회화(일반회화)
#15	조각(불교조각)
#16	공예(생활공예)
#17	회화(일반회화)
#18	공예(생활공예)
#19	과학기술(지도)
#20	과학기술(선사유물)

〈표 2〉 태그 수집을 위해 사용된 조사지의 예

work 8



keyword	no.
1.	9.
2.	10.
3.	11.
4.	12.
5.	13.
6.	14.
7.	15.

4. 태그 데이터 분석 결과

4.1 태그의 일반적인 패턴

본 절에서는 수집된 태그의 수, 태그의 문법적인 형태,

16) 디지털 유물에 대한 '유물상세정보'에서는 물리적 유물 컬렉션의 확대이미지와 함께 유물명칭, 국적/시대, 재질, 크기, 지정 구분, 작자/필자, 용도/기능, 소장처 정보 등과 같은 상세수준의 서지정보를 확인할 수 있으며, 유물에 따라 추가적인 해설을 제공하기도 한다.

17) 과학기술에는 선사유물, 자연과학유물, 생업기술용구, 천문지리기구, 계측교역기용구, 이동운송기술, 공예기술용구, 음식제조용구, 인쇄기술용구, 놀이, 유희용구, 무기병기류 등이 포함된다.

각 유물의 빈도수가 높은 상위 태그를 조사해봄으로써, 박물관 디지털 문화유산 태그가 가지는 일반적인 양상에 대해 분석해 보았다.

100명의 이용자 집단으로부터 총 20개의 유물에 대해 8,834개의 태그를 수집하였다. 유물 이미지 당 평균 442개의 태그를 부여받았고, 1명의 이용자가 작품 당 평균적으로 4.4개의 태그를 달았다. 한편, 전체 태그 8,834개에서 중복 사용된 용어를 제외한 유니크(unique)한 태그의 총수는 3,832개였으며, 이는 유물 이미지 당 평균 192개의 유니크한 태그가 부여된 것이다. 수집된 태그의 수를 중복된 태그를 포함한 전체 태그 수와 중복된 태그를 제외하였을 때로 구분하여 나타낸 결과는 <표 3>과 같다.

수집된 태그 수를 유물 유형별로 살펴보면, 회화의 경우 평균적으로 유물 당 483(중복 태그 제외 시 234)개의 태그가 수집되었으며, 조각의 경우 437(중복 제외 시 206)개, 공예의 경우 436(중복 제외 시 181)개, 과학기술¹⁸⁾의 경우 405(중복 제외 시 168)개, 기록의 경우 381(중복 제외 시 137)개의 태그가 수집되었다.

중복 태그를 포함하는 전체 태그의 수를 비교해보면 이용자들이 회화유물에는 상대적으로 많은 태그를 달았으며, 기록유물에는 상대적으로 적게 태깅한 것으로 나타났다. 중복 태그를 제외한 유니크한 태그(unique tags)의 수와 태그의 중복도를 살펴보면 기록유산의 경우 전체 태그 수는 적은 반면 태그의 중복률은 상대적으로

<표 3> 수집된 태그의 수

ID	유물의 유형	수집된 태그의 총 수		태그 중복률 ¹⁹⁾
		중복 태그 포함 시 (total tags): 8834개	중복 태그를 제외한 유니크 한 태그 (unique tags): 3832개	
#2	회화(불교회화)	462개	238개	
#4	회화(일반회화)	469개	213개	
#5	회화(일반회화)	416개	196개	
#7	회화(일반회화)	454개	227개	
#11	회화(일반회화)	601개	246개	
#14	회화(일반회화)	493개	281개	
#17	회화(일반회화)	478개	239개	
회화류		유물 당 평균 483개(100%)	234개(48%)	52%
#3	조각(불교조각)	454개	206개	
#15	조각(불교조각)	420개	205개	
조각류		유물 당 평균 437개(100%)	206개(47%)	53%
#16	공예(생활공예)	431개	193개	
#18	공예(생활공예)	441개	169개	
공예류		유물 당 평균 436개(100%)	181개(42%)	58%
#8	과학기술(인장)	415개	177개	
#9	과학기술(약장)	441개	175개	
#10	과학기술(평교자)	435개	200개	
#13	과학기술(화폐)	432개	144개	
#19	과학기술(지도)	340개	172개	
#20	과학기술(선사유물)	365개	141개	
과학기술류		유물 당 평균 405개(100%)	168개(41%)	59%
#1	기록유산(서간류)	380개	127개	
#6	기록유산(서간류)	408개	151개	
#12	기록유산(서간류)	356개	132개	
기록유산류		유물 당 평균 381개	137개(36%)	64%

18) 과학기술에는 선사유물, 자연과학유물, 생업기술용구, 천문지리기구, 계측교역기용구, 이동운송기술, 공예기술용구, 음식제조용구, 인쇄기술용구, 놀이, 유희용구, 무기병기류 등이 포함된다.

19) 태그 중복률이란 전체 태그 중 한번씩만 태깅된 유니크한 태그(unique tags)를 제외하고 두 번 이상 부여된 태그들의 비율을 나타낸다.

로 높고, 유니크한 태그의 수는 적은 것으로 나타났다. 반면, 회화류와 조각류의 경우 태그의 중복률은 다른 유형에 비해 상대적으로 낮으며, 유니크한 태그의 수가 높은 것으로 나타났다.

다음으로 수집된 태그 데이터 중 중복 사용된 태그를 제외한 유니크한 태그를 대상으로 문법적인 형태를 분석한 결과는 다음의 <표 4>와 같다. 한글태그의 문법적인 형태는 단일명사, 복합명사(명사 + 명사), 명사구, 형용사, 부사, 동사, 기타로 조사되었다. 기타에 해당하는 태그에는 '어디에 쓰는 물건인가?', '12지가 있다'와 같은 문장형태의 태그, 단순한 숫자태그, 영어단어 태그와 같이 한글의 특정 문법 형태에 해당되지 않는 경우이다.

태그의 문법적인 형태를 살펴보면, 전체 유물유형에 걸쳐 '단일명사', '복합명사', '명사구'의 명사류태그가 전체

태그의 대부분을 차지하고 있는 것으로 나타났다. '형용사'태그의 경우, 다른 유형의 유물에 비해 '회화류'와 '조각류'에서 상대적으로 그 비율이 높은 것으로 나타났다.

아울러, 유물 이미지에 부여된 태그들을 빈도순으로 순위화하여 빈도수가 높은 상위 태그 5개를 순위화 결과는 <표 5>와 같다.

각 유물 별로 빈도수가 높은 상위 태그들을 조사해본 결과 Munk & Mork(2007)의 선행연구에서 언급한 것과 같이 광범위하고 전반적인 내용을 범주화하는 태그는 자주 사용되고 그 외 다수의 태그는 매우 드물게 사용되는 양상을 띄고 있는 것으로 나타났다. 전체 유물에 대한 태그에서 '유물', '문화재'와 같은 태그의 빈도수가 높은 것과, '그림', '서예', '토기', '동양화', '인물화', '도자기', '수묵화' 등과 같이 내용을 범주화하는 태그의 사용

<표 4> 태그의 문법적인 형태

ID	유물의 유형	단일명사	복합명사	명사구	동사	형용사	부사	기타	합계
#2	회화(불교회화)	133	29	35	0	36	1	4	238
#4	회화(일반회화)	121	27	34	4	26	0	1	213
#5	회화(일반회화)	88	38	38	0	30	0	2	196
#7	회화(일반회화)	113	32	50	2	29	0	1	227
#11	회화(일반회화)	123	33	47	3	36	0	4	246
#14	회화(일반회화)	128	41	48	4	58	0	2	281
#17	회화(일반회화)	103	56	30	2	46	0	2	239
회화류		49 (809)	16 (256)	17 (282)	0 (15)	16 (261)	0 (1)	0 (16)	100% (1640)
#3	조각(불교조각)	100	33	24	5	40	2	2	206
#15	조각(불교조각)	89	41	38	2	31	1	3	205
조각류		46 (189)	18 (74)	15 (62)	0 (7)	17 (71)	0 (3)	0 (5)	100% (411)
#16	공예(생활공예)	117	33	20	1	21	0	1	193
#18	공예(생활공예)	98	31	17	0	23	0	0	169
공예류		35 (125)	18 (64)	10 (37)	0 (1)	12 (44)	0 (0)	0 (1)	100% (362)
#8	과학기술(인장)	101	28	25	1	19	1	2	177
#9	과학기술(약장)	98	28	33	0	15	0	1	175
#10	과학기술(평교자)	122	19	33	4	21	0	1	200
#13	과학기술(화폐)	95	10	28	0	9	0	2	144
#19	과학기술(지도)	111	19	31	0	11	0	1	172
#20	과학기술(선사유물)	78	18	30	0	14	0	1	141
과학기술류		60 (605)	12 (122)	18 (180)	0 (5)	9 (89)	0 (1)	0 (8)	100% (1009)
#1	기록유산(서간류)	53	19	49	2	3	0	1	127
#6	기록유산(서간류)	64	30	51	1	3	0	2	151
#12	기록유산(서간류)	69	20	36	0	4	1	2	132
기록유산류		45 (186)	17 (69)	33 (136)	0 (3)	2 (10)	0 (1)	0 (5)	100% (410)

〈표 5〉 빈도수가 높은 상위 태그

유물의 유형	ID	1st.	2nd.	3rd.	4th.	5th.
회화류	#2	불교	부처님	그림	빨기*	화려*
	#4	수목화	신선	바위	조선	그림
	#5	부채	산수화	소나무	풍경화	수목화
	#7	달마	수염	달마도	인물화	수목
	#11	서당	훈장*	김홍도	천자문	글공부
	#14	초가집	산속	겨울	선비	그림
	#17	눈	겨울	벚꽃	눈꽃	빨기*
조각류	#3	불상	금	금동	왕관	유물
	#15	불상	불교	서있*	금	연꽃
공예류	#16	화로	청동	향로	향	유물
	#18	주전자	백자	도자기	뚜껑	꽃
과학기술류	#8	조각	두꺼비	도장	옥	도자기
	#9	청동기	청동	칼	녹*	무기
	#10	가마	나무	양반	의자	귀족
	#13	엽전	동전	동그*	돈	화폐
	#19	지도	섬	옛날지도	고지도	산
	#20	토기	흙	빗살무늬	깨진*	그릇
기록유산류	#1	한자	서예	붓글씨	낙관	김정희
	#6	한자	붓글씨	효	떡	가훈
	#12	서예	한자	붓글씨	한문	세로쓰기

이 빈번하게 사용된 것을 알 수 있다. 특히, 태그 중복률이 높은 기록유산의 경우 조사에 이용된 기록유산들에 공통적으로 이용자들이 가장 많이 태깅한 단어는 ‘한자’, ‘서예’, ‘붓글씨’인 것으로 나타났다.

박물관 디지털 문화유산에 관한 태그의 일반적인 패턴을 분석한 결과를 종합해 보면 첫째, 회화류나 조각류의 유물은 조각이나 기록류, 공예, 과학기술류 유물에 비해 많은 수의 이용자생성 태그가 부여되었으며, 태그의 중복도가 낮고, 유니크한 태그의 수도 많은 것으로 나타났다. 회화류나 조각류의 경우 대부분이 예술작품에 해당되므로, 이 경우 감상에 기반하여 작품을 보고 이용자가 느낀 감정을 기술하고자 하거나, 작품에 사용된 다양한 색채와 질감, 모양 그리고 내포된 의미와 주제에 관해 표현하기 위한 태그의 비율이 높아지므로 태그의 중복률이 낮게 나타나는 것으로 분석된다. 반면, 기록유산의 경우 유물의 이미지가 비교적 단순하고 디테일이 적어서 다양하고 유니크한 태그의 비율이 적은 것으로 분석된다. 이러한 경향은 빈도수가 높은 상위태그의 분석 결과를 통해서도 뒷받침되는데, 기록유산의 경우 광범위하고 전반적인 내용을 범주화하는 일부 한정된 단어의 사용빈도가 높은 것으로 나타났다.

그러나, 이미 사전에 많은 이용자들이 알고 있는 유명

한 회화 유물의 경우에는 예외적으로 이미 사전정보를 가지고 친숙함을 느끼는 이용자들에 의해 많은 수의 태그가 부여되지만, 이미 많은 사람들이 잘 알고 있는 내용일수록 여러 이용자에 의해 중복될 가능성이 높아지므로 태그의 중복도가 높게 나타나기도 한다. 예를 들면 #11은 김홍도의 서당도로 회화류 평균 중복도인 52%보다 높은 60%의 중복도를 나타냈다.

둘째, 태그의 문법적인 형태 분석의 결과 전체 태그 중 명사류(단일명사, 복합명사, 명사구)태그의 비율이 전체 유물 유형 모두에서 우세한 것으로 나타났다. 회화류와 조각류의 경우, 작품에 대한 색채와, 질감, 모양, 느낌 등을 표현하기 위해 형용사 태그의 사용이 상대적으로 높은 것으로 나타났다.

4.2 태그의 기능별 구분

이용자가 직접 입력한 태그는 해당 자원에 대한 이용자의 견해를 반영하고 있으므로, 태그 분석을 통해 이용자가 콘텐츠를 바라보는 관점을 알아낼 수 있다. 따라서 본 절에서는 박물관 디지털 문화유산에 대한 이용자생성 태그를 통해 이용자들이 유물을 바라보는 관점, 혹은 어떠한 관점으로 태그를 부여하는지를 분석해 보았다.

이용자들이 유물의 어떠한 측면을 기술하기 위해 태깅을 하는지를 알아보기 위해 수집된 태그들을 그 기능에 따라 <표 6>과 같이 구분해 보았다. 구체적으로 살펴보면 태그들의 기능을 크게 여덟 가지로 구분하였으며 그 내용을 살펴보면 전기적 태그, 주제태그, 특성묘사태그, 자기참고태그, 복합태그, 특수문자태그, 오탈자태그 등이다. 이런 기준의 선정은 사용자 태그를 기능별로 구분

하기 위해 태그의 기능적 분류 기준을 제시한 Golder & Huberman(2006)과 Angus, Thelwall & Stuart(2008)의 선행연구와 이미지의 분류와 조직화 시에 고려되어지는 네 가지 속성 범주를 제시한 Shatford-Layne, S. (1994)의 선행연구를 참고하였다.

유물별로 태그의 기능을 분석한 결과를 종합하면 다음의 <표 7>과 같다.

<표 6> 태그의 기능별 구분 기준

구분	기능
전기적 태그(Biographical)	유물 자체에 대한 창조자, 생성자, 생성시간과 장소, 생성자가 부여한 이름이나 타이틀 등(유물의 탄생과 관련)에 관한 태그.
	현재 소장되어 있는 곳, 현재 소유자, 유물의 현재 가치, 변경 여부, 현존 혹은 훼손 여부 등(외부로의 여행)에 관한 태그.
주제 태그(Subject & Topic)	무엇(누구)에 관한 것인가에 관한 태그로 유물이 표현하고 있는 시간(time), 공간(space), 활동과 사건(Activities&event), 객체(objects) 등에 대한 태그.
특성 묘사 태그 (Impression & characteristics)	유물과 관련하여 이용자의 감정을 나타내는 태그.(주로 형용사)
	유물의 질감, 색채, 재질 등에 관해 표현하는 태그.(주로 형용사)
자기 참고 태그(Self reference)	이용자 자신의 필요와 관련하여 자기 참고를 위한 태그.
복합태그	2개 이상의 기준에 중복 해당되는 태그.
특수문자 태그	특수문자가 있는 태그.
오탈자 태그	오탈자가 포함된 태그.
기타	어느 범주에도 포함되지 않는 그 밖의 태그.

<표 7> 기능별 태그 구분

ID	유물의 유형	기능별 태그 구분								합계
		전기적	주제	특성묘사	자기참고	복합	특수문자	오탈자	기타	
#2	회화류	28	47	89	0	74	0	0	0	238
#4		19	68	91	1	34	0	0	0	213
#5		30	45	83	0	38	0	0	0	196
#7		51	56	67	3	49	1	0	0	227
#11		37	60	81	3	30	0	4	0	246
#14		46	61	108	1	64	0	1	0	281
#17		62	49	92	0	36	0	0	0	239
#3	조각류	47	34	62	2	51	1	0	0	206
#15		61	38	93	0	32	0	1	0	205
#16	공예류	42	81	19	0	51	0	0	0	193
#18		38	57	44	1	29	0	0	0	169
#8	과학 기술류	61	65	39	0	12	0	0	0	177
#9		64	54	27	0	28	1	1	0	175
#10		67	67	31	0	31	3	1	0	200
#13		51	48	30	0	13	2	0	0	144
#19		38	61	42	0	30	1	0	0	172
#20		58	43	24	0	15	1	0	0	141
#1	기록 유산류	65	19	29	0	12	0	2	0	127
#6		76	45	23	0	7	0	0	0	151
#12		23	61	21	0	27	0	0	0	132

회화류와 조각류의 경우 색채와 질감 그리고 모양 등을 표현하기 위한 '특성묘사 태그'와 '주제태그'의 사용 비율이 높은 것으로 나타났으며, 과학기술류와 공예류에서는 전기적 태그와 주제 태그의 비율이 보다 높은 것으로 나타났다.

보다 자세히 살펴보면, 회화와 조각의 경우 감상에 기반하여 전달하고자 하는 느낌이 중요시되기 때문에 특히 색채, 내포된 의미와 주제에 관해 표현하거나, 이미지를 보고 이용자가 느낀 감정을 기술하기 위한 '특성 묘사 태그'의 비율이 상대적으로 높은 것으로 나타났다. 이러한 특성 묘사 태그의 경우 색 표현이나 감정 표현을 위해 주로 형용사 태그를 사용하였는데, 이 경우 동일한 의미를 가지지만 다양한 형태를 띠는 형용사의 사용(ex. 고요한/조용한/적막한, 빨간/붉은)과 어미의 형태가 조금씩 다른 형용사적 표현(ex. 빨가*: 빨강다/빨간/빨강, 복잡*: 복잡/복잡한/복잡하다)이 빈번하였다. 또한 이들 형용사와 결합된 '복합태그'의 비율도 높은 것으로 나타났다.

공예나 과학기술류 유물의 경우에는 그 쓰임새나 사용 목적을 나타내는 '주제 태그'나 유물의 현재 가치나 현존 여부 등을 나타내는 기능을 하는 '전기적 태그'의 사용 비율이 높은 것으로 나타났다. 이러한 전기적 태그나 주제 태그의 경우 대부분 명사 혹은 명사구 형태의 태그가 대부분을 차지했다. 그 밖에도 주제 태그의 사용과 관련하여서는 너무 일반적이거나 혹은 너무 구체적인 형태를 띠는 경우나 같은 개념을 지칭하는 한자와 한글 표현, 명사의 단수와 복수 표현이 다른 경우가 많았다. 기록유산류에 속하는 유물의 경우에는 무엇에 관한 기록인가에 대한 기술을 위해 '주제 태그'의 이용이 빈번한 반면, 기록유산과 관련하여 이용자의 감정이나 유물의 상태에 대해 기술하는 '특성 묘사 태그'의 비율은 저조한 것으로 나타났다.

한편, 태그의 기능별 구분과 관계없이 해당 유물에 대해 틀린 내용을 태깅하는 사례 또한 빈번한 것으로 조사되었는데, 예로 기록유산의 경우 잘못된 서체정보나 틀린 한자독음 기재와 같은 경우가 많았으며, 고대 유물의 경우 시대구분에 관해 틀린 내용을 태깅하는 경우가 있었다. 하지만, 실제 유물에 대해 틀린 내용의 태그 역시 그것이 많은 이용자들이 의해 중복적으로 태깅된 것이 라면 유물 검색 시에 의미 있는 태그로 받아들여질 수 있을 것이며, 이러한 내용은 이후 박물관에서의 이용자 교육에 참고자료로 활용될 수 있을 것이다.

따라서, 이용자생성 태그를 기능별로 구분한 후 유물

별로 정리해보면, 유물의 유형에 따라 디지털 문화유산에 대한 협력태깅 패턴과 태그의 특성이 서로 다르다는 것을 확인할 수 있다. 아울러 폭소노미 태그의 단점인 단어의 모호성 및 다의성과 관련하여 유사한 개념의 다양한 형태를 가지는 형용사 태그와 명사 태그의 단·복수형, 한자와 한글 표현 문제와 같이 태그의 공유와 재사용을 어렵게 만드는 형태가 많았는데 이는 태그 검색을 저해하는 요인이 된다.

4.3 검색에서의 활용 가능성

본 절에서는 디지털 문화유산에 대한 이용자생성 태그와 박물관 메타데이터 내용을 서로 비교해보고 이용자생성 태그 중 박물관 메타데이터의 내용과 상이한 부분이 박물관 디지털 문화유산 검색에 도움을 줄 수 있는지, 그 활용가능성에 대해 살펴보았다.

Trant(2006)의 미술관에서의 폭소노미 적용 가능성에 대한 연구와 이성숙(2009)의 도서관 폭소노미 태그의 패턴 분석에 관한 연구에 따르면 이용자집단이 생성한 태그가 기관에서 제공하는 기존의 전문가 메타데이터와는 일치하지 않지만, 유용한 내용을 이용자가 태깅한 경우, 이러한 태그는 기존 분류 체계를 보완하고 검색을 향상시킬 가능성이 있는 것으로 평가하였다. 따라서 본 연구에서도 분석 대상인 이용자생성 태그와 박물관 메타데이터와의 비교를 통해 박물관 디지털 문화유산 검색에서의 태그의 활용가능성을 살펴보았다.

'디지털 문화유산에 대한 이용자생성 태그가 박물관 메타데이터의 내용과 다른가?', '만약 다르다면 박물관 메타데이터의 내용과 상이한 이용자생성 태그는 박물관 디지털 문화유산 검색에서 활용될 가능성이 있다고 봐도 좋은 것인가?'와 같은 문제들에 관한 논의를 위해 수집된 이용자생성 태그 중 '유물상세정보'에는 없는 단어 들 중에 많은 이용자들이 중복 사용된 것이라면 유물 검색에서 활용될 가능성이 있는 태그로 보고 이들을 분석해 보았다. '유물상세정보'의 내용과 일치하지 않는 태그들 중에서도 앞서 분석하였던 기능별 태그 분류의 '자기참고 태그'나 '오탈자 태그' 범주에 속하는 태그들을 제외하였으며(Golder & Huberman 2006), 또한 최소한 두 명 이상에게 중복 사용되지 않은 태그는 검색에서 활용될 가능성이 적은 것으로 보고 제외하였다. 검색에서의 태그 활용 가능성에 대해 논의해보기 위해 20개의 유물에 대한 이용자생성 태그를 <표 8>과 같은 방식으로

분석하였다. <표 8> 상단에는 유물 이미지와 박물관에서 제공하고 있는 ‘유물상세정보’가 나타나 있으며, 아래의 ‘이용자 생성 태그’란에는 조사에서 수집된 이용자생성 태그들 중 상단의 유물상세정보와 중복되지 않는 것들을 기재하였다. 전체 유물 20개에 대한 종합 결과는 아래의 <표 9>에 제시된 것과 같다. 표에서 ‘사용불가능’한 태그는 ‘메타데이터와 일치하지 않는 태그’ 중에서 ‘자기참고 태그’나 ‘오탈자 태그’들과 ‘사용빈도가 2번 미만인 태그’ 범주에 해당되는 태그 수를 나타내고 있다.

분석 기준에 따라 전체 태그를 박물관 측 메타데이터 내용과 일치하는 태그와 일치하지 않는 태그로 구분하여, 일치하지 않는 태그 중 ‘자기참고 태그’나 ‘오탈자

태그’ 범주에 속하지 않으며 2명 이상의 사용자에 의해 중복 태깅된 단어들을 검색에 활용될 가능성이 높은 유용한 태그로 간주하여 각 유물 별로 이러한 조건을 만족하는 태그들에 대한 분석을 실시해 보았다.

그 결과 이용자생성 태그 중 기존 박물관 메타데이터에는 존재하지 않는 태그이면서도 많은 이용자들이 의해 중복 사용된 태그가 의외로 많은 것으로 나타났다. 하지만 이러한 조건만 만족한다고 해서 실제 유물정보 검색에서 활용될 가능성이 높은 유용한 태그라고 선별리 판단하기는 힘들다. 향후 실제 시스템 사용자들의 질의 분석 내용과의 비교 분석이 이루어진다면 보다 신뢰로운 결과의 도출이 가능할 것으로 본다.

<표 8> 이용자생성 태그와 유물상세정보 비교 분석 예



이뮤지엄의 유물정보검색의 유물상세정보	
유물명칭	강희안필고사관수도(姜希顔筆高士觀水圖)
국적/시대	한국(韓國) / 조선(朝鮮)
크기	세로: 37.6 cm /가로: 31.3 cm /두께: 2.6 cm
재질	지(紙)
작자/필자	강희안(姜希顔)
용도/기능	문화예술(文化藝術)/서화(書畫)/회화(繪畫)/일반회화(一般繪畫)
참고문헌	명품도감, 국립중앙박물관 편, 삼화출판사, 1985, 도154
소장처	국립1(國立1) / 중앙(中央)
유물번호	본관(本館) 2504

해설

강희안(姜希顔, 1417-1464)은 조선 초기의 대표적인 문인 화가로 자는 경우(景愚), 호는 인재(仁齋) 또는 청천자(靑川子)이다. 시·서·화(詩書畫)에 모두 뛰어나 삼절(三絶)로 일컬어졌으며, 23세에 문과에 급제하여 집현전 부제학, 호조 참의 등의 벼슬을 지냈다. 그림 그리는 일을 천기(賤技)로 생각하는 것은 당시의 사회적 통념이었는데, 강희안도 자신의 서화(書畫)가 후세에 전해지는 것을 굴욕으로 여겼다는 기록이 있다. 따라서 그의 작품은 전래되는 것이 드물다. 그러나 문헌상에 나타나는 작품들로는 산수 외에 매화, 대나무, 화훼(花卉), 초충(草蟲) 등 수십 점에 이른다. 바위에 기대어 흐르는 물을 조용히 바라보며 명상에 잠긴 고사(高士)를 그린 것이다. <개자원화전(芥子園畫傳)>의 <고운공편심(高雲共片心)>과 사뭇 닮고 있으나 허리를 더욱 굽혀 느슨한 모습이다. 남송(南宋) 원체(院體)에서 유래된 잔산잉수(殘山剩水)의 정형을 보여주며, 당시 상류 사회 사인화(士人畫)의 기법과 취향을 시사하는 바가 크다. 강희안은 조선 전기 화단에 있어서 사대부로서 지도적 위치에 있었다. 이 작품에 보이는 대로 조선 중기를 풍미한 절파(浙派) 화풍의 선구로서 영향력이 큰 문인화가였다.

* 유물 ID: #4, * 유물의 유형: 회화(일반회화)

이용자 생성 태그

유물정보검색의 유물상세정보와 중복되지 않는 이용자 생성 태그(오탈자, 자기참고 태그 제외)

수목화(수목화), 신선, 신선놀음, 강(강가), 호수(가), 계곡(가), 휴식, 유유자적, 무릉도원, 연못(연못가), 바위, 자연, 동양화, 먹물, 풍경, 물(물가), 나무, 노인, 낮잠, 낙관, 산수화, 안빈낙도, 선인, 대머리, 붓, 인생무상, 인물화, 화선지, 망중한, 적막감, 안식처, 풍류, 절벽, 바위 위 사람, 바위 위 신선, 바위 위 노인, 바위 위 스님, 누워있는 사람, 옆드려있는 사람, 한가*로움/롭다/함, 여유*로움/롭다/, 편안*하다/한/함, 고요*하다/한/함, 잔잔*하다/한/함, 즐리*다/운/림, 흑백의

〈표 9〉 검색에서 활용가능성이 있는 태그

ID	유물의 유형	태그의 총수 (중복 태그 제외)	메타데이터와 일치하는 태그	메타데이터와 일치하지 않는 태그	
				사용불가능	사용가능
#2	회화류	238	94	81	53
#4		213	62	85	46
#5		196	75	62	69
#7		227	67	100	60
#11		246	89	99	58
#14		281	62	111	68
#17		239	26	103	81
#3	조각류	206	31	132	43
#15		205	29	131	45
#16	공예류	193	24	125	44
#18		169	29	103	37
#8	과학 기술류	177	31	102	44
#9		175	30	106	39
#10		200	34	131	35
#13		144	23	90	31
#19		172	18	121	33
#20		141	21	91	29
#1	기록 유산류	127	22	88	57
#6		151	30	83	44
#12		132	21	88	23

5. 결론 및 제언

콘텐츠 이용자가 주체가 되어 콘텐츠 정보를 조직하는 방식인 태그시스템과 폭소노미는 이용자의 다양한 관점을 반영하여 웹상에 흩어진 정보들을 보다 효율적으로 연결해주고 사회적 네비게이션을 활발하게 해줌으로써, 기존의 정보조직과 검색이 가지는 문제점을 보완할 수 있는 웹 2.0 기술로 주목받고 있다. 물론 아직까지 정확성과 신뢰성 면에서 여러 가지가 약점으로 지적되고 있지만, 웹상의 특정 주제 및 특정 형태의 정보 접근에 부분적으로 효과적인 것으로 평가되고 있다.

박물관 유물 정보 검색서비스에 폭소노미를 이용한 태그시스템이 도입되면 태그는 박물관 자체에서 제공하는 전문가 메타데이터와는 다른 측면에서 유물 정보를 설명해줌으로써 기존 메타데이터를 보완해 주는 역할과 함께 이용자들이 박물관 디지털 문화유산을 보다 쉽고 다양한 방법으로 검색할 수 있도록 해 줄 것이다. 또한 박물관 측면에서는 이용자 생성 태그를 통해 제공 서비스에 대한 이용자의 요구사항 및 소장 자료에 대한 이용자의 인식수준을 파악할 수 있어서, 이를 바탕으로 한

서비스 개선과 박물관 이용자 교육에 기초자료로 활용할 수 있다는 이점이 있다.

이에 본 연구는 박물관 디지털 문화유산 검색에서의 태그와 폭소노미 적용에 앞서 그 활용가능성을 검토해 보기위해 디지털 문화유산에 부여된 태그들의 양상과 특성을 분석해 보았으며, 그 결과는 다음과 같다.

첫째, 기존의 태그 분석 연구들에서 제시된 태그의 전형적인 패턴들 중 광범위하고 전반적인 내용을 범주화하는 태그가 빈번하게 사용되며, 매우 빈번하게 사용되는 소수의 태그들이 있으며 그 외 대다수 태그는 매우 드물게 사용된다는 특성 등이 본 연구의 박물관 디지털 문화유산 태그에서도 비슷한 양상으로 나타났다.

둘째, 디지털 문화유산 태그의 일반적 패턴과 특성분석을 통해 얻어낸 결과들이 물리적인 유물의 유형에 따라 서로 다른 것으로 나타났다. 이와 같은 연구결과를 토대로, 박물관 메타데이터가 미처 기술하지 못한 박물관 컬렉션의 유형별 특성정보를 이용자생성 태그에서 잘 반영할 수 있도록 해준다면 박물관 디지털 문화유산 검색에서 유용한 역할을 할 수 있을 것으로 본다.

마지막으로, 이용자생성 태그와 박물관 메타데이터와

의 비교를 통해 박물관 디지털 문화유산 검색에서의 태그의 활용가능성을 검토해 보았다. 수집된 전체 태그 중 박물관 메타데이터 내용과 일치하지 않으면서, '자기 참고 태그'나 '오타자 태그' 범주에 속하지 않는 내용이 여러 사용자들에 의해 태그된 것이라면 이들을 검색에서 활용될 가능성이 있는 것으로 분석되었다.

본 연구에서는 조사 대상이 실제 태깅 시스템을 도입한 사례를 대상으로 이루어진 것이 아니라 임의로 구성된 이용자집단을 대상으로 하여 한시적으로 수집된 태그를 분석하였다. 때문에 본 분석에 사용된 태그 데이터로는 박물관 유물 태그의 형태적 특성 및 일반적인 태그 패턴 위주의 분석만이 가능하였으며, 서비스 이용자들의 자발적 태깅 참여도와 실제적 태깅 동기 등을 분석하기에는 한계가 있었다. 하지만, 특정 기관에 적합한 태깅 서비스의 도입을 위해서는 해당 기관 및 제공 콘텐츠의 특성을 고려한 태그 분석이 반드시 필요함으로 박물관 문화유산포털의 디지털 문화유산 검색에 태깅 시스템과 폭소노미 도입의 필요성을 언급하였고, 박물관 디지털 문화유산에 대한 태그가 갖는 몇몇 특성을 분석해 보았는데 본 연구의 의의가 있다.

박물관의 디지털 문화유산 태깅에 대한 보다 면밀한 분석을 위해서는 실제 태깅 시스템을 도입하고 있는 박물관 문화유산포털의 보다 다양한 형태의 디지털 문화유산을 대상으로, 광범위한 이용자에 의해, 장기간에 걸쳐 부여된 실제 태그의 수집이 중요하다. 따라서 태깅 시스템을 도입한 문화재검색서비스에서 수집된 실제 태그 데이터와 사용자들의 태깅 행태를 분석하는 후속연구가 필요하다.

【참고 문헌】

- 김녹환, 김태수. 2002. XML을 활용한 전자박물관시스템 구현에 관한 연구. 『지식처리연구』, 3(1): 41-62.
- 김운용, 박석규. 2007. 웹2.0의 참여형 아키텍처 환경에서 그래픽 기반 포크소노미 태그 연관 검색의 설계 및 구현. 『한국인터넷정보학회』, 8(51): 1-10.
- 문화재청. [cited in 2009.07.21].
<http://www.cha.go.kr/korea/heritage/knowledge/notion_01.jsp?mc=KS_01_01_01>.
- 박정희. 2008. 문화재보호의 법리에 관한 연구. 『법학연구』, 31: 77-96.
- 박찬표, 반영환, 정지홍. 2007. 웹 이미지의 협력적 폭소노미에 관한연구. 『한국디자인학회 봄 학술발표대회 논문집』, 5: 58-59.
- 배정현. 2007. 『디지털 문화유산 콘텐츠 공동 활용을 위한 협력 방안 연구』. 연세대학교 대학원 석사학위 논문.
- 배주희, 이경원. 2009. 소셜태깅에서 관심사로 바라본 태그특징. HCI 학술대회.
- 유동환. 2008. 한국 전통 문화유산 콘텐츠개발 현황과 과제, 한국국학진흥원, 『국학연구』, 12: 6:5-49.
- 유시내. 2007. 『기록검색서비스 개선을 위한 폭소노미 도입에 관한 연구』. 명지대학교기록과대학원 석사학위논문.
- 이성숙. 2008. 대학도서관 폭소노미 태그의 형태적 특성에 관한 연구. 『한국문헌정보학회지』, 42(4): 463-480.
- 이성숙. 2009. 국내 도서관 폭소노미 태그의 일반적 패턴 연구. 『한국비블리아학회지』, 20(1): 137-150.
- 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.
- 장호수. 2006. 문화재 활용론. 인문콘텐츠학회, 『인문콘텐츠』, 7: 155-173.
- 조재인. 2008. 도서관 정보 수요자를 위한 소셜 네트워크 서비스 도입에 관한 연구 『한국도서관·정보학회지』, 39(2): 169-186
- 한문희. 2004. 국가문화유산 종합정보시스템 구축 사업: 문화유산 콘텐츠, '축적'과 '활용'의 전략적 코어에 대한 인식. 인문콘텐츠학회. 『인문콘텐츠』, 3: 367-377.
- Angus, E., M. Thelwall, and D. Stuart. 2008. General patterns of tag usage among university groups in Flickr. *Online Information Review*, 32(1): 89-101.
- Golder, S. A., & B. A. Huberman. 2006. Usage patterns of collaborative tagging systems. *Journal of Information Science*32(2): 198-208.
- Guy, M. and E. Tonkin. 2006. Folksonomies: tidying up tags? *D-Lib Magazine*, 12(1).
<<http://www.dlib.org/dlib/january06/guy/01guy.html>>
- Holt, B. L Hartwick, 1994. Visual image retrieval for applications in art and art history. *Proc. SPIE*,

- Vol.2185(70).
- Shatford-Layne, S. 1994. Some issues in the indexing of images. *Journal of the American Society for Information Science*45(8): 583-8.
- Trant, Jennifer. 2006. Exploring the potential for social tagging and folksonomy in art museums: proof of concept. *A paper for the new review of hypermedia and multimedia*.

