

폭소노미 연구 문헌에 대한 자아 중심 주제 인용 분석*

Ego-centered Topic Citation Analysis on Folksonomy Research Documents

이재윤 (Jae Yun Lee)**

초 록

이 연구에서는 White가 제안한 자아 중심 인용 분석을 응용하여 연구 주제에 대한 다층적인 분석을 가능하게 해주는 자아 중심 주제 인용 분석 기법을 제안하였다. 시험적으로 폭소노미에 대한 연구문헌을 Web of Science 데이터베이스로부터 검색한 후 이에 대한 주제 인용 분석을 수행해보았다. 폭소노미 주제에 대한 자아 중심 인용 분석은 자아 문헌 집단 분석, 주제 인용 정체성 분석, 주제 인용 이미지 분석의 세 단계로 나뉘어 수행되었다. 분석 결과 이 연구에서 제안된 자아 중심 주제 인용 분석을 통해서 폭소노미 연구의 내부 지적 구조와 외부 지적 구조를 함께 파악하는 것이 가능함이 확인되었다.

ABSTRACT

This research aims to present the ego-centered topic citation analysis, which is a new application of White's ego-centered citation analysis, for analyzing multilayered knowledge structure of a subject domain. An experimental topic citation analysis was carried out on the folksonomy research documents retrieved from Web of Science. Ego-centered topic citation analyses on folksonomy research domain were conducted in three stages: ego-documents set analysis, topic citation identity analysis, and topic citation image analysis. The results showed that the ego-centered topic citation analysis suggested in this study was successfully performed to illustrate the inner and the outer knowledge structures of folksonomy research domain.

키워드: 자아중심 인용분석, 주제 인용분석, 연구영역 분석, 지적구조 분석, 폭소노미, 인용정체성, 인용이미지

ego-centered citation analysis, topic citation analysis, research domain analysis, intellectual structure analysis, folksonomy, citation identity, citation image

* 본 연구는 2012학년도 경기대학교 학술연구비(일반연구과제) 지원에 의하여 수행되었음.

본 연구의 일부 내용은 제19회 한국정보관리학회 학술대회에서 발표된 바 있음.

** 경기대학교 문헌정보학과 부교수(memexlee@kgu.ac.kr)

■ 논문접수일자: 2012년 12월 8일 ■ 최초심사일자: 2012년 12월 8일 ■ 게재확정일자: 2012년 12월 24일
■ 정보관리학회지, 29(4), 295-312, 2012. [<http://dx.doi.org/10.3743/KOSIM.2012.29.4.295>]

1. 서론

학문영역의 지적 구조와 연구 동향에 대한 계량서지적 분석은 Kessler(1963)의 서지결합 분석부터 따졌을 때 50년의 역사를 가지고 있다. 지적 구조 분석을 위한 주요 기법인 서지결합과 동시인용분석(Small, 1973), 그리고 저자동시인용분석(White & Griffith, 1981)은 모두 처음 발표된지 최소 30년 이상이 흘렀지만 Lee와 Choi(2011)의 분석에 따르면 21세기에 들어서도 연구영역 분석(research domain analysis)에서 활발히 사용되고 있다. 이런 기법들을 적용하는 대부분의 지적 구조 분석 연구에서는 일정한 범위의 저자 집합이나 문헌 집합을 대상으로 인용을 통해 드러난 관계를 분석한다.

상대적으로 최근이라고 할 수 있는 2000년에 개발된 자아 중심 인용 분석(ego-centered citation analysis)은 White(2000)가 제안한 기법으로서, 저자 집단이나 문헌 집단을 미리 확보하고 집단의 내부를 분석하는 기존 방식과 달리 특정 연구자를 중심으로 주변 연구자를 여러 관점으로 파악하여 개인 연구자의 연구활동에 대한 다면적인 분석을 수행하는 방법이다. White(2000)는 사회 네트워크 분석 분야에서 정립된 자아 중심 네트워크 분석(ego-centered network analysis) 기법을 응용하여 개인 연구자를 자아(ego)로 설정하고 자아의 인용 정체성, 인용 이미지, 인용 이미지 구축자, 공저자까지 네 측면의 분석이 가능하다고 제안하였다. 각각의 기본 개념은 다음과 같다(이재윤, 2012).

- 인용 정체성: 연구자 자신(ego)에 의해서 인용되는 저자 집합(alters)

- 인용 이미지: 연구자와 동시인용된 저자 집합
- 인용 이미지 구축자: 연구자를 인용하는 저자 집합
- 공저자: 연구자와 공저한 저자 집합

이중에서 개인 연구자의 연구에 대해서 분석할 때 가장 직접적으로 정보를 제공하는 것은 인용 정체성과 인용 이미지이다. 연구자 A의 인용 정체성을 구성하는 저자들은 A의 논문에서 A 본인이 직접 인용하기로 선택한 저자들이며, 연구자 A의 인용 이미지를 구성하는 저자들은 다른 사람들이 A를 인용한 논문에서 함께 인용하기로 선택한 저자들이다. 인용 정체성은 연구자 본인의 생각 속에 있는 자신의 모습이고 인용 이미지는 타인의 눈에 비친 연구자의 모습이라고 할 수 있다(이재윤, 2012).

자아 중심 인용 분석 기법은 연구자 개인을 대상으로 적용하도록 제안되었지만, 저자처럼 논문 집합의 범위를 정할 수 있는 대상인 저널에 대해서 적용한 연구도 다수 발표되었다(Bonnevie-Nebelong, 2006; Nebelong-Bonnevie & Frandsen, 2006; Tsay, 2011; Yanping, 2011). 연구자에 대해서 분석할 때에는 해당 연구자의 저작에서 인용되었거나 해당 연구자의 저작을 인용한 저자를 분석하듯이, 저널에 대해서 분석할 때에는 해당 저널의 논문에서 인용되었거나 해당 저널의 논문을 인용한 저널을 분석하는 방법이 사용되었다.

자아 중심 인용 분석 기법처럼 본래 개인 연구자를 대상으로 개발된 기법을 개인이 아닌 다른 대상에 적용한 사례로는 h-지수(Hirsch, 2005)를 들 수 있다. 인용을 통해 개인 연구자

를 평가하는 지표로 제안된 h-지수는 저널 영향력 평가(김판준, 이재운, 2010; Braun, Glänzel, & Schubert, 2005)나 연구 주제의 영향력 평가(Banks, 2006)에 응용되었다. 이는 h-지수의 산출 방식이 어느 정도 규모의 문헌 집합으로 대표되는 대상이라면 무엇에나 적용할 수 있음을 의미한다. 자아 중심 인용 분석 기법은 개인 이외에 저널에만 응용되었지만, h-지수는 저널 뿐만 아니라 연구 주제를 대상으로도 적용된 바 있다. 개인을 대상으로 자아 중심 인용 분석을 수행하는 것이 연구자에 대한 다면적인 분석을 가능하게 한 점을 감안하면, 자아 중심 인용 분석을 연구 주제에 응용하는 것을 고려해볼 여지가 있다.

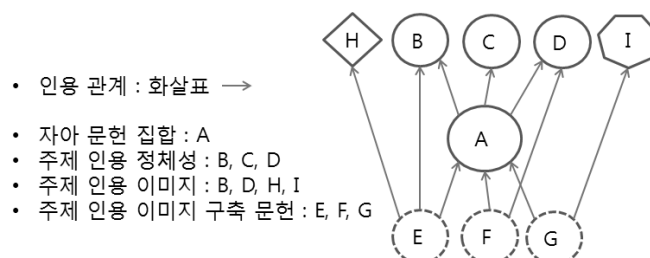
이 연구에서는 특정 저자나 저널을 다루는 것과 유사하게 특정한 연구 주제를 자아로 설정하고 해당 주제와 관련된 문헌 집합을 대상으로 자아 중심 인용 분석을 수행하는 자아 중심 주제 인용 분석 기법을 제안하고, 폭소노미 연구 문헌들에 대해서 적용해보았다. 이 연구에서 제안하는 자아 중심 주제 인용 분석 기법은 특정 주제 분야의 연구 동향에 대한 다면적인 분석을 가능하게 할 것으로 기대된다.

2. 자아 중심 주제 인용 분석의 개념

자아 중심 주제 인용 분석에서는 선행연구에서 자아(ego)를 연구자 개인이나 저널로 설정했던 것과 달리 특정 주제 연구문헌 집합으로 설정한다. 사실상 연구자 개인이나 저널도 단일 논문이 아닌 여러 논문의 집합이므로 ego의 설정에 있어서는 크게 다르지 않다. 일단, 주요 문헌 집합을 자아로 설정하였을 때 다음과 같이 기본 개념을 설정할 수 있다.

- 자아 문헌 집합(ego documents): 특정 주제와 관련된 문헌 집합
- 주제 인용 정체성(topic citation identity): 주제 문헌 집합에서 인용된 주요 문헌 집합
- 주제 인용 이미지(topic citation image): 주제 문헌 집합과 동시인용된 주요 문헌 집합
- 주제 인용 이미지 구축자(topic citation image maker): 주제 문헌 집합을 인용한 문헌 집합

〈그림 1〉의 예를 보면, 중앙의 A가 특정 주제 분야의 문헌들로 구성된 자아 문헌 집합일



〈그림 1〉 자아 중심 주제 인용 분석의 구성

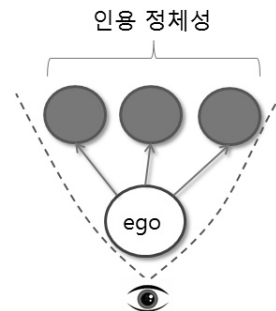
경우에 이들로부터 인용된 문헌 B, C, D가 주제 인용 정체성에 해당한다. 자아 문헌 집합 A에 속한 문헌과 동시인용된 문헌 B, D, I, H는 주제 인용 이미지에 해당하고, 자아 문헌 집합 A에 속한 문헌을 인용한 문헌 E, F, G는 인용 이미지 구축자에 해당한다.

자아 문헌 집합 자체에 대한 분석은 특정 주제 영역의 문헌들을 수집하여 인용분석을 수행하는 전통적인 방법과 유사하다. 이를 통해서 주제 영역 내에서 어떤 연구들이 이루어지는지 미시적으로 파악할 수 있다. 자아 문헌 집합의 범위는 특정 주제에 대한 문헌을 망라적으로 검색하는 검색식을 설정하고 탐색을 수행함으로써 결정할 수 있다. 자아 중심 주제 인용 분석에서는 자아보다 타자를 의미있는 단위로 식별하는 것이 어렵다. 분석 대상이 연구자나 저널인 경우에는 서지사항에 명시되므로 자아를 인용하거나 자아가 인용하는 타자를 식별하기가 용이하지만, 주제를 단위로 하는 분석에서는 타자, 즉 자아가 아닌 다른 주제를 기계적으로 식별하기는 어렵다. 자아 중심 주제 인용 분석에서 타자인 문헌 집합은 자아를 인용하거나 자아가 인용한 문헌들 중에서 인용 빈도가 높은 문헌들을 다양한 방법으로 군집화해서 해석하기로 한다.

주제 인용 정체성을 파악할 때에는 주제 문헌 집합 내에서 빈번히 인용된 문헌들을 동시인용 분석(Small, 1973)하여 결과를 해석한다. 주제 인용 이미지를 파악할 때에는 주제 문헌 집합을 인용한 문헌 집합 내에서 자주 인용된 문헌들을 동시인용 분석하여 결과를 해석한다.

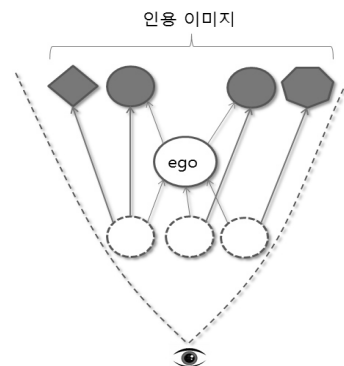
이와 같이 연구 주제를 대상으로 한 인용 정체성과 인용 이미지는, 연구 주제에 대한 좁은 시각의 접근과 넓은 시각의 접근으로 비유할 수

있다. 주제 인용 정체성은 특정 연구 주제 문헌들에서 인용된 문헌 집합이므로 해당 주제의 연구 수행에 필수적인 핵심 문헌 집합을 드러내 준다. 이는 <그림 2>와 같이 해당 주제 가까에서 바라보는 근거리 관찰이라고 할 수 있다.



<그림 2> 주제 인용 정체성의 개념도

이와 달리 인용 이미지는 특정 연구 주제 문헌들을 인용한 문헌들이 함께 인용한 문헌 집합, 즉 자아 문헌과 동시인용된 문헌의 집합이므로 해당 연구 주제에 한정되지 않고 관련된 연구 문헌들을 폭넓게 포괄하게 된다. 따라서 주제 인용 이미지 분석은 <그림 3>과 같이 해당 주제에서 어느 정도 떨어져서 살펴보는 원거리 시점의 관찰이라고 할 수 있다.



<그림 3> 주제 인용 이미지의 개념도

〈그림 2〉와 〈그림 3〉에서 ego(자아)는 특정 주제의 문헌 집합을 나타내고, 다른 도형은 문헌, 화살표는 인용관계를 의미한다. 〈그림 2〉의 주제 인용 정제성에서 최종 분석 대상 문헌 집합은 특정 주제 문헌 집합에서 인용된 문헌들(원으로 표시)로만 한정된다. 이와 달리 〈그림 3〉의 주제 인용 이미지에서는 특정 주제 문헌 집합에서 직접 인용되지는 않았으나, 주제 인용 이미지 구축자로부터 특정 주제 문헌 집합과 동시 인용된 문헌들(다각형으로 표시)도 분석에 포함된다. 이처럼 주제 인용 이미지 분석에서는 특정 주제를 연구할 때 함께 인용되는 다른 주제의 문헌들도 살펴볼 수 있다.

3. 폭소노미 연구 문헌에 대한 분석

3.1 폭소노미 연구 문헌 검색

자아 중심 주제 인용 분석의 시험적인 적용 대상으로는 문헌정보학을 비롯하여 물리학, 사회학, 컴퓨터과학 등의 여러 분야에서 탐구되고 있는 폭소노미를 선정하였다. 2011년 11월 12일, Web of Science 데이터베이스에서 topic 필드 검색창을 대상으로 〈그림 4〉의 1번 검색식과 같이 Topic=(Folksonomy OR Folksonomies)로 검색한 결과 139건, 2번 검색식에서 Topic= (“Social Bookmark” OR “Social Bookmarks” OR “Social Bookmarking”)으로 검색한 결과

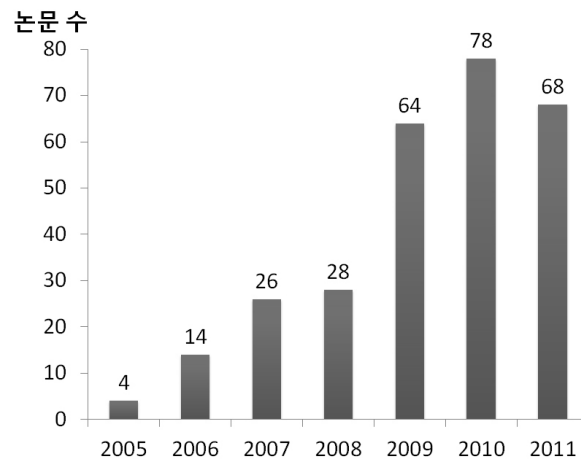
# 1	139	Topic=(Folksonomy OR Folksonomies) Databases=SCI-EXPANDED, SSCI, A&HCI Timespan=All Years Lemmatization=On
# 2	89	Topic=(“Social Bookmark” OR “Social Bookmarks” OR “Social Bookmarking”) Databases=SCI-EXPANDED, SSCI, A&HCI Timespan=All Years Lemmatization=On
# 3	147	Topic=(“collaborative tagging” OR “user tagging” OR “social tagging” OR “Social tag” OR “Social tags”) Databases=SCI-EXPANDED, SSCI, A&HCI Timespan=All Years Lemmatization=On
# 4	300	#3 OR #2 OR #1 Databases=SCI-EXPANDED, SSCI, A&HCI Timespan=All Years Lemmatization=On
	283	Refined by: Document Type=(ARTICLE OR PROCEEDINGS PAPER OR MEETING ABSTRACT OR REVIEW)
	282	1997 exclude

〈그림 4〉 WoS에서 폭소노미 관련 문헌을 검색한 검색식과 검색 건수

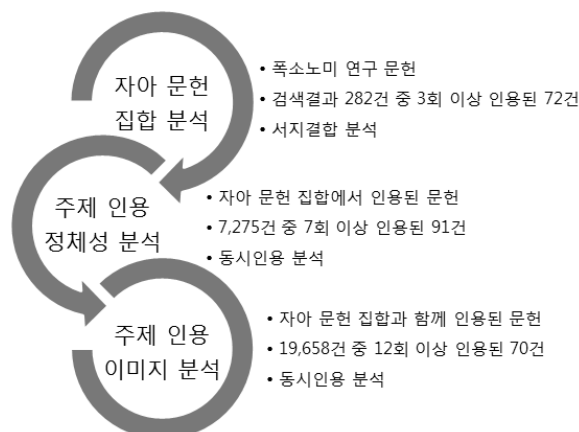
89건, 3번 검색식에서 Topic = ("collaborative tagging" OR "user tagging" OR "social tagging" OR "Social tag" OR "Social tags")로 검색한 결과가 147건이었으며, 세 차례 검색 결과의 중복 논문을 제외하고 통합한 결과 300건이 되었다. 이 중에서 문헌 유형이 학술지나 학술대회 논문, 리뷰 논문인 경우로 제한하여 283건을 획득하였다. 283건 중에서 1997년에 발표된 문헌

이 1건 있어 확인해본 결과 2000년대에 제안된 폭소노미의 개념과 전혀 다른 문헌이어서 이를 제외한 282건이 최종 폭소노미 주제 자아 문헌 집합으로 선정되었다. 282건의 연도별 논문 건수는 <그림 5>와 같다.

폭소노미 주제 자아 문헌 집합을 확정한 이후에는 <그림 6>과 같이 자아 문헌 집합 분석, 주제 인용 정체성 분석, 주제 인용 이미지 분석



<그림 5> WoS의 연도별 폭소노미 관련 논문 수
(2011년 11월 12일 검색)



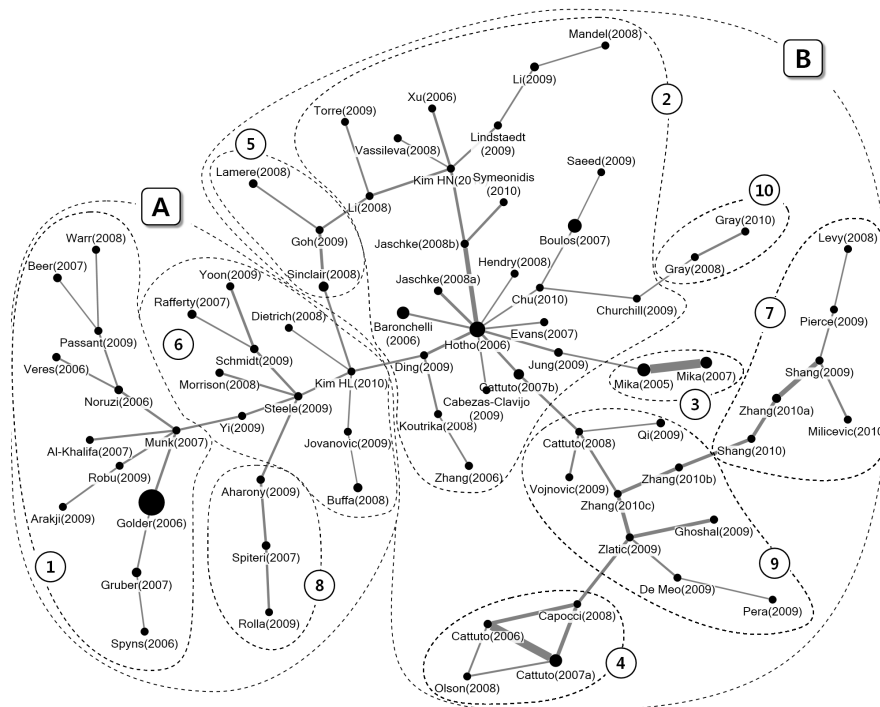
<그림 6> 폭소노미에 대한 자아 중심 인용 분석 절차

의 순서로 진행된다. 자아 문헌 집합에 대한 세부 주제 분석은 각 문헌의 인용 빈도가 낮기 때문에 서지결합(Kessler, 1963)을 이용해서 수행한다. 자아 문헌 집합에서 인용된 문헌들로 구성되는 주제 인용 정체성은 인용 빈도 상위 문헌들 사이의 동시인용(Small, 1973)을 이용해서 분석한다. 자아 문헌 집합의 문헌들과 동시인용된 주제 인용 이미지를 분석할 때에도 인용 빈도 상위 문헌들 사이의 동시인용을 단서로 삼는다.

3.2 폭소노미 주제 자아 문헌 집합 분석

폭소노미 주제의 자아 문헌 집합에 해당하는
검색된 282건 중에서 인용빈도가 3회 이상이고

타 문헌과 공유하는 참고문헌이 있는 76건을 폭소노미 주제 주요 문헌으로 선정하고 이 문헌들의 서지결합(Kessler, 1963) 네트워크를 분석해 보았다. 일단 두 문헌 사이의 서지결합도를 측정하기 위해 일치하는 참고문헌의 수를 분자로 하고 두 문헌 각각의 참고문헌 수를 곱한 것의 제곱근을 분모로 하는 코사인 유사도 공식으로 산출하였다. 모든 문헌 사이의 서지결합도를 산출하여 서지결합도 행렬을 구축하고 이를 입력 자료로 하여 WNET 프로그램(이재윤, 2011)으로 <그림 7>과 같은 패스파인더 네트워크를 생성하였다. 또한 병렬 최근접 이웃 클러스터링 PNNC(이재윤, 2006)을 수행하여 10개의 소군집과 2개의 대군집을 도출하였다. 각 군집의 주요 문헌은 <표 1>과 같다.



〈그림 7〉 폭소노미 주제 주요 문헌의 서지결합 패스파인더 네트워크

〈표 1〉 폭소노미 주제 문헌 집합의 군집별 주요 문헌

PNNC군집		제1저자(연도)	인용 빈도*	제목
C2	C10			
A	1	Golder(2006)	345	Usage patterns of collaborative tagging systems
		Gruber(2007)	23	Ontology of folksonomy: A mash-up of apples and oranges
		Noruzi(2006)	12	Folksonomies: (Un) controlled vocabulary?
		Beer(2007)	12	Sociology and, of and in Web 2.0: Some initial considerations
	6	Rafferty(2007)	17	Flickr and democratic indexing: Dialogic approaches to indexing
		Buffa(2008)	17	SweetWiki: A semantic wiki
		Schmidt(2009)	10	Collective indexing of emotions in images: A study in emotional information retrieval
		Morrison(2008)	8	Tagging of and searching: Search retrieval effectiveness folksonomies on the World Wide Web
	8	Spiteri(2007)	12	The structure and form of folksonomy tags: The road to the public library catalog
		Aharony(2009)	9	Librarians and information scientists in the blogosphere: An exploratory analysis
		Rolla(2009)	3	User tags versus subject headings: Can user-supplied data improve subject access to library collections?
B	2,3	Hotho(2006)	114	Information retrieval in folksonomies: Search and ranking
		Boulos(2007)	82	The emerging Web 2.0 social software: an enabling suite of sociable technologies in health and health care education
		Mika(2005)	70	Ontologies are us: A unified model of social networks and semantics
		Baronchelli(2006)	63	Sharp transition towards shared vocabularies in multi-agent systems
	4	Cattuto(2007a)	66	Semiotic dynamics and collaborative tagging
		Cattuto(2006)	9	Semiotic dynamics in online social communities
		Olson(2008)	4	Syntagmatic relationships and indexing consistency on a larger scale
		Capocci(2008)	3	Folksonomies and clustering in the collaborative system CiteULike
	5	Sinclair(2008)	28	The folksonomy tag cloud: when is it useful?
		Lamere(2008)	12	Social tagging and music information retrieval
		Goh(2009)	5	Resource discovery through social tagging: a classification and content analytic approach
	7	Zhang(2010a)	17	Personalized recommendation via integrated diffusion on user-item-tag tripartite graphs
		Shang(2009)	7	Diffusion-based recommendation in collaborative tagging systems
		Shang(2010)	6	Collaborative filtering with diffusion-based similarity on tripartite graphs
		Levy(2008)	4	Learning latent semantic models for music from social tags
	9	Qi(2009)	12	Web page classification: features and algorithms
		Ghoshal(2009)	7	Random hypergraphs and their applications
		Zlatic(2009)	6	Hypergraph topological quantities for tagged social networks
		Zhang(2010b)	5	Solving the cold-start problem in recommender systems with social tags
	10	Gray(2010)	6	Students as Web 2.0 authors: Implications for assessment design and conduct
		Gray(2008)	4	Web 2.0 authorship: Issues of referencing and citation for academic integrity

* 인용빈도는 Web of Science 전체에서 인용된 횟수임

자아 문헌 집합의 대군집 A는 군집 1, 6, 8로 구성된다. 군집 1은 Golder(2006)-“Usage patterns of collaborative tagging systems”이 대표문헌으로서 폭소노미의 개념과 이용이 주제이다. 군집 6은 Rafferty(2007)-“Flickr and democratic indexing: Dialogic approaches to indexing”이 대표문헌으로서 색인으로서의 폭소노미에 대한 군집이다. 군집 8은 Spiteri(2007)-“The structure and form of folksonomy tags: The road to the public library catalog”이 대표문헌으로서 도서관 입장에서 폭소노미를 바라보는 문헌들이 포함되어 있다. 이와 같이 대군집 A는 폭소노미 사용, 색인, 도서관 등과 같이 폭소노미 자체의 특성에 집중한 연구들이라고 할 수 있다.

자아문헌 집합의 대군집 B는 군집 2, 3, 4, 5, 7, 9, 10으로 구성된다. 군집 2는 Hotho(2006)-“Information retrieval in folksonomies: Search and ranking”가 대표문헌이며 폭소노미를 응용한 검색을 다루고 있다. Mika의 연구만 포함된 작은 군집인 군집 3은 시멘틱 웹과의 융합을 다루고 있다. 군집 4는 Cattuto(2007a)-“Semi-otic dynamics and collaborative tagging”가 대표 문헌이며 폭소노미 태그의 분석을 주제로 한다. 군집 5는 Sinclair(2008)-“The folksonomy tag cloud: when is it useful?”가 대표 문헌이며 폭소노미를 이용한 브라우징과 검색을 다루고 있다. 군집 7은 Zhang(2010a)-“Personalized recommendation via integrated diffusion on user-item-tag tripartite graphs”가 대표 문헌이며 폭소노미를 이용한 개인화 추천 시스템이 연구 주제이다. 군집 9는 Qi(2009)-“Web page classification: Features and algorithms”가 대

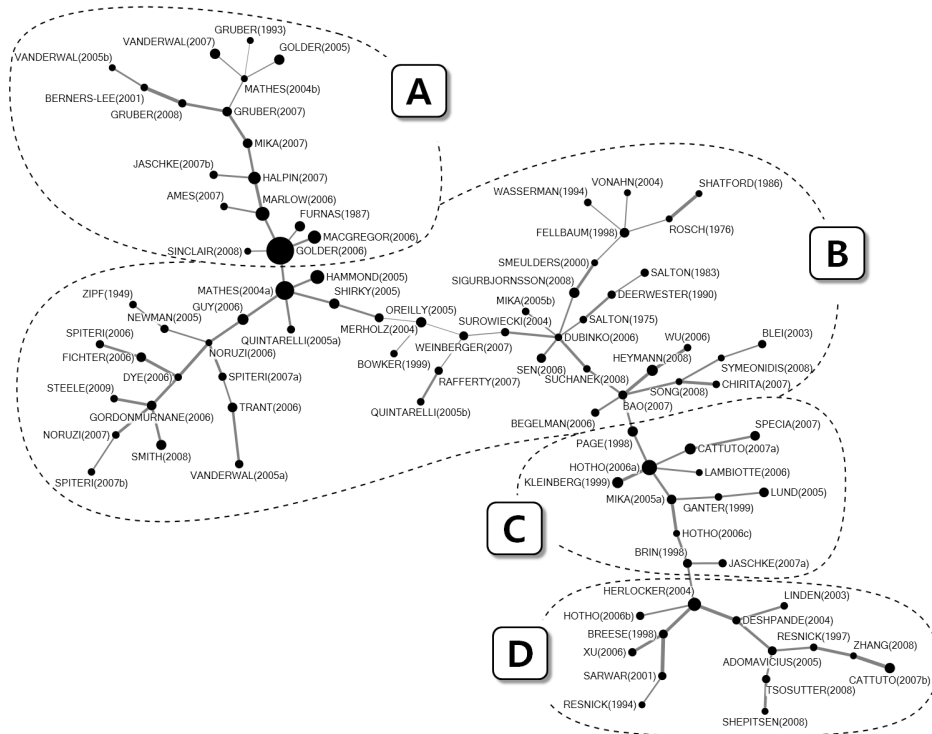
표 문헌이며 폭소노미를 이용한 분류와 검색을 다루고 있다. 군집 10은 Gray(2010)-“Students as Web 2.0 authors: Implications for assessment design and conduct”가 대표 문헌이며 폭소노미를 통한 사람들의 문헌 접근과 이용을 다루고 있다. 이와 같이 대군집 B는 폭소노미를 이용한 검색, 분류, 문헌 접근과 같은 응용에 초점을 맞춘 연구들이라고 할 수 있다.

결국 폭소노미 자아 문헌 집합에서 다루고 있는 주제는 폭소노미의 본질과 관련된 영역(대군집 A)과 응용과 관련된 영역(대군집 B)으로 크게 양분되는 것으로 나타났다.

3.3 폭소노미 주제 인용 정체성 분석

폭소노미 주제의 인용 정체성을 분석하기 위해서 폭소노미와 관련된 자아 문헌 집합 282건의 참고문헌에서 자주 등장한 문헌, 즉 폭소노미 연구에서 자주 인용된 문헌들을 파악하였다. 폭소노미 주제 자아 문헌집합 282건에서 인용된 문헌은 총 7,275건이었다. 이중에서 자아 문헌집합으로부터 인용된 횟수가 7회 이상인 문헌이 91건이었으며, 이 문헌들을 폭소노미 주제 인용 정체성 대표 문헌집합으로 하였다.

인용 정체성 대표 문헌인 91건의 구성을 살펴보기 위해서 문헌동시인용분석을 수행하였다. 이를 위해서 동시인용빈도를 코사인유사도로 정규화한 코사인유사도 행렬을 입력 데이터로 하여 WNET 프로그램으로 패스파인더 네트워크를 구축하고 PNNC 알고리즘으로 군집을 도출하였다. 분석 결과 4개 대군집과 19개 소군집이 도출되었으며, 4개 대군집을 패스파인더 네트워크와 함께 <그림 8>과 같이 표현하였다. 각



〈그림 8〉 패스파인더 네트워크로 표현한 폭소노미 연구의 인용 정체성 지도

군집별 주요 문헌은 〈표 2〉에 제시하였다.

대군집 A는 17개의 문헌으로 구성되어 있으며, VANDERWAL(2007)-“Folksonomy coinage and definition”, GOLDER(2005)-“The structure of collaborative tagging systems”와 같은 폭소노미 관련 핵심 개념을 다루는 문헌들이 포함되어 있다. 여기에는 BERNERS-LEE(2001)-“The semantic web”과 같이 시멘틱웹의 기본 개념을 다룬 문헌이 포함되어 있으므로 시멘틱웹이 일부 폭소노미 연구에서 함께 고찰되고 있음을 알려준다.

대군집 B는 48개의 문헌으로 구성되어 가장 큰 군집이며 MATHES(2004a)-“Folksonomies: Cooperative classification and communication

through shared metadata”, HAMMOND(2005)-“Social bookmarking tools (I): A general review”, HEYMANN(2008)-“Can social bookmarking improve web search?”, SMITH(2008)-“Tagging: People-powered meta-data for the social Web”와 같이 폭소노미의 활용과 관련된 논문이나 단행본이 다수 포함되어 있다. 이와 함께 FELLBAUM(1998)-“WordNet: An electronic lexical database”이나 ROSCH(1976)-“Basic objects in natural categories”와 같은 인지심리학 분야의 선행 연구와 WASSERMAN(1994)-“Social network analysis: Methods and applications”와 같은 사회네트워크분석의 선행 연구, ZIPF(1949)-

〈표 2〉 폭소노미 주제 인용 정체성의 군집별 주요 문헌

PNNC군집		제1저자(연도)	인용 빈도*	제목
C4	C19			
A	1	GOLDER(2006)	97	Usage patterns of collaborative tagging systems
		MARLOW(2006)	24	HT06, tagging paper, taxonomy, Flickr, academic article to read
		MACGREGOR(2006)	22	Collaborative tagging as a knowledge organisation and resource discovery tool
	10	VANDERWAL(2007)	13	Folksonomy coinage and definition
		GOLDER(2005)	13	The structure of collaborative tagging systems
		MIKA(2007)	12	Ontologies are us: A unified model of social networks and semantics
	18	GRUBER(2008)	8	Collective knowledge systems: Where the social web meets the semantic web
		BERNERS-LEE(2001)	7	The semantic web
B	2	MATHES(2004a)	44	Folksonomies: Cooperative classification and communication through shared metadata
		HAMMOND(2005)	24	Social bookmarking tools (I): A general review
		GUY(2006)	16	Folksonomies: Tidying up tags?
	6	HEYMANN(2008)	15	Can social bookmarking improve web search?
		BAO(2007)	10	Optimizing Web search using social annotations
	8	SIGURBJORNSSON(2008)	14	Flickr tag recommendation based on collective knowledge
		FELLBAUM(1998)	12	WordNet: An electronic lexical database
	9	OREILLY(2005)	13	What is Web 2.0
	11	SMITH(2008)	13	Tagging: People-powered meta-data for the social Web
		GORDONMURNANE(2006)	12	Social bookmarking, folksonomies, and web 2.0 tools
		FICHTER(2006)	11	Intranet applications for tagging and folksonomies
	13	TRANT(2006)	11	Exploring the potential for social tagging and folksonomy in art museums: Proof of concept
		VANDERWAL(2005a)	9	Explaining and showing broad and narrow folksonomies
	14	SEN(2006)	11	Tagging, communities, vocabulary, evolution
		WEINBERGER(2007)	10	Everything is miscellaneous: The power of the new digital disorder
		SUROWIECKI(2004)	9	The wisdom of crowds: Why the many are smarter than the few
C	16	BLEI(2003)	8	Latent dirichlet allocation
		CHIRITA(2007)	7	P-tag: Large scale automatic generation of personalized annotation tags for the web
	17	DEERWESTER(1990)	8	Indexing by latent semantic analysis
	3	HOTH(2006a)	28	Information retrieval in folksonomies: Search and ranking
		KLEINBERG(1999)	16	Authoritative sources in a hyperlinked environment
		PAGE(1998)	13	The PageRank citation ranking: Bringing order to the Web
	5	CATTUTO(2007a)	16	Network properties of folksonomies
	12	LUND(2005)	12	Social bookmarking tools (II): A case study-Connotea
D	4	HERLOCKER(2004)	21	Evaluating collaborative filtering recommender systems
		ADOMAVICIUS(2005)	10	Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions
		DESHPANDE(2004)	9	Item-based top-N recommendation algorithms
	7	CATTUTO(2007b)	14	Semiotic dynamics and collaborative tagging
	15	BRESE(1998)	10	Empirical analysis of predictive algorithms for collaborative filtering
		XU(2006)	8	Towards the semantic web: Collaborative tag suggestions
	19	TSOSUTTER(2008)	8	Tag-aware recommender systems by fusion of collaborative filtering algorithms

* 인용빈도는 폭소노미 주제 자아 문헌 집합으로부터 인용된 횟수임

“Human behavior and the principle of least effort”와 같은 지프법칙(먹함수법칙)의 선행 연구, SALTON(1983)-“Introduction to modern information retrieval”과 같은 정보검색 분야의 선행 연구가 포함되어 있다. 결국 폭소노미의 이해 및 활용과 관련하여 인지심리학, 사회네트워크분석, 먹함수법칙, 정보검색 분야의 고전적인 선행 연구들이 인용되고 있는 것으로 나타났다.

대군집 C는 12개의 문헌으로 구성되어 있으며, HOTH(2006a)-“Information retrieval in folksonomies: Search and ranking”와 같은 폭소노미를 이용한 정보검색에 관련된 문헌과 함께 PAGE(1998)-“The PageRank citation ranking: Bringing order to the Web”, KLEINBERG(1999)-“Authoritative sources in a hyperlinked environment”와 같은 네트워크 구조를 이용한 정보검색 관련 문헌이 동시인용되는 것으로 나타났다. 이는 폭소노미를 이용한 정보검색 연구에서 네트워크 정보검색 분야의 문헌들이 선행 이론으로 인용되었음을 의미한다.

대군집 D는 14개의 문헌으로 구성되어 있으며, HERLOCKER(2004)-“Evaluating collaborative filtering recommender systems”이 대표 문헌이며 협력적 필터링과 관련된 문헌, 그리고 태그를 이용한 개인화 필터링과 관련된 문헌들이 포함되어 있다. 이는 폭소노미 태그를 이용한 필터링 연구에서 협력적 필터링 분야의 문헌들이 선행 연구로 인용되었음을 의미한다.

폭소노미 관련 연구에서 인용된 문헌들로 폭소노미 주제의 인용 정체성을 분석해본 결과, 폭소노미에 대한 핵심 문헌들과 주요 선행 문헌

들이 파악되었다. 폭소노미에 대한 핵심 문헌들은 폭소노미의 개념적 이해와 활용 방안, 검색 및 추천 시스템에서의 응용 등과 같은 주제를 다루고 있었다. 폭소노미 연구에서 인용된 주요 선행 연구들은 크게 나누어보면 인지심리, 정보검색, 먹함수법칙 분야의 고전이 된 문헌들과, 시멘틱 웹이나 개인화 추천 관련 문헌과 같은 2000년대 이후의 선행 연구들이 포함된 것으로 나타났다. 결과적으로, 폭소노미 연구의 인용 정체성을 살펴본 결과 자아 문헌 집합에 대한 분석보다는 덜 포괄적이지만 어느 정도 주요 연구 분야를 파악할 수 있었으며, 폭소노미 연구에 활용된 다른 주제의 선행 연구도 일부 파악할 수 있었다.

3.4 폭소노미 주제 인용 이미지 분석

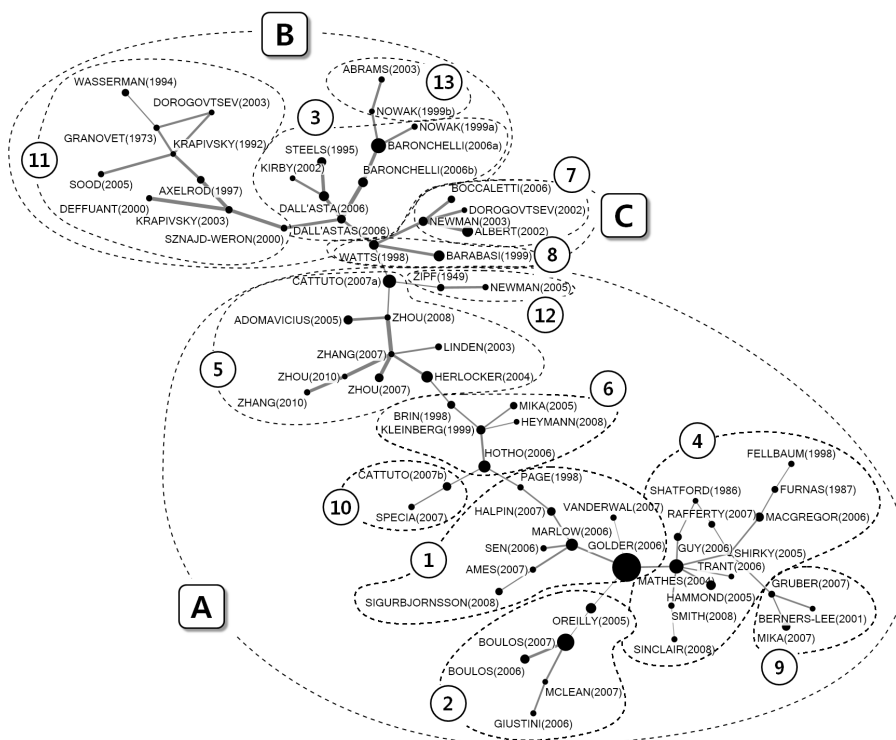
폭소노미 주제 문헌들과 동시인용된 문헌들은 폭소노미 주제 인용 이미지에 해당한다. 최초 검색된 자아 문헌 집합 282건을 인용한 572건은 인용 이미지 구축자가 되며, 이들 인용 이미지 구축자 572건으로부터 인용된 문헌 19,658건이 인용 이미지 집합이 된다. 인용 이미지 집합에 속한 문헌 중 12회 이상 인용된 70건의 문헌을 인용 이미지 대표 문헌들로 간주하였다. 대표 문헌 70건을 통해 폭소노미 주제의 인용 이미지 구조를 살펴보기 위하여 70개 문헌 사이의 동시인용분석을 수행하였다. 앞 절의 인용 정체성 대표 문헌에 대한 동시인용분석과 마찬가지로 동시인용빈도를 코사인유사도로 정규화한 코사인유사도 행렬을 입력 데이터로 하여 WNET 프로그램으로 패스파인더 네트워크를 구축하고 PNNC 알고리즘으로 군집을 도출하

였다. 분석 결과 파악된 3개 대군집과 13개 소군집을 70개 문헌의 패스파인더 네트워크와 함께 <그림 9>와 같이 표현하였다. 각 군집별 주요 문헌은 <표 3>과 같다.

<그림 9>의 인용 이미지 지도에서 아래쪽에 위치한 대군집 A를 살펴보면 폭소노미 관련 문헌은 대부분 중앙의 군집 1, 4, 10에 소속되어 있고 다른 군집은 타 주제의 문헌들이 대부분이다. 대군집 A의 중앙 군집에는 GOLDER(2005)-“The structure of collaborative tagging systems”, MATHES(2004)-“Folksonomies: Cooperative classification and communication through shared metadata”와 같이 폭소노미 자아 문헌 집합이나 인용 정체성에도 나타났던

폭소노미 관련 핵심 문헌들이 포함되어 있다.

대군집 A의 아래 부분에 해당하는 군집 2는 OREILLY(2005)-“What is Web 2.0”을 비롯한 웹 2.0 관련 문헌들이고, 군집 9는 BERNERS-LEE(2001)-“The semantic web”을 비롯한 시멘틱 웹 관련 문헌들이어서 폭소노미의 인접 주제가 지도의 아래쪽에 위치한 것으로 나타났다. 이는 웹 2.0 및 시멘틱 웹에 대한 연구가 폭소노미 연구와 함께 다루어지고 있음을 의미한다. 웹 2.0은 폭소노미를 포함하는 개념인만큼 폭소노미에 대한 폭넓은 시각의 연구가 이루어지고 있으며, 시멘틱 웹을 폭소노미와 함께 고려하는 융합 연구가 시도되는 것도 짐작할 수 있다.



<그림 9> 패스파인더 네트워크로 표현한 폭소노미 연구의 인용 이미지 지도

〈표 3〉 폭소노미 주제 인용 이미지의 군집별 주요 문헌

PNNC군집		제1저자(연도)	인용 빈도*	제목
C3	C13			
A	1	GOLDER(2006)	163	Usage patterns of collaborative tagging systems
		MARLOW(2006)	36	HT06, tagging paper, taxonomy, Flickr, academic article to read
		HALPIN(2007)	22	The complex dynamics of collaborative tagging
		SIGURBJORNSSON(2008)	18	Flickr tag recommendation based on collective knowledge
	2	BOULOS(2007)	67	The emerging Web 2.0 social software
		OREILLY(2005)	27	What is Web 2.0
		BOULOS(2006)	24	Wikis, blogs and podcasts: A new generation of Web-based tools for virtual collaborative clinical practice and education
		GIUSTINI(2006)	13	How Web 2.0 is changing medicine
	4	MATHES(2004)	47	Folksonomies: Cooperative classification and communication through shared metadata
		HAMMOND(2005)	25	Social bookmarking tools (I): A general review
		MACGREGOR(2006)	24	Collaborative tagging as a knowledge organisation and resource discovery tool
		GUY(2006)	20	Folksonomies: Tidying up tags?
	5	CATTUTO(2007a)	41	Semiotic dynamics and collaborative tagging
		HERLOCKER(2004)	34	Evaluating collaborative filtering recommender systems
		ADOMAVICIUS(2005)	23	Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions
		ZHOU(2007)	22	Bipartite network projection and personal recommendation
	6	HOTH(2006)	35	Information retrieval in folksonomies: Search and ranking
		KLEINBERG(1999)	23	Authoritative sources in a hyperlinked environment
		BRIN(1998)	19	The anatomy of a large scale hypertextual web search engine
	9	MIKA(2007)	21	Ontologies are us: A unified model of social networks and semantics
		GRUBER(2007)	15	Ontology of folksonomy: A mash-up of apples and oranges
		BERNERS-LEE(2001)	12	The semantic web
	10	CATTUTO(2007b)	21	Network properties of folksonomies
		SPECIA(2007)	14	Integrating folksonomies with the semantic web
	12	ZIPF(1949)	17	Human behavior and the principle of least effort
		NEWMAN(2005)	14	Power laws, Pareto distributions and Zipf's law
B	3	BARONCHELLI(2006a)	51	Sharp transition towards shared vocabularies in multi-agent systems
		DALL'ASTA(2006)	27	Nonequilibrium dynamics of language games on complex networks
		STEELS(1995)	23	A self-organizing spatial vocabulary
	11	AXELROD(1997)	18	The dissemination of culture: A model with local convergence and global polarization
		KRAPIVSKY(2003)	17	Dynamics of majority rule in two-state interacting spin systems
		WASSERMAN(1994)	17	Social network analysis: Methods and applications
		GRANOVET(1973)	13	The strength of weak ties
	13	ABRAMS(2003)	14	Linguistics: Modelling the dynamics of language death
		NOWAK(1999b)	12	The evolution of language
C	7	ALBERT(2002)	32	Statistical mechanics of complex networks
		NEWMAN(2003)	23	The structure and function of complex networks
		BOCCALETTI(2006)	17	Complex networks: Structure and dynamics
		DOROGVTSEV(2002)	12	Evolution of networks
	8	BARABASI(1999)	30	Emergence of scaling in random networks
		WATTS(1998)	25	Collective dynamics of 'small-world' networks

* 인용빈도는 폭소노미 주제 인용 이미지 구축자 문헌 집합으로부터 인용된 횟수임

대군집 A의 윗 부분에 해당하는 군집 5는 HERLOCKER(2004)-“Evaluating collaborative filtering recommender systems”와 같은 추천 시스템 관련 문헌들이고, 군집 6은 KLEINBERG(1999)-“Authoritative sources in a hyperlinked environment”와 같은 네트워크 검색 시스템 관련 연구이다. 이는 개인화 추천 시스템 분야의 선행 연구와 네트워크 검색 분야의 선행 연구가 폭소노미 분야에 응용되고 있음을 의미한다. 또한 군집 12는 지프 법칙을 소개한 ZIPF(1949)-“Human behavior and the principle of least effort”와 지프 법칙을 더 일반화한 멱함수 법칙에 대한 NEWMAN(2005)-“Power laws, Pareto distributions and Zipf’s law”로 구성되어 있다. 이는 폭소노미에 대한 분석에서 빈도 분포에 대한 지프 법칙이나 멱함수 법칙의 검증이 이루어지고 있음을 의미한다.

대군집 B와 C의 경우는 폭소노미 주제 문헌이 전혀 포함되어 있지 않으며 타 주제의 문헌들만으로 구성되어 있다. 대군집 B는 WASSERMAN(1994)-“Social network analysis: Methods and applications”나 GRANOVIET(1973)-“The strength of weak ties”와 같은 사회네트워크 분야의 고전들이 포함된 군집 11, ABRAMS(2003)-“Linguistics: Modelling the dynamics of language death”와 같은 언어 시스템의 동역학에 대한 문헌들이 포함된 군집 13, BARONCHELLI(2006a)-“Sharp transition towards shared vocabularies in multi-agent systems”, DALL’ASTA(2006)-“Nonequilibrium dynamics of language games on complex networks”와 같이 어휘 시스템에 대한 연구 문헌이 포함된 군집 3으로 구성되어 있다. 종합적으로 보면 대군집 B는 언어

네트워크와 사회 네트워크, 복잡계 분야의 이론을 도입하면서 폭소노미에 대한 연구가 심화되고 있음을 보여준다.

대군집 C는 ALBERT(2002)-“Statistical mechanics of complex networks”, NEWMAN(2003)-“The structure and function of complex networks”와 같은 네트워크 동역학 연구가 포함된 군집 7과, BARABASI(1999)-“Emergence of scaling in random networks”, WATTS(1998)-“Collective dynamics of ‘small-world’ networks”와 같은 네트워크 모델 연구가 포함된 군집 8로 구성된다. 이는 폭소노미에 대한 연구가 네트워크 과학 분야의 대표적인 문헌들과 함께 인용되고 있음을 나타내는 것으로서, 폭소노미가 네트워크 과학 연구의 대상도 되고 있음을 시사한다.

폭소노미 주제에 대한 인용 이미지 구조가 앞 절의 인용 정체성 구조와 다른 점은 주요 문헌 중에서 폭소노미 주제 문헌의 비중이 더 낮다는 점, 그리고 타 주제 문헌이 폭소노미 연구 문헌과 섞이지 않고 독립적으로 형성한 군집이 많다는 점이다. 이는 폭소노미 주제의 구체적인 연구동향을 살펴보기에는 불리하지만 폭소노미 연구에 응용되거나 폭소노미 연구와 융합되는 타 주제를 파악하는데 유리한 조건이 된다.

4. 결 론

전통적인 계량서지적 주제분석 기법인 문헌 동시인용 기법은 특정 주제 문헌 집합 내부를 미시적으로 분석하는 방법이었다. 이 연구에서 제안한 자아 중심 주제 인용 분석 기법은 주제

영역 내부에 대한 미시적인 분석과 함께 주제 영역의 외부에서 살펴보는 거시적인 분석을 추가하여 다각적이고 포괄적인 해석을 가능하게 하는 기법이라고 할 수 있다.

폭소노미 연구 영역을 대상으로 연구 문헌을 검색하고 이들을 인용한 문헌과 이들로부터 인용된 문헌을 파악하여 자아 중심 주제 인용 분석을 수행한 결과는 다음과 같다.

자아 문헌 집합에 대한 미시적인 분석에서는 폭소노미 관련 연구가 폭소노미의 본질이나 특성을 연구하는 영역과 폭소노미의 응용을 연구하는 영역으로 크게 양분되는 것으로 나타났다. 폭소노미의 본질 및 특성 영역에는 폭소노미의 개념과 이용, 색인으로서의 폭소노미, 도서관에서의 수용과 폭소노미 입장에서 폭소노미를 바라보는 문헌들이 포함되어 있었다. 폭소노미의 응용을 연구하는 영역에는 폭소노미를 응용한 검색, 브라우징, 개인화 추천, 분류 등에 대한 문헌이 포함되어 있었다. 이처럼 자아 문헌 집합에 대한 분석을 통해 폭소노미 연구 영역의 세부 연구 주제와 주요 문헌이 파악되었다.

폭소노미 관련 연구에서 인용된 문헌들로 폭소노미 주제의 인용 정체성을 분석해본 결과, 폭소노미에 대한 핵심 문헌 및 관련된 타 분야의 주요 선행 문헌들이 파악되었다. 폭소노미에 대한 핵심 문헌은 폭소노미의 개념적 이해와 활용 방안, 검색 및 추천 시스템에서의 응용 등과 같은 주제를 다루고 있었다. 폭소노미 연구에서 인용된 주요 선행 연구들은 크게 나누어보면 인지심리, 정보검색, 맥함수법칙 분야의 고전이 된 문헌들과, 시맨틱 웹이나 개인화 추천 관련 문헌과 같은 2000년대 이후의 선행 연구들이 포함된 것으로 나타났다. 이런 타 주제

의 선행 문헌들이 폭소노미를 다룬 관련 문헌의 군집에 나누어 배치됨으로써 폭소노미 연구의 세부 주제별로 활용된 다른 주제의 선행 연구가 어떤 것인지에 대해서 부분적인 파악이 가능하였다.

폭소노미 연구 문헌과 동시인용된 문헌들로 구성된 인용 이미지를 분석한 결과 폭소노미 연구에 응용되거나 폭소노미 연구와 융합되는 타 주제를 파악할 수 있었다. 인용 이미지를 구성하는 문헌 중에는 폭소노미가 아닌 타 주제 문헌의 비중이 높았으며, 이들이 별도의 독립된 군집으로 파악되므로 주제를 식별하기도 용이하였다. 폭소노미 자아 문헌 집합이나 인용 정체성에도 나타났던 폭소노미 관련 핵심 문헌들은 인용 이미지 지도에서는 타 분야의 문헌들과 독립적으로 군집을 형성하였다. 인용 이미지 분석에서는 웹 2.0 및 시맨틱 웹, 개인화 추천 시스템, 네트워크 검색, 지프 법칙과 맥함수법칙, 사회네트워크 분석, 언어 체계의 동역학, 복잡계로서의 어휘 체계, 네트워크 모델, 네트워크 동역학 등과 같은 타 분야의 이론과 연구가 폭소노미에 도입되거나 적용되면서 폭소노미에 대한 연구가 심화되고 있음이 드러났다.

이상과 같이 폭소노미에 대한 자아 문헌 집합 분석, 주제 인용 정체성 분석, 주제 인용 이미지 분석은 폭소노미 연구 주제의 미시적인 구조를 드러내줄 뿐만 아니라 폭소노미와 관련되는 타 주제를 포함한 거시적인 해석도 가능하게 하였다. 자아 문헌 집합 분석에서는 폭소노미 연구가 본질 및 특성 영역과 응용 영역으로 구성된다는 점이 드러났고, 주제 인용 정체성 분석에서는 세부 연구 주제별로 인용되는 외부 주제 연구들이 나타났으며, 주제 인용 이미지

분석에서는 폭소노미 연구에 도입되거나 적용되는 타 분야의 이론과 연구가 구체적으로 파악되었다. 향후에는 시간 흐름에 따른 미시적인 변화와 거시적인 변화를 함께 포착하여 특정 주제 영역의 연구가 발전, 성장하는 과정을 다면적으로 해석하는 접근도 가능할 것으로 기대된다.

참 고 문 헌

- 김관준, 이재운 (2010). 학술지 영향력 측정을 위한 h-지수의 응용에 관한 연구. *정보관리학회지*, 27(1), 269-287. <http://dx.doi.org/10.3743/KOSIM.2010.27.1.269>
- 이재운 (2006a). 지적 구조의 규명을 위한 네트워크 형성 방식에 관한 연구. *한국문헌정보학회지*, 40(2), 333-355.
- 이재운 (2006b). 지적 구조 분석을 위한 새로운 클러스터링 기법에 관한 연구. *정보관리학회지*, 23(4), 215-231.
- 이재운 (2011). WNET. (Version 0.4) (Software).
- 이재운 (2012). 자기 인용 네트워크와 인용 정체성을 이용한 연구자의 연구 이력 분석에 관한 연구. *정보관리학회지*, 29(1), 157-174.
- Banks, M. G. (2006). An extension of the Hirsch Index: Indexing scientific topics and compounds. *Scientometrics*, 69(1), 161-168. <http://10.1007/s11192-006-0146-5>
- Braun, T., Glänzel, W., & Schubert, A. (2005). A Hirsch-type index for journals. *The Scientist*, 19(22), 8.
- Kessler, M. M. (1963). Bibliographic coupling between scientific papers. *American Documentation*, 14(1), 10-25.
- Lee, Jae Yun, & Choi, Sanghee (2011). Intellectual structure and infrastructure of informetrics. *Journal of the Korean Society for Information Management*, 28(2), 11-36. <http://dx.doi.org/10.3743/KOSIM.2011.28.2.011>
- Small, H. (1973). Co-citation in the scientific literature: A new measure of the relationship between two documents.
- Tsay, Ming-yueh (2011). A bibliometric analysis and comparison on three information science journals: JASIST, IPM, JOD, 1998-2008. *Scientometrics*, 89(2), 591-606.
- White, H. D. (2000). Toward ego-centered citation analysis. In B. Cronin, & H.B. Atkins (Eds.), *The web of knowledge: A festschrift in honor of Eugene Garfield* (pp. 475-496). Medford,

New Jersey: Information Today, Inc.

White, H. D., & Griffith, B. (1981). Author cocitation: A literature measure of intellectual structure. *Journal of the American Society for Information Science*, 32(3), 163-171.

Yanping, L. (2011). Research on the journal -based citation identity and citation image. *Information Studies: Theory & Application*, 2011-2. Retrieved from http://en.cnki.com.cn/Article_en/CJFDTotat-QBLL201102011.htm

• 국문 참고문헌에 대한 영문 표기 (English translation of references written in Korean)
--

Kim, Pan-Jun, & Lee, Jae Yun (2010). A study on journal impact measurement with Hirsch-type indices. *Journal of the Korean Society for Information Management*, 27(1), 269-287. <http://dx.doi.org/10.3743/KOSIM.2010.27.1.269>

Lee, Jae Yun (2006a). A study on the network generation methods for examining the intellectual structure of knowledge domains. *Journal of the Korean Society for Library and Information Science*, 40(2), 333-355.

Lee, Jae Yun (2006b). A novel clustering method for examining and analyzing the intellectual structure of a scholarly field. *Journal of the Korean Society for Information Management*, 23(4), 215-231.

Lee, Jae Yun (2011). WNET. (Version 0.4) (Software).

Lee, Jae Yun (2012). Exploring a researcher's personal research history through self-citation network and citation identity. *Journal of the Korean Society for Information Management*, 29(1), 157-174. <http://dx.doi.org/10.3743/KOSIM.2012.29.1.157>