과학계량학 지표 및 Emerging Topic 탐색 기법 연구 동향 분석

구본진, 장덕현, 정영임, 김재훈, 강무영

목차

1. 연구 영향력, 과학적 수월성에 대한 계량학적 지표 ····· 1

1.1 인용기반 지표 동향 ····· 1

1.2 이용기반 지표 동향 ····· 10

1.3 Altmetrics ····· 13

2. Emerging Research Topic 탐지 모델의 연구 동향 분석 ····· 21

2.1 유망 기술 및 유망 연구 영역 탐지의 필요성 ····· 21

2.2 유망 기술 및 유망 연구 영역 탐지 모델의 연구 동향 분석 ····· 22

3. 요약 및 시사점 ····· 33

참고문헌 ····· 35

표 목차

<표 1> 인용기반 지표 정리 ····· 9

<표 2> 계량서지적 기법 분류 ····· 24

그림 목차

<그림 1> h-index ····· 4

<그림 2> 인용행태에 따른 저널의 명성 ····· 7

<그림 3> 문헌종류별 평균 피인용과 다운로드 건수 추이 ····· 11

<그림 4> 연구분야별 평균 피인용과 다운로드 건수 추이 ····· 11

<그림 5> Altmetric.com의 유형별 데이터 규모 ····· 17

<그림 6> Altmetric.com의 Coverage ····· 17

<그림 7> 유럽지역 생화학, 유전학 및 분자생물학 저자들의 지리적 분포 ····· 25

<그림 8> 2010년 국내 바이오분야 논문등고선 ····· 25

<그림 9> 세 지표간 시간적 변화 추이 비교 ····· 29

<그림 10> 유망 연구 영역의 식별을 위한 클러스터 유형 ····· 30

<그림 11> Nano Pollution(좌)과 Biofuel(우) 분야의 출판활동 및 협력 현황 ····· 30

1. 연구 영향력, 과학적 수월성에 대한 계량학적 지표

1.1 인용기반 지표 동향

1) Impact Factor(IF)

- Impact Factor(IF)는 1955년, Eugene Garfield가 제안한 저널 평가 지표로, 가장 많이 알려져 있고, 가장 대중적으로 널리 사용되는 지표 중 하나

- 최근 2년간 해당 저널에 수록된 논문이 당해 연도에 몇 번의 피인용을 받았는가를 측정

$$IF = \frac{\text{지난 2년 동안 출판된 논문이 해당 년도에 인용된 총 횟수}}{\text{최근 2년간 해당 저널에 실린 (인용 가능한)논문의 수}}$$

- Impact Factor는 도서관에서 저널 수서를 위한 의사결정 도구, 저널의 순위 결정, 원문(full-text) 데이터베이스의 선정과 평가, 저널의 학술적 영향력 측정, 편집자나 출판사 입장에서의 저널 평가 또는 연구자의 입장에서 자신의 논문 기고를 위한 저널 선정, 저널 연구, 성과평가(연구자, 대학, 국가의 연구 성과 평가 등) 등 다양한 영역에서 사용됨(Nisonger 2004)

- Impact Factor의 장점

  - 상당한 권위를 가지고 있음
  - 잘 알려져 있고, 다수의 목적을 위해 사용됨
  - 계산이 용이하며, 이해하기 쉬움
  - 객관적인 판단이 가능
  - IF가 저널의 질과 영향력을 나타낸다는 가정을 내포하고 있음

- Impact Factor의 단점/문제점

- 분모의 부적합성 : “인용 가능한” 논문에 연구논문, 리뷰 등이 포함되며, 레터, 편집자란, 학술회의 초록 등은 포함되지 않음(Moed and Van Leeuwen 1995)
- 학문 혹은 학술지간 인용 특성을 반영하지 못함 : 학문 분야 또는 연구 분야에 따라 포함하고 있는 학술지의 수나 논문수가 다르고, 인용빈도 또는 인용의 관행이 다름, 즉, 연구 분야나 학술지의 종류에 따라 2년 이상의 시간이 필요할 수도 있음
- 자기 인용과 부정적 인용(비판, 반박 등)을 구분할 수 없음
- 최신 논문의 경우 평가가 불가능함
- 조작 가능성 : 리뷰논문 작성을 유도하거나 자기 학술지 또는 같은 출판사나 단체에서 발행하는 학술지의 인용을 권고하기도 함
- 성과평가에 무분별하게 사용[1)]

2) Impact Factor의 수정 및 대안

- Impact Factor의 경우 Garfield가 처음 고안할 당시 도서관에서 학술지의 구매를 위해 의사결정 도구로 활용되었으나, 최근에는 연구개발 투자에 대한 정당성과 근거를 위한 연구 성과 평가 기준으로 많이 활용됨

- 이처럼 R&D 투자에 대한 기준으로 학술지의 영향력을 측정하고 평가하기 위해 사용하는 Impact Factor에 대한 활용 범위가 가장 넓지만 많은 한계를 내포하고 있어 객관적인 지표로 부적합하다는 의견이 있음

- 특히, 분모의 부적합성, 인용 속도 및 양의 문제, 논문의 유형에 따른 여러 가지 문제점을 해결하기 위하여 Impact Factor의 공식을 수정하거나 정교한 척도를 개발할 필요성이 제기되었으며(Glänzel and Moed 2002), 대표적으로 피인용 기간을 조정하거나 Impact Factor를 정규화하려는 노력이 이루어짐

---

1) Thomson사에서도 Impact Factor를 “연구자 개인의 실적을 평가하는데 사용하지 말 것. JCR에서 제공되는 다른 모든 수치와 마찬가지로 “저널” Impact Factor는 저널을 평가하는 데에만 사용되어야 한다.” 라고 공식 입장을 밝히고 있음에도 연구 성과의 측정에 사용됨에 따라 많은 문제점이 지적됨

- 피인용 기간의 조정

  - 피인용 기간을 조정하는 것은 학술지간 노화의 차이를 두드러지게 하며, 피인용 기간을 조정한다고 하더라도 학술지 순위에는 크게 영향을 미치지 않는다(Garfield 1998)는 연구 결과가 있으나, 피인용 기간의 확대 필요성을 주장하는 연구들이 꾸준히 이루어짐

  - Adler, Ewing and Taylor(2008)는 자신들의 연구에서 생명과학은 수학이나 사회과학보다 높은 Impact Factor를 가지는 경향이 있음, 특히 수학자들은 논문 한 편 당 참고문헌 수가 적으며, 오래된 문헌을 중요하게 여기므로 수학자들이 인용하는 대다수의 문헌은 2년의 범위는 넘어가는 경우가 많아 인용 시기를 2년으로 제한하는 것은 문제가 있다고 지적하였으며, 이 외에도 5년으로 조정(Butler 2002) 및 3년 또는 4년으로 조정(Glänzel 2006)할 것을 제안하기도 함

- IF의 정규화

  - 분야별 영향력 지표(DIF, Discipline Impact Factor) : 단일 세부 분야에 속하는 핵심 문헌에서 학술지 한 종이 인용되는 횟수를 측정(Hirst 1978)

  - DIS(Discipline Influence Score) : 특정 분야에 속한 학술지들이 인용한 횟수만 반영하는 방식(He and Pao 1986)

  - 학술지 분야별 영향력 지표(JFIS, Journal to Field Impact Factor) : 분자와 분모를 모두 동일 아이템으로 집계, 개별 학술지가 가진 영향력을 해당 연구 분야의 전체 인용 평균과 비교(Van Leeuwen and Moed 2002)

  - RCII(Relative Citation Impact Index) : 분야 간 비교가 가능하도록 상대적인 영향력 측정을 시도(Glanzel 2003), RCR(Relative Citation Rate)[2] 값을 정규화한 지수

- 이처럼 Impact Factor의 단점을 보완하기 위한 많은 노력에도 불구하고 여전히 성과평가에 무분별하게 사용됨, 특히 Impact Factor가 분야 간 평

2) 인용지수의 기댓값 대비 실제 값을 구하는 것으로, 인용지수의 기댓값은 해당 주제 분야 학술지 집단의 평균값을 사용함

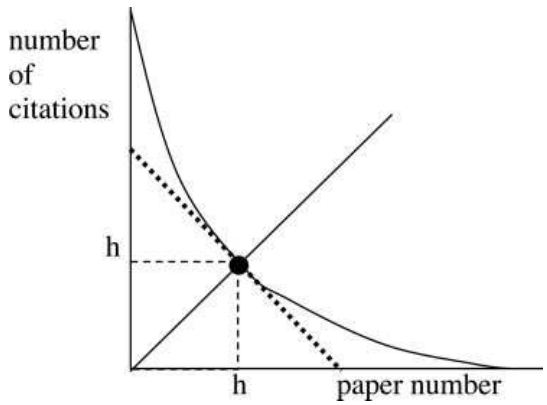
가가 어렵고, 단순히 Impact Factor를 비교하는 것에 문제가 있다는 지적에 따라 연구자, 기관, 국가 등을 안정적으로 비교할 수 있는 추가 지표에 대한 수요가 발생

3) h-index

- 연구성과를 평가하기 위한 추가 지표에 대한 수요 발생과 함께 가장 큰 주목을 받은 지표인 h-index는 2005년에 물리학자 조지 허쉬가 제안

- h-index는 과학자의 생산력과 영향력을 측정하기 위한 지표로 "어떤 연구자의 h-index가 h라는 것은 그가 출간한 N_p개의 논문 가운데 최소한 h회 이상의 인용을 받고 있는 논문의 수가 h개 이다."라고 정의(Hirsch 2005)

논문수	인용수
A1	30
A2	10
A3	8
A4	6
A5	5
A6	1
A7	0
h=5	



<그림 1> h-index (Hirsch 2005 재정리)

- h-index의 장점

  - 계산이 용이하고, 직관적 이해가 가능
  - 생산성과 누적 연구의 영향력을 함께 제공
  - 연구자 뿐 아니라 학술지, 국가 단위에도 적용 가능
  - 지수의 강건성 : 피인용도가 낮은 여러 논문이나 높은 몇 논문에 영향을 받지 않음

- h-index의 단점(이재윤 2006)

  - 지수값이 출판물의 총 수를 넘지 못함
  - 자기 인용 및 공저자의 영향을 많이 받음
  - 상이한 학문 분야의 연구자에 대한 비교가 어려움
  - 연구생산성이 높을수록 h-index가 높아질 가능성이 있음
  - 변별력이 낮음

- 이후 학술지 평가를 위한 h-index와 h-index의 변형 지수(Hirsch 유형 지수) 등 h-index를 개선하거나 한계를 보완한 다양한 지표의 개발이 활발히 이루어짐

  - Tibor, Glänzel, and Schubert(2006)은 학술지 평가를 위한 h-index를 제안하였으며, "특정 학술지에 게재된 N개의 논문 중에서 h개의 논문은 최소 h번 이상 인용되었고, 나머지 (N-h)개의 논문은 h번 이하 인용되었다면 그 학술지의 h-지수는 h이다."라고 정의

  - Rousseau(2006a)는 h-index이상의 인용을 받은 학술지 게재 논문들을 h-core 또는 Hirsch-core라고 명명함

  - 학술지 평가를 위한 h-index 역시 산출방법이 단순하고, 직관적인 해석이 가능하지만 변별력이 낮고, 소규모 학술지에 불리하게 작용하는 단점이 있음(Batista, Campiteli and Kinouchi 2006; Braun et al. 2006)

  - Rousseau(2006b)는 학술지 h-index의 정규화를 위해 h-index를 전체 논문수로 나누어 상대적 h-index를 제안

  - Egghe(2006)는 h-index가 인용빈도 상위논문들(h-core)에 대하여 이후에 추가 인용이 발생하더라도 지수 값의 증감에 거의 영향을 미치지 못한다는 문제를 보완하기 위해 g-index를 제안하였으며, "특정 학술지에 게재된 Np개의 논문 중에서 인용빈도 상위 g개 논문의 피인용 횟수의 합이 g의 제곱 이상인 최하위 순위 g가 그 학술지의 g이다."라고 정의

  - 이외에도, 단일저자가 공저자보다 불리하다는 문제를 보완하기 위해 제안된 hI-지수, h-지수가 h-core에 속한 논문들에 대한 전체 인용빈도를 반영

하지 못하는 문제를 보완하기 위해 제안된 A-지수, g-지수를 개선하기 위해 제안된 R-지수, h-지수의 최대값 문제를 해결하기 위해 제안된 tapered h-지수, h-지수와 g-지수의 절충안으로 두 지수의 기하평균을 구하는 hg-지수 등 다양한 Hirsch 유형 지수들이 연구되었음(Batista, Campiteli and Kinouchi 2006; 김판준, 이재윤 2010)

4) SNIP

- Leiden대학 CWTS(Center for Science & Technology Studies)의 Henk Moed(2010)에 의해 개발된 평가지표로 학문분야별 상대적 지표를 나타냄

SNIP = 해당 연도의 Impact / 해당연도의 인용 잠재력(Citation Potential)

※ 인용잠재력 : 최근 3년 내의 논문을 인용한 참고문헌의 논문 당 평균

- 주제분야의 인용패턴을 고려하고, 논문이 해당 주제 분야에 영향을 미치는 속도, 주제분야 학술지의 커버리지를 고려하며, 편집자의 조작이 지수에 영향을 미치지 못하도록 고안

- 서로 다른 분야의 학술지를 비교하는데 강점이 있으며, 동일 분야의 학술지는 IF와 거의 동일한 값을 산출한다는 특징이 있음

- Waltman, Ludo, et al.(2013)은 에서는 분모의 값을 낮출 수 있는 방안을 도입하여 SNIP2를 제시, 참고문헌 수가 적은 학술지로부터의 인용에 대한 가중치 부여가 적절하게 되었는지 확인하기 위한 계수를 보정하고, 학술지 지수간의 비교를 용이하게 하기 위해 Scopus에 등재된 학술지의 평균 SNIP값을 1로 책정함

$$\text{SNIP2} = \frac{\text{RIP(Raw Impact Per Paper)}}{\text{RDCP(Relative Database Citation Potential)}}$$

※ RIP : 특정 학술지에서 A년 이전 3개년 간에 출판된 논문들이 A년에 인용된 횟수/A년 이전 3개년 간에 발표된 논문 수

※ RDCP : 데이터베이스에 포함된 학술지들에 대해 DCP 순서로 나열했을 때의 중간값에 해당하는 학술지의 DCP

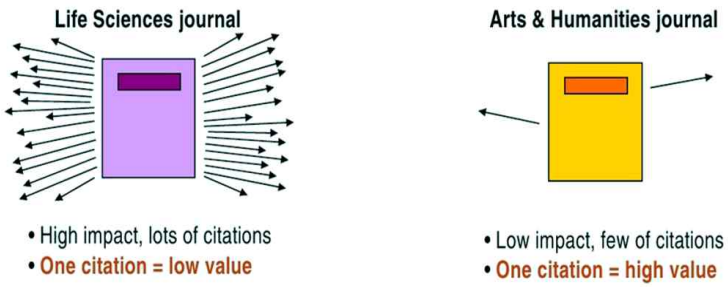
※ DCP : 특정 학술지에서 A년 이전 3개년 간에 출판된 논문들을 A년에 인용한 모든 논문에서 인용한 참고문헌들 중 A년 이전 3개년 간 출판된 참고문헌들의 총 수/인용한 논문들의 총수

5) SJR

- '모든 인용을 동등하게 평가하면 안된다'는 전제 하에 페이지랭크 알고리즘을 기반으로 개발되었으며, 인용학술지의 명성(prestige)을 피인용횟수에 반영

  - 학술지로부터 받은 인용 횟수와 인용학술지의 영향력을 모두 고려하며, 인용 학술지의 중요도에 따라 인용에 다른 값을 가중함으로써 영향력이 높은 학술지의 인용을 더 가치 있게 함(Guerrero et al. 2009)

  - 저널의 명성은 인용을 통해 이전(transfer)하며, 명성의 이전은 즉, SJR로 표현됨(SJR이 높다는 것은 명성이 높다는 것을 의미), 또한 저널의 명성은 인용횟수를 나누어 가짐(주제 분야의 인용 행태에 따라 다르게 표준화 함)



<그림 2> 인용행태에 따른 저널의 명성

- SJR의 특징

  - 페이지랭크 알고리즘 사용(네트워크 기반의 지수)
  - 피인용 기간이 3년으로 조금 더 안정적이며, 자기인용은 1/3을 허용
  - 높은 피인용 횟수를 보이면서도 SCI에 등재되지 않은 오픈액세스 저널의

평가에 유용

- 쉽게 해석이 어려운 결과를 도출

6) Eigenfactor와 AI(Article Influence)

- 역시 페이지랭크 알고리즘을 기반으로 하며, 하나의 저널이 다른 저널에 비해 상대적으로 어느 정도의 영향력을 가지를 가를 평가함

- Eigenfactor score의 총 합은 100으로 직관적으로 이해하기 용이하며, AI는 개별 논문의 영향력을 파악하기 위하여 Eigenfactor score를 게재된 논문의 수로 나눈 값으로, 한 저널에서 출판된 후 5년 동안의 논문에 대한 상대적인 평균 영향력 값이라 할 수 있음

- AI의 피인용 기간은 5년이며, 자기 인용은 제외함, 평균 AI는 1로 1보다 큰 값은 전체 논문의 평균 영향력보다 크고, 1보다 작은 값은 전체 평균보다 영향력이 적음을 의미

※ 각 지표별 비교

<표 1> 인용기반 지표 정리

	IF	SNIP	SJR	AI
피인용기간	2년	3년	3년	5년
인용기간	1년	1년	1년	1년
자기인용포함여부	포함	포함	1/3 포함	제외
정규화여부	×	○	○	○
데이터베이스	JCR	Scopus	Scopus	JCR
분자에 사용하는 문헌의 유형	모든 아이템	논문, 컨퍼런스 페이퍼, 리뷰	논문, 컨퍼런스 페이퍼, 리뷰	모든 아이템
분모에 사용하는 문헌의 유형	논문, 리뷰	논문, 컨퍼런스 페이퍼, 리뷰	논문, 컨퍼런스 페이퍼, 리뷰	논문, 레터, 리뷰
인용 문헌의 영향력(중요도)의 역할	없음	없음	인용문헌이 게재된 저널의 명성에 근거하여 인용을 가중	인용문헌이 게재된 저널의 명성에 근거하여 인용을 가중
리뷰 논문을 많이 포함할 때의 효과	지표값을 증가시킴	지표값을 증가시킴	리뷰를 인용하는 저널의 SJR값에 의해 수정됨	리뷰를 인용하는 저널의 Eigenfactor값에 의해 수정됨

1.2 이용기반 지표 동향

- 처음에는 전자저널의 급격한 증가로 인해 전자저널을 포함한 전자정보원 구입 예산이 도서관의 자료구입비에서 상당한 비중을 차지하게 되면서 저널의 이용기록을 분석하려는 시도가 이루어짐

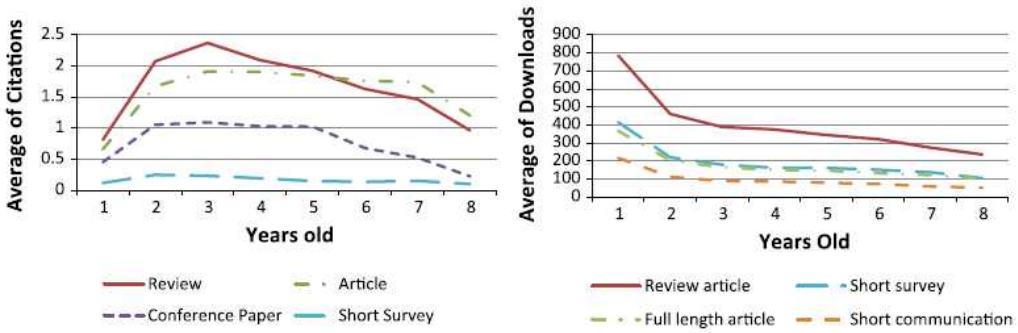
- 이후, 기존의 인용 기반의 지표들의 한계를 보완하기 위하여 이용량에 기반하여 저널의 영향력을 측정하려는 연구가 진행됨

  - Moed, H. F. (2005)는 그의 연구에서 이용은 출판 이후 수개월 내에 즉시 정점을 찍고 하락세로 돌아서지만 인용은 더 긴 시간 분포를 보여 인용과 이용의 연령 분포가 다르게 나타남을 확인했으며, ScienceDirect의 논문 다운로드 횟수와 학술지 인용횟수 사이의 상관관계 분석을 통해, 특정 분야에서는 이용이 인용을 예측하는 데 도움을 줄 수 있다고 주장

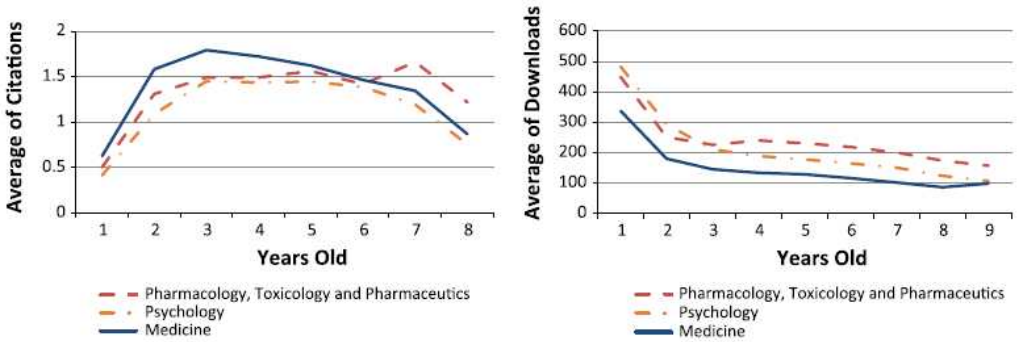
  - 이외에도 Schloegl, C., & Gorraiz, J. (2010), Brody, T., Harnad, S., & Carr, L. (2006), McDonald, J. D. (2007), Duy, J., & Vaughan, L. (2006) 등이 인용과 이용을 비교하는 연구를 진행하였으며, 학문 영역의 범위나 주제에 따라 상관관계가 다르게 나타나지만, 인용과 이용 간에는 유의미한 상관성이 있으며, 이용통계가 논문의 인용 영향력을 예측할 수 있는 도구임을 제안

- 학술 영향력의 측정이라는 측면에 있어서, 아직까지는 여전히 인용에 초점이 맞추어져 있으나, 새로운 주제나 연구 영역의 경우 인용 데이터가 축적 되는데 걸리는 시간을 고려한다면, 새로운 연구 주제나 이슈를 모니터링 하는데 이용량 기반 지표는 유용성이 있음

  - Guerrero-Bote, V. P., & Moya-Anegón, F.(2013)의 연구에서 피인용과 이용의 다운로드 건수 추이를 보면, 피인용과 이용간의 잠복기가 현저하게 차이가 남. 즉 이용량 기반 지표의 경우 출판과 동시에 이용이 이루어지므로 인용데이터에 비해 급부상하는 연구 주제나 이슈를 빠르게 탐지할 수 있다는 장점이 있음



<그림 3> 문헌종류별 평균 피인용과 다운로드 건수 추이
(Guerrero-Bote, V. P., & Moya-Anegón, F. 2013)



<그림 4> 연구분야별 평균 피인용과 다운로드 건수 추이
(Guerrero-Bote, V. P., & Moya-Anegón, F. 2013)

- 각 출판사들은 저널 이용과 관련하여 자체 제작한 이용 데이터를 사용하는 보고용 형식을 가지고 있었으나, 이에 대한 신뢰성과 일관성의 문제가 대두되고, 출판사별로 각기 다른 데이터 형식으로 인해 이를 통일하고 비교하는 데 많은 어려움이 있음, 따라서 이러한 이용 통계를 일관성 있고, 신뢰할만하며, 호환 가능한 방식으로 표준화 하려는 노력이 시도

- COUNTER 프로젝트

  - Counting Online Usage of Networked Electronic Resources의 약자로, 출판사, 도서관 단체 및 정보중개업체가 주축이 되어 현실적으로 이용 가능한 전자정보통계자료 및 보고양식에 대한 표준을 개발하기 위한 프

로젝트

- 여러 벤더들로부터 이용통계를 제공 받더라도 서로 비교 가능할 수 있게 하기 위한 목적으로 2002년 3월 공식적으로 수립

• MESUR 프로젝트

- 인디애나 주립대의 Bollen교수가 진행한 프로젝트로 학술영향력 지수로서의 이용데이터의 가능성을 살펴본 결과, 이용량 기반의 지수가 다양한 형태의 학술 커뮤니케이션을 평가하는데 새롭고 광범위한 시각을 제공한다고 주장

• SERUM 프로젝트

- 전세계적인 도서관들 사이의 협력 네트워크 구축을 통해 전자자원의 이용기록을 통합하고, 개별 기관의 이용 기록을 타 기관이나 전체 기록과 비교할 수 있게 하며, 적절한 이용 지표와 저널 순위를 산출하는 것을 목적으로 함

- 총다운로드 횟수(DT), 저널 이용 지수(JUF), 즉시 다운로드 횟수(DI), 다운로드 반감기(DHL)의 4가지 지표를 저널별로 산출하며, 이를 논문, 리뷰논문, 학술대회 발표논문, 레터, 노트 및 기타의 5가지 유형별로 구분하여 제공

• Journal Usage Factor

- COUNTER그룹에서 개발한 저널 이용 지표

$$JUF = \frac{\text{Total usage over period x of items published during period y}}{\text{Total items published online during period y}}$$

1.3 Altmetrics

- 논문의 영향력을 측정하는데 있어서 기존에 주로 사용되어오던, 특히 Impact Factor, H-index, 그리고 이 두 지표를 기반으로 변형시킨 다양한 인용기반의 지표들이 가지고 있는 여러 가지 문제점들로 새로운 학술 영향력 측정을 위한 움직임이 일어남

- Altmetrics라는 용어가 사용된 것은 2010년으로, Priem이 소개(Priem et al. 2010)하였고, 다른 도구를 사용하여 소셜 웹에서의 광범위한 연구 영향력을 측정하는 대안적 방법으로 정의함

- 초기에는 이전에 사용되던 지표들의 일부 한계를 극복하고 연구 평가에 새로운 관점을 제공하기 위해 제안(Priem and Hemminger 2010; Galligan and Dyas-Correia 2013; Bornmann 2013)되었으며, 이용데이터 분석, 웹 인용과 링크 분석, 소셜 웹 분석을 포함하는 개념이었음

- 특히, 최근 학술문헌의 온라인 출판 증가와 함께 블로그, SNS, 온라인 토론 사이트 등 다양한 채널을 통해 논문이 공유되고, 논문에 대한 의견 교환이 이루어지고 있으며, 소셜 북마크, 서지정보 관리 도구 등을 활용해 연구자들 간 서지정보의 교환이나 추천 등이 이루어지기도 함

- 논문의 영향력을 측정하는 새로운 방법인 Altmetrics는 alternative metrics의 합성어로, 소셜 미디어를 이용하여 논문의 영향력(Impact)을 분석하는 것, 논문 공개 후 트위터 등 여러 소셜 미디어로부터의 반응을 정량화 하는 것을 의미

- Altmetrics 분석 대상은 뉴스 및 언론보도, 블로그와 트위터 등의 소셜 미디어 뿐 아니라 참고문헌 관리 도구, 피어리뷰 사이트, 정책문서 등 다양함

  - Priem & Hemminger(2010)는 소셜 북마크, 참고문헌 관리도구, 웹사이트

추천 도구, 댓글, Microblogging, 위키백과, 블로그, 기타 비디오 및 데이터 세트 등을 기반으로 Altmetrics 지표를 확장해 갈 수 있을 것이라고 하였음

- Altmetrics 지표는 인용 횟수 뿐 아니라 보기, 저장, 공유, 토론, 태그의 횟수를 포함하며, Altmetrics의 개념이 발달하면서 소셜 미디어에서 과학 출판물 및 자료에 대한 공유, 좋아요, 코멘트, 리뷰, 토론, 북마크, 저장, 트윗, 멘션을 찾아내는 것으로 연구 성과의 실시간 영향력을 추적하려는 목적을 가진 다양한 측정도구의 발달을 수반함(Wouters and Costas 2012)

- Wouters and Costas(2012)는 Altmetrics의 네 가지 논점을 언급
  - 다양성과 필터링 : 블로그, 트위터 등 소셜 웹 등 다양한 자료의 분석이 가능하며, 그렇기 때문에 다양한 측정 도구 간 균형을 맞추어야 함
  - 신속성 : 데이터 수집 및 검색이 용이하며, 즉시 사용 가능
  - 개방성 : Altmetrics 데이터는 다운로드와 이용이 쉬움
  - 많은 웹 기반의 과학 커뮤니케이션 활동 추적이 인용분석이나 피어리뷰와 같은 활동으로 측정할 수 없는 과학적 성과의 측면을 나타내는 척도로 사용 가능

- Piwowar(2013)는 Altmetrics의 장점으로 다음의 네 가지를 언급
  - 영향력의 미묘한 차이를 나타낼 수 있으며, 어떤 학술자료가 인용되는지 뿐만 아니라 어떤 것을 읽고, 토론하고, 저장하는지를 알 수 있음
  - 연도가 아닌 매일의 영향력을 알 수 있어 더 적시성 있는 데이터를 얻을 수 있음
  - 데이터 세트, 소프트웨어, 블로그, 비디오 등과 같은 웹을 이용한(web-native) 학술자료의 영향력을 나타낼 수 있음
  - 학자 뿐만 아니라 실무자, 임상자(clinicians), 교육자, 일반 대중 등 다양한 집단에 의한 영향력을 파악할 수 있음

- 최근 Altmetrics에 관해 학술적 영향력을 평가하는 지표로서의 Altmetrics의 활용, 오픈액세스 관점에서 Altmetrics에 대해 논의한 연구 등이 진행되었으며, Altmetrics가 기존의 인용분석을 대체하거나 보완할 수 있는지를 검증하기 위한 연구 등 다양한 측면에서 연구가 진행되고 있음

- 영향력 평가 도구로서의 Altmetrics의 유효성을 검증한 연구

  - Li et al.(2012)는 연구 성과 평가를 위해 참고문헌 관리도구인 CiteULike와 Mendeley를 대상으로 장·단점 및 유용성을 평가하였으며, Mendeley가 연구 성과 평가에 있어 더 유용함을 주장

  - Shuai et al.(2012)는 arXiv.org를 대상으로 한 연구에서 소셜 미디어가 논문의 영향력을 결정하는 중요한 지표가 될 수 있다고 주장

- Altmetrics의 가장 큰 장점중 하나는 기존의 피인용 중심의 평가 방식을 보완할 수 있다는 점

  - Hammarfelt(2014)는 인문학 분야의 논문 및 단행본을 대상으로 Altmetrics 분석을 수행한 결과 그동안 피인용되지 않았던 다수의 인문학 분야의 논문 및 단행본이 연구의 초기 단계 및 기타 전문가적 활동에 활용되고 있음을 밝힘

  - Mohammadi and Thelwall(2014) 또한 기존에 Bibliometrics로는 나타나지 않았던 인문학 분야 및 사회과학 분야 논문의 영향력을 Altmetrics 도구인 Mendeley로 밝혀냄

  - Zahedi, Costas and Wouters(2014)는 피인용도와 Mendeley와의 상관성을 분석한 연구에서 Altmetrics가 기존의 피인용 중심의 평가 방식을 보완할 수 있음을 주장

  - Bornmann(2014)는 트위터, 블로그, CiteULike, Mendeley와 전통적 피인용도간의 상관성을 분석한 연구에서 CiteULike, Mendeley와 같은 참고문헌 관리도구의 경우 피인용도와 강한 상관관계를 갖는 것으로 나타나 참고문헌 관리도구의 경우 기존의 피인용 방식을 보완할 수 있는 가능성

이 있다고 주장

- 또한, Altmetrics는 웹에서의 반응을 지표로 하고 있기 때문에 소셜미디어의 초기 언급수 관측을 통해 피인용수의 예측이 가능하며, 이를 통해 연구 트렌드를 파악하고 향후의 동향 예측에도 도움이 될 수 있음
  - Eysenbach(2011)은 기존에 영향력을 평가하는 척도로 사용되어 온 인용데이터와 트위터에 근거한 사회적 영향력과의 상관관계를 분석하는 연구를 수행한 결과, 소셜미디어를 통해 어느 정도의 인용 건수에 대한 예측은 가능하지만 전통적인 인용데이터를 보완하는 용도로 사용하는 것이 적당하다는 결론을 내림
  - Shema et al.(2014)는 출판 직후 블로그에 인용된 논문의 경우 그렇지 않은 경우보다 더 많은 인용을 받는다고 주장

- Liu(2014)는 학술연구의 관계자들의 관점에서 Altmetrics의 유용성을 제시
  - 사서의 경우, 연구자들이 자신의 논문을 향한 관심의 정도를 알 수 있도록 도와주고, 기관 리포지토리의 가치를 높일 수 있음
  - 연구지원기관의 경우, 부서별 출판물의 이용량, 영향력을 모니터할 수 있으며, 자금 지원 기관 및 정부의 위임 사항을 준수할 수 있음
  - 홍보기관의 경우, 기관의 성공 사례를 공유할 수 있고, 성과를 최대한 외부로 홍보 가능
  - 연구자의 경우, 자신의 영향력 지표를 이력서, 펀딩 지원서에 제시할 수 있고, 연구 결과를 어디에 출판할지에 대한 결정에 도움을 줄 수 있음

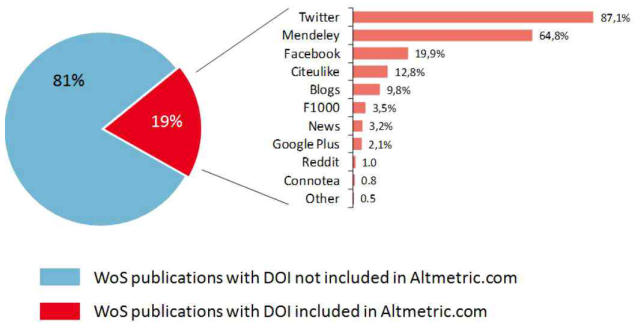
- Altmetrics의 단점(조재인 2014)
  - 소셜미디어가 지표가 되므로 온라인으로 접근이 용이한 연구 성과 외에는 높은 반응을 나타내기 어려움
  - 간혹 학술적 가치와 무관한 평가가 이루어질 수 있는 가능성이 있음

- 국가와 언어에 따라 측정 결과가 다르게 나타날 수 있음

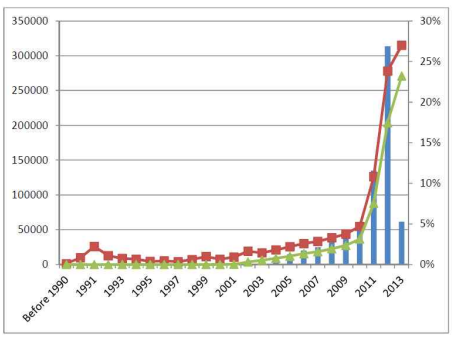
- Altmetrics 측정도구 간에도 결과가 다르게 나타날 수 있음

• Altmetrics 측정 도구로는 Altmetric.com, Impact Story, Plum X Metrics, PLOS ALM (Article-Level-Metrics) 등이 있으며, 모두 웹을 기반으로 하고 있음, 각 측정 도구별 정보원으로는 블로그, 페이스북, 트위터, Google+, F1000, Figshare, Slideshare, CiteULike, Mendeley[3] 등을 사용

- Altmetric.com
  · 2011년 Euan Adie에 의해 개발되었으며, 소셜 미디어에서 논문 기사 및 연구 데이터가 얼마나 주목받고 있는지를 그림으로 시각화하거나 점수로 산출하여 표시함
  · 저널 웹페이지나 기관 리포지터리에 삽입 가능한 API를 제공하고 있으며, 북마클렛, 탐색기, 뱃지 서비스를 제공
  · Altmetrics 개념이 도입되기 이전의 데이터는 거의 없어 커버리지가 너무 적다는 것이 큰 단점



<그림 5> Altmetric.com의 유형별 데이터 규모 (Robison-Garcia et al. 2014)



<그림 6> Altmetrics.com의 Coverage (Rodrigo Costas et al. 2014)

3) 특히 많은 연구에서 Mendeley와 피인용도와의 상관분석을 수행한 결과 Mendeley가 피인용도를 대체 또는 보완할 수 있는 도구임을 밝힘(Henning 2010 ; Priem et al. 2012 ; Li et al. 2012 ; Bar-Ilan 2012 ; Zahedi et al. 2013 ; Schlogl et al. 2013 ; Thelwall et al. 2013 ; Haustein et al. 2013)

- ImpactStory.org
  · 연구자들이 그들의 연구 성과에 대한 다양한 영향력을 탐색하고 공유할 수 있도록 하는 오픈 소스의 웹 기반 도구로, 프로필을 등록한 사용자에게만 제공됨
  · 다른 웹 사이트(Figshare, GitHub, Google Scholar, ORCID, Slideshare)의 저자 계정이나 논문 DOI, URL 등 온라인 식별자를 연동하여 프로파일을 만들고 이를 기반으로 점수를 산출함

- Plum X Metrics
  · Plum Analytics사가 개발한 프로그램으로 기관 차원의 구독에 주로 사용됨
  · 연구의 영향력을 이용량(보기, 다운로드 등), 수집(북마크, 즐겨찾기 등), 코멘트(댓글, 리뷰, 링크 등), 소셜미디어(좋아요, 공유, 트윗 등), 인용(PubMed, Scopus, 특허 등)의 5가지 기준으로 분류하여 이를 '원'으로 표시, 각 기준별 값에 따라 원의 크기가 다르며, 이를 통해 시각적으로 기준별 연구 영향력을 즉시 파악할 수 있음

- PLOS ALM(Article-Level-Metrics)
  · PLOS(Public Library of Science)에 논문 단위의 지표를 제공하는 도구
  · 이용량(보기 및 다운로드), 인용(PMC, CrossRef, Scopus, WoS), POLS(코멘트, 노트, ratings), 소셜 네트워크(CiteULike, Mendeley, 트위터, 페이스북), 블로그 및 미디어의 데이터를 사용함

• Altmetrics를 적용하고 있는 전자저널, 데이터베이스가 다수 존재하며, 현재 Altmetrics는 각계에서 많은 주목을 받고 있고, 사회 전반에 걸쳐 많은 영향을 미치고 있음

- ScienceDirect, Scopus, Wos, Wiley 등에 이미 적용되고 있으며, 다수의 기관 리포지토리에도 Altmetrics가 적용되고 있음

- 일본에서는 자국의 학술 논문을 대상으로 Altmetrics를 측정하는 'Ceek.jp

Altmetrics'가 추진되고 있으며, 일본의 Agriknowledge, CiNii, J-STAGE, JAIRO, NII ID(JARIO Cloud, 정보학광장), 일본국립국회도서관 디지털화 자료, 일부 기관리포지토리를 대상으로 CiteULike, 레퍼런스협동데이터베이스, Google+, 하테나북마크, 트위터, 위키피디아를 조사하고 리뷰를 수집하는 것으로 영향력을 측정하고 있음

2. Emerging Research Topic 탐지 모델의 연구 동향 분석

2.1 유망 기술 및 유망 연구 영역 탐지의 필요성

- 20세기 이후 과학기술지식의 양이 폭발적으로 증가하면서 현재 부상하고 있는 새로운 연구주제의 발견과 시간에 따른 연구 동향 파악, 연구 분야 간의 연관성을 파악하는데 어려움을 겪고 있으며, 국가적 차원에서도 국내외 과학기술의 연구 상황을 객관적으로 파악하고, 향후 발전 방향을 결정하기 위한 목표 설정에 상당한 어려움이 있음

- 또한 대부분의 연구자들도 유망한 새로운 연구 주제에 대한 관심이 있으나, 해당 주제가 새로운 분야의 시작인지 아닌지 판단하기 어려움. 따라서 새롭게 부상하고 있는 연구 주제를 발굴하여 경쟁력을 선점하기 위한 노력들이 현재 진행되고 있으며, 과학기술의 현황 파악 및 미래 예측을 위하여 국내외적으로 다양한 활동이 이루어지고 있음

- 이처럼 개인적 차원에서 뿐만 아니라 국가적 차원에서도 새로운 연구 주제의 동태(動態)와 발전을 모니터링 하는 것은 주요한 관심사가 됨

- 과거에는 주로 학문 영역 간 관계와 구조를 이해하기 위한 시도 또는 주로 국가 또는 기관, 연구자, 연구 영역의 연구 실적을 분석하여 생산성과 영향력을 측정하기 위한 목적으로 과학계량학이 많이 이루어졌으나, 최근에는 유망연구 또는 유망기술 분야를 탐색하기 위한 목적으로 많은 연구가 이루어지고 있음

- 따라서 현재 부상하는 새로운 연구 주제의 탐지와 관련하여 현재까지 진행된 연구의 동향을 파악하고자 함

2.2 유망 기술 및 유망 연구 영역 탐지 모델의 연구 동향 분석

- 과학 영역의 출현 또는 성장을 조사하기 위해 개발된 지표들은 Price 지수(Price 1970), 즉시성 지수(immediacy index)(Garfield and Small 1989), currency index(Small 2006)와 같은 지표들이 있음

- 처음에는 문헌빈도, 인용빈도 등의 양적 지표들을 사용한 연구는 유망한 새로운 연구주제를 파악하기 위한 것보다는 주로 국가 또는 기관, 연구자, 연구 영역의 생산성과 영향력을 측정하기 위한 목적으로 많이 사용되었으나, 일부 연구에서 발표되는 논문의 수와 인용데이터를 통해 특정 연구 분야의 과거와 현재 상황을 파악하고, 미래를 예측하고자 하는 시도가 있음

  - Hsieh and Chang(2009)은 생산 및 공정관리 분야(POM, Production and operations)의 연구 성과 측정을 위해 해당 분야의 20개 저널을 대상으로 IF, 논문수, h-index를 사용하여 연간 논문수, 연구 생산성, 가장 많은 연구를 수행한 연구자와 소속 기관 및 국가(지리적 분포)를 분석하였으며, 논문의 타이틀에 나타는 키워드 분석을 통해 연구 주제에 대한 분석을 시도

  - 안세정 등(2012)은 SCOPUS 데이터베이스에 수록된 2000년부터 2011년까지 발표된 그래핀 관련 논문으로 부터 키워드를 추출한 후 키워드 맵을 작성하여 해당 연구영역에서의 성장, 전이, 분화과정을 분석한 연구사례가 있음

- 초기에는 과학 커뮤니티에서 어떻게 아이디어가 확산되는지, 또한 과학의 영역이 시간이 지남에 따라 어떻게 발전해 가는지 연구하기 위해 양적, 수학적 모델을 사용하였으며, 유망 연구 영역의 탐지를 위해 양적 측면에서 수행된 연구들은 Chen(2002), Ball and Tunger(2006), Luis M. A. Bettencourt et al.(2008) 등의 연구가 있음

  - Chen(2002)은 주류영역의 지식과 잠재영역의 지식을 적합성(relevance)와

인용(citation)의 2가지 차원으로 구분하였으며, 높은 적합성과 높은 이용을 보이는 영역을 주류 영역으로, 높은 적합성과 낮은 인용을 보이는 영역을 잠재영역으로 구분하여, 해당 영역이 향후 주류 연구 분야가 될 가능성이 있음을 제시

- Ball and Tunger(2006)는 논문 수와 인용데이터 분석을 통해 특정 연구 분야의 과거와 현재 상황, 미래 트렌드를 예측하였으며, 구체적으로 특정 연구 영역을 다루는 논문의 수를 연도별로 측정하여 과거로부터의 발전 과정을, 현재의 피인용 횟수를 측정하여 현재의 반응을 파악하고, 논문수와 피인용 횟수의 변동폭을 통해 미래, 즉 앞으로의 발전가능성을 가늠하는 연구를 통해, 계량서지학을 통해 특정 연구 영역이 향후 어떻게 발전할 것인가에 대한 정보를 파악할 수 있어 미래의 연구 트렌드를 분석할 수 있다고 주장

- 이처럼 논문의 생산 건수 및 논문의 피인용 횟수는 해당 연구 영역에 대한 관심의 정도를 파악할 수 있을 뿐만 아니라 성장 역동성을 보여주는 지표로 분명히 활용 가능하지만 인용이 가지고 있는 문제점과 논문의 출판시점과 인용을 받기까지 걸리는 시간차로 인해 즉시성이 중요한 쟁점이라고 할 수 있는 새로운 유망 주제의 탐지에 있어서는 이들 지표의 사용이 적절하지 않음

- 따라서 다양한 분석 단위(저자, 학술지, 국가 등) 간의 상호 연간관계를 측정하는 관계형 지표(<표 1> 참조)들이 등장하기 시작했으며, 주로 고피인용논문을 추출하고 동시인용 또는 동시단어 분석을 사용하여 현재 연구자들의 관심이 집중되고 있는 중요한 연구 분야를 선정하는 방법으로 다수의 연구가 진행됨

- KISTI는 자체 연구개발한 Macro Monitoring 분석방법으로 2006년부터 유망기술을 선정하고 미래유망기술세미나를 통해 발표하고 있음. SCOPUS 데이터베이스의 고피인용 학술 논문(상위 0.1%)을 인용관계 유사도로 군집분석을 하고 핵심 연구영역을 도출하여 동일 연구영역에서 고피인용 논문으로 구성된 컴포넌트를 인용한 문헌집합을 추출함. 각 문

헌집합에서 최빈도 키워드를 추출하여 해당 기술의 급부상하는 영역을 확인함

<표 2> 계량서지적 기법 분류(Cobo, Lopez et al. 2011)

분석기법		분석단위	관계의 종류
서지결합	저자	저자의 모든 연구성과	저자의 연구성과 중 공통의 참고문헌
	문헌	문헌	문헌간 공통의 참고문헌
	저널	학술지의 모든 논문	학술지의 모든 논문 중 공통의 참고문헌
동시저자	저자	저자명	저자 동시 출현
	국가	소속 국가	국가 동시 출현
	기관	소속 기관	기관 동시 출현
동시인용	저자	저자의 참고문헌	동시인용 저자
	문헌	참고문헌	동시인용 문헌
	저널	학술지의 참고문헌	동시인용 저널
동시단어	-	타이틀, 초록, 본문에서 추출한 키워드	단어의 동시출현

- 이 중에서도 새로운 연구 주제의 파악을 위해 가장 많이 사용된 방법 중의 하나가 동시인용(co-citation) 및 동시단어(co-word)분석으로, 이를 통해 연구자들의 관심이 증가하는 주제나 학문의 진화를 감지할 수 있으며, 동시저자 분석을 통해서는 학제간 특성, 지리적 근접성 등을 분석할 수 있음

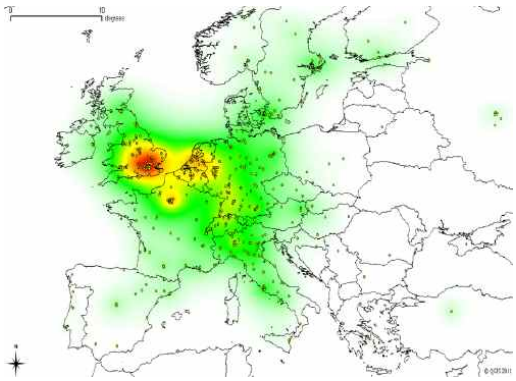
  - Callon(1991)이 연구한 co-word 분석 기법은 키워드를 연결강도에 따라 군집화한 뒤, 각 키워드를 군집내, 군집간 연결 정도에 따라 전략 다이어그램에 배치하고 각 키워드의 위상을 해석함. 주요 시구간을 설정하고, 각 시구간별 전략 다이어그램을 도출한 뒤, 해당 키워드의 위치변화 추적하는 방법임. 키워드의 위상변화에 대한 해석에 강점이 있으나 해당 키워드의 동태적 변화를 추적하는데는 한계도 있음

  - 정의섭 외(2008)는 건설교통관련 분야의 고피인용빈도 논문을 대상으로 국가별 논문 수 동향을 분석하였으며, 분야 간 융·복합관계 확인을 위해 상관관계 분석을, 서지결합분석을 통해 유망연구 영역을 도출하는 연

구를 수행

- Kajikawa et al.(2008)는 에너지 연구에 있어서 새로운 연구 영역을 추적하기 위해 인용네트워크분석을 수행하였으며, 연료, 태양전지의 연구가 빠른 속도로 성장하는 영역임을 나타냄

- 이러한 관계형 지표를 사용하는 경우 특정 연구 영역의 구조 및 진화를 모니터링하기 위해 이를 시각화하여 표현하는 기법들과 함께 사용되고 있음



<그림 7> 유럽지역 생화학, 유전학 및 분자생물학 저자들의 지리적 분포 (Bornmann and Waltman 2011)

<그림 8> 2010년 국내 바이오분야 논문등고선 (감주식 외 2012)

- 특히, 다양한 분석 단위(저자, 학술지, 국가 등) 간의 상호 연간관계를 측정하는 관계형 지표의 경우 노드와 링크로 구성된 네트워크로 가장 많이 표현되고 있으며, 공간적 분석 기법을 통해 특정 연구의 지리적 분포를 분석하여 특정 지역으로 수행된 연구 및 해당 연구가 주변 지역에 미치는 영향을 파악하는 연구도 진행됨

- 네트워크 분석의 경우, 네트워크를 통해 분야 간 상호연관성을 이해하고 이를 통해 미래의 유망분야를 예측하거나 제한하는 데 도움이 되며, 따라서 유망분야가 확산되는 경로의 예측 및 연구자들에게 연구의 방향성 제시, 그리고 연구과제의 방향과 목표가 올바르게 진행되고 있는지에 대한 판단 근거 제시 등의 장점이 있음(이우형 2005)

- Guimera & Amaral(2005)는 Callon의 co-word 분석에서 군집내/군집간 연결관계에 대한 구분을 확장하여, 네트워크 상의 위상을 세분화하고 위상변화를 추적하는 P-z 다이어그램과 주요 문헌 위상변화 추적 방법을 제시함

- 특허는 산업과의 높은 연관성이 있어 특허 분석을 통해 미래 유망기술 아이템을 발굴하려는 방법론이 연구되어지고 있음. 이재민 등(2011)은 검색을 통해 OLED 관련 특허를 조사하고 핵심 특허를 선별한 후 이들간의 인용 네트워크를 분석하여 OLED 분야의 기술 진화 경로를 분석함

• 동시단어분석을 사용한 연구들은 Ding et al.(2000), Lee(2008) 등이 있음

- Ding et al.(2000)의 경우 동시단어분석을 수행하면서 네트워크 분석 기법을 활용하여 분석하였으며, 동시단어 분석을 통해 새로운 영역의 출현을 감지할 수 있고, 지역별 동시단어분석을 통해서는 미처 발견하지 못했거나 주력하지 않았던 연구주제를 발견할 수 있다고 밝힘

- Lee(2008)는 정보 보안 분야를 대상으로 동시단어분석을 수행하여 해당 분야의 최신 연구 동향과 유망 주제를 추적

• 클러스터링 기법을 사용한 연구들은 Lamirel, J. C., Ta A. P., Attik, M.(2008), Shibata, N. et al.(2008), Takeda, Y., & Kajikawa, Y.(2009) 등이 있음

- Shibata, N. et al.(2008)는 과학 출판물의 인용 네트워크를 위상 클러스터링 방법을 사용하여 나타내고, 각 클러스터 내에 문헌의 위치를 추적함. 각 클러스터를 대표하는 특징이 있는 용어를 선정하여 클러스터의 진화를 모니터링 했으며, 유사한 주제는 강한 연결을 보이는 반면, 다른 주제를 다루는 논문은 약한 연결을 보여 지식 영역에서 강한 연결을 보이는 클러스터가 출현(emergence)을 탐지하는데 필요하다는 것을 밝힘

- Takeda, Y., & Kajikawa, Y.(2009)의 연구에서는 광학 분야에 있어 연구영역의 구조와 유망 연구 영역을 탐지하기 위해 학술문헌을 대상으로 인

용분석을 수행하였으며, 이를 위상적 클러스터링으로 군집화 함

- 텍스트 마이닝 방법을 활용하여 연구 영역을 클러스터링으로 표현한 경우 연구 트렌드 변화를 파악하기 위해 시계열적 분석을 함께 수행

  - Iwami, S., Mori, J., Sakata, I., & Kajikawa, Y.(2014)는 4개의 학술 도메인을 대상으로 인용 네트워크가 구축된 이후 시간 전이 분석(time transition analysis)을 통해, 각 문헌의 중심성을 계산, 그 변화를 추적하였고, 이를 통해 유망 선도 문헌을 탐지 할 수 있음을 주장

  - 감주식 외(2012)의 연구에서는 바이오 분야를 대상으로 텍스트마이닝 기법으로 논문등고선을 분석하여 바이오 분야의 중점 연구분야를 도출하였음

- Lucio-Arias and Leydesdorff(2007)은 과학적 발견으로부터 지식의 출현과 과학 커뮤니케이션 구조에서 그들의 영향에 대해 네트워크 분석을 수행하였으며, 저널, 단어, 인용의 관점에서 시각화하여 표현하였고, Leydesdorff and Schank(2008)는 새로운 영역이 출현하는 동안의 저널 인용 패턴의 변화를 살펴봄

- Jo, Y., Lagoze, C. & Giles, C. L.(2007)는 인용 그래프에서 용어 배분과 링크 배분 사이의 상관관계를 분석하였으며, 이를 통해 주제를 탐지하는 알고리즘을 제안하였음. 이 알고리즘은 단일 용어로 표현되는 주제어와 용어의 집합으로 표현되는 주제어의 순위 리스트를 생성하며, 이는 유행의 변화와 특수성을 가진 주제의 연속적 스펙트럼을 보여주어 이를 통해 주제 탐지에 있어 새로운 측면을 나타냄을 발견

- Kleinberg(2003)는 그의 연구에서 유망 주제의 영역을 식별하기 위한 'burst 탐지 알고리즘'을 제안하였고, Mane and Borner(2004), Chen(2006, 2009) 등의 연구에서 이를 적용하거나 응용함

- 현재는 유망 연구 주제의 탐지를 위해 2개 이상의 방법론을 결합한 형태의 연구들이 다수 존재함

  - Braam et al.(1991)과 Zitt & Bassecoulard(1994)는 2개 이상의 방법을 결합하여 사용하는 것을 처음으로 주장하였으며, 특히 텍스트 및 인용기반의 방법은 서로를 보완할 수 있는 방법으로 많이 사용됨(Glanzel and Thijs 2011a, b; Glanzel 2012)

  - 이외에도 서지결합분석과 단어빈도분석을 결합하여 사용한 연구(Glanzel and Thijs 2011b), 저자 및 텍스트 기반 접근 방법을 사용한 연구(Scharnhorst and Garfield 2010) 등이 있음

  - 특히, 최근에는 텍스트 기반 및 링크 기반 접근을 결합한 연구들이 이루어지고 있으며, 이를 통해 하이브리드 방법의 우수성을 보여주고 있음(Glenisson et al. 2005; Boyack, K. W., & Klavans, R. 2010; Liu, X. et al. 2010)

  - Roche et al.(2010)은 해당 분석 영역에서 키워드의 성격을 해당 영역 신규 등장, 해당 영역 특화성 확립, 다영역 횡단 등으로 구분하고, 각 키워드의 해당 특성이 시구간별로 변화하는지를 분석하는 Pathway-3 모델을 제시함

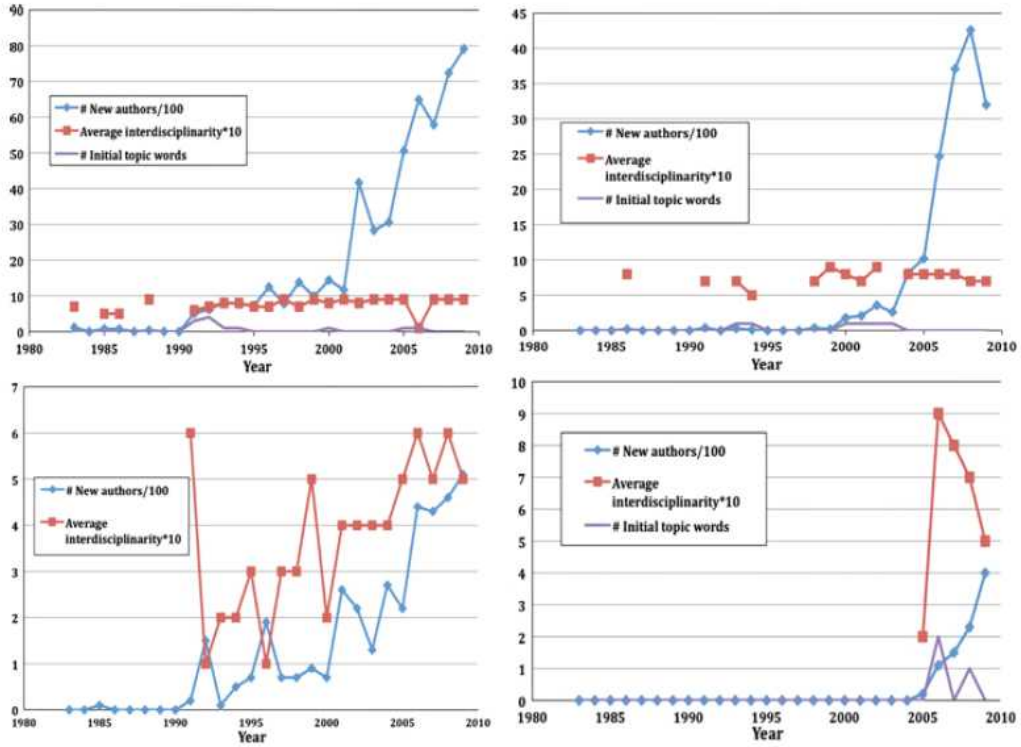
  - Cunningham & Kwakkel(2011)은 핵심 키워드 모니터링을 위한 프레임워크로 개별 키워드 자체의 증감 데이터에 기반해 성장곡선을 추정하고 각 키워드의 속도와 모멘텀 변수를 추정하고, 키워드가 갖는 위상/중요성을 구분하는 분석틀을 제시함

- Guo, H., Weingart, S., & Börner, K.(2011)는 (1) 어떤 단어가 광범위하게 사용되기 이전에 나타나는 'word burst'는 새로운 연구의 경향을 나타내고, (2) 유망 연구 영역은 새로운 저자를 유입시키며, (3) 학제적 참고문헌을 인용한다는 것을 전제로 하는 3개의 지표를 혼합하여 새로운 연구 분야의 주요 구조 및 기능을 예측하고자 하였음

  - 이를 위해, 특정 단어의 출현 빈도수의 급격한 증가, 유망 연구 영역으로

유입되는 새로운 저자의 수와 유입되는 속도, 인용한 참고문헌의 학제성 변화 양상을 살펴보았으며, 연구 결과 유망 연구 영역은 'word burst'가 선행하며, 새로운 저자를 빠른 속도로 유입하고, 학제적 성격의 참고문헌의 인용이 상당히 높은 것으로 나타남

- 또한 이 세 지표를 비교했을 때, 이들은 새로운 저자의 유입, 학제적 참고문헌의 인용 증가, word burst의 발생의 순으로 시간적 상관관계가 있는 것으로 나타나 해당 지표들이 유망 연구 영역을 식별하고 예측하는데 유용한 지표임을 제시함

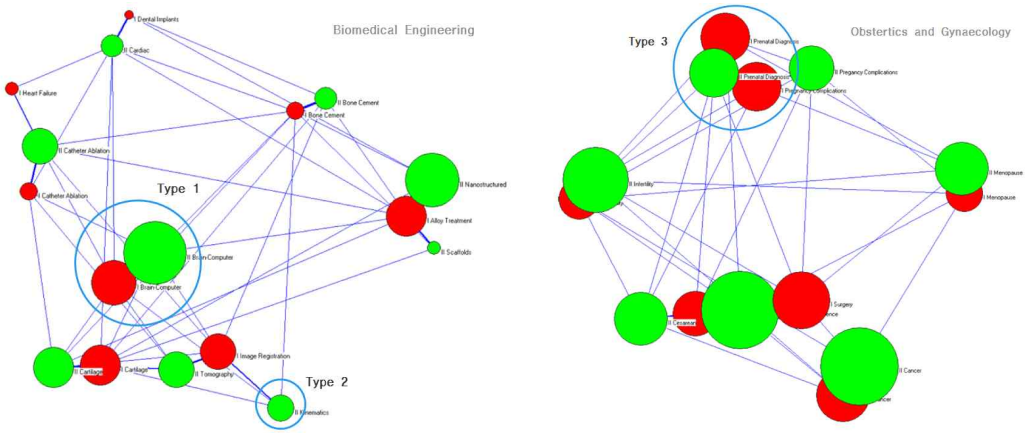


<그림 9> 세 지표 간 시간적 변화 추이 비교(Guo, H., Weingart, S., & Börner, K. 2011)

- Glanzel and Thijs(2012)는 새로운 연구 분야를 탐지하고 표현하기 위한 '핵심 문헌'의 사용에 대한 연구에서 하이브리드 접근방식을 사용하였으며, 클러스터링 방법을 통해 유망 연구 영역을 도출하고, 해당 영역에서 다른 문헌들과 강한 연결을 보이는 핵심 문헌을 선정

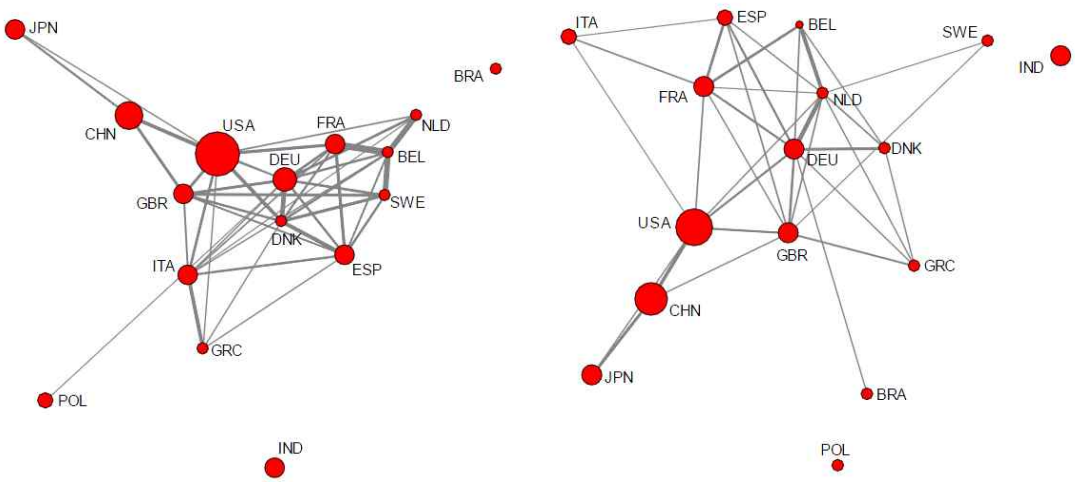
- 이들은 유망 연구 영역의 식별에 있어서 클러스터의 진화를 감지하여 새롭게 부상하는 주제를 나타내는 클러스터의 유형을 '(1) 예외적인 성장을

보이는 기존의 클러스터, (2) 다른 클러스터에 뿌리를 둔 완전히 새로운 클러스터, (3) 주제가 전환된 기존의 클러스터'의 3가지로 제시하면서, 선정한 4개 주제 분야에 대한 클러스터 분석을 수행하고 획득한 클러스터에서 핵심 문헌을 선정하였음



<그림 10> 유망 연구 영역의 식별을 위한 클러스터 유형(Glanzel and Thijs 2012)

- 유망 연구 영역은 클러스터링의 결과로 나온 키워드와 핵심 문헌으로 설명할 수 있으며, 이들 핵심 문헌들이 클러스터링의 결과로 나온 유망 영역에 대한 키워드를 더 보충한다는 것을 발견



<그림 11> Nano Pollution(좌)과 Biofuel(우) 분야의 출판활동 및 협력 현황(Glanzel and Thijs 2011a)

- 또한 유망 연구 영역의 핵심 문헌의 출판량이나 협력 현황 등의 다른 계량서지적 지표들을 활용하여 해당 영역에 대한 연구가 활발히 이루어지는 지역 또는 국가를 파악할 수 있음(Glanzel and Thijs 2011a)

- Scientometric이 시작된 초기에는 인용분석 등이 주로 방법론으로 사용되었으며, 방법론의 한계를 보완하기 위해 네트워크 분석을 결합한 연구가 많이 이루어져 오다 현재는 2개 이상의 방법론을 결합한 형태의 연구들이 다수 존재함

- 2개 이상의 지표나 방법을 사용한 연구의 경우에도 그 결과를 네트워크로 표현하는 경우가 대다수이며, 공간적 분석 기법을 사용하여 특정 지역을 중심으로 지리적 분포를 표현하거나 기타 다른 시각화 기법을 활용하기도 함

- 이외에 미래유망기술 탐색 모형 연구 및 서비스 몇 가지 사례를 참고로 제시함
  - 미국 SciTech Strategies는 과학기술 영역 매핑 및 신흥 기술 탐색과 관련한 실험적 프로젝트 중심으로 미래 유망기술을 발굴하여 선정함
  - 일본 NISTEP은 문헌 DB 계량분석 기반 핵심 연구영역을 탐색하는 Science Map 프로젝트를 수행함
  - 스페인 SCIMAGO는 Scopus DB 기반으로 저널/기관 랭킹 분석 기반 경쟁력을 평가함
  - 네덜란드 CWTS는 Web of Science DB 기반으로 대학 랭킹 분석, 과학기술 엑설런슨 매핑 연구을 수행함
  - 미국의 TechCast는 전문가 기반 델파이 등의 방법을 통해 미래기술을 선정하고 전문가 시각의 리포트를 발간함
  - 독일 Z-punkt는 Trenddatabase라는 자체 선별 트렌드에 대해 사회, 기술, 환경, 정치 등 차원의 분석 리포트를 제공함

- 영국 HSC는 시그마 스캔(Sigma Scan)을 통한 미래 이슈를 발굴·분석하고, 영국 정부의 미래전략과 정책적 의사결정의 방향을 제시함

3. 요약 및 시사점

- 과학기술지식의 양이 폭발적으로 증가하면서 현재 새로이 부상하는 연구 주제의 발견과 연구동향 파악, 연구 분야 간의 연관성 파악의 필요성이 점점 증대하고 있으며, 따라서 연구의 영향력, 학술적 우수성을 측정하고 평가하며, 나아가 발전 가능성이 높은 학문을 선별하는 것이 중요함

- 인용은 연구 영향력, 과학적 수월성을 측정하는 주요 지표로 여전히 널리 활용되고 있으며, 학문별, 연구 영역 별로 다른 인용의 특성을 반영하기 위한 정규화가 이루어져 왔으며, 더불어 h-index 등 새로운 인용기반 지표들이 등장하는 등 다양한 분석지표가 개발되었음

- 또한 기존의 인용 기반 지표들의 단점을 보완하기 위해 관계형 지표들이 등장했으며, 이는 네트워크 분석과 함께 발전함, 네트워크 분석 기법을 통해 연구 결과를 시각적으로 표현함을 통해 학술공동체의 구조와 연구의 동향에 대한 이해를 도울 수 있으며, 다른 다양한 시각화 기법과 함께 발전해 옴

- 이외에 이용량 기반 지표에 대한 연구와 새로운 학술 영향력을 측정하기 위한 Altmetrics 관련 연구가 진행되어 왔으며 이후 발전가능성이 있는 연구 영역을 탐지하고 추적하기 위한 도구로 영역을 확장해 가고 있음. 이러한 시도는 초기의 인용 지표를 도입하여 미래를 예측하고자하는 시도에서 나아가 현재는 클러스터링을 이용한 시계열 분석이나 2개 이상의 지표를 결합한 복합 지표를 사용한 연구들이 많이 수행되고 있음

- 국내의 경우, 학제적 성격이 강한 일부 분야에서 이루어졌으며, 인용분석을 기반으로 한 연구들이 대부분임

- 유망 연구 주제를 발굴하는 것은 학술적으로 우수한 연구의 조기 발굴 뿐 아니라 미래의 성과 예측, 발전 잠재성이 높은 연구의 선별 등에 방향

성을 제시할 수 있음

- 과학기술의 급부상성을 따라가기 위해서는 즉시성을 담보할 수 있는 이용량 지표를 함께 검토하는 것이 필요하며, 또한 최근 활발한 연구가 이루어지고 있는 Altmetrics 지표의 가능성을 검토하고 학술적 영향력 뿐 아니라 사회적 영향력을 평가할 수 있는 지표의 적용 가능성을 검토할 필요가 있음

참고문헌

Adler, R., Ewing, J., & Taylor, P. 2009. “Citation statistics.” Statistical Science, 24(1): 1.

Batista, Pablo D., Mônica G. Campiteli, and Osame Kinouchi. 2006. “Is it possible to compare researchers with different scientific interests?.” Scientometrics, 68(1): 179-189.

Bettencourt, L., Kaiser, D., Kaur, J., Castillo-Chavez, C., & Wojick, D. 2008. “Population modeling of the emergence and development of scientific fields.” Scientometrics, 75(3): 495-518.

Bollen, J., & Sompel, H. V. D. 2008. “Usage impact factor: the effects of sample characteristics on usage based impact metrics.” Journal of the American Society for Information Science and technology, 59(1): 136-149.

Bornmann, L. 2014. “Alternative metrics in scientometrics: A meta-analysis of research into three altmetrics.” Scientometrics, 103(3): 1123-1144.

Boyack, K. W., & Klavans, R. 2010. “Co-citation analysis, bibliographic coupling, and direct citation: Which citation approach represents the research front most accurately?” Journal of the American Society for Information Science and Technology, 61(12): 2389-2404.

Braam, R. R., Moed, H. F., & van Raan, A. F. J. (1991a). Mapping of science by combined cocitation and word analysis, part 1: Structural aspects. Journal of the American Society for Information Science, 42(4), 233-251.

Braam, R. R., Moed, H. F., & van Raan, A. F. J. (1991b). Mapping of science by combined cocitation and word analysis, part II: Dynamical aspects. Journal of the American Society for Information Science, 42(4), 252-266.

Braun, Tibor, Wolfgang Glänzel, and András Schubert. 2006. “A Hirsch-type index for journals.” Scientometrics 69(1): 169-173.

Brody, T., Harnad, S., & Carr, L. 2006. “Earlier web usage statistics as predictors of later citation impact.” Journal of the American Society for Information Science and Technology, 57(8): 1060-1072.

Chen, C. 2006. “CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature.” Journal of the American Society for information Science and Technology, 57(3): 359-377.

Cobo, M. J., López Herrera, A. G., Herrera Viedma, E., & Herrera, F. 2011. "Science mapping software tools: Review, analysis, and cooperative study among tools." Journal of the American Society for Information Science and Technology, 62(7): 1382-1402.

Ding, Y., Chowdhury, G. and Foo, S. 2000. "Journal as markers of intellectual space: Journal co-citation analysis of information retrieval area: 1987-1997." Scientometrics, 47(1): 55-73.

Duy, J., & Vaughan, L. 2006. "Can electronic journal usage data replace citation data as a measure of journal use? An empirical examination." The Journal of Academic Librarianship, 32(5): 512-517.

Egghe, Leo. 2006. "Theory and practise of the g-index." Scientometrics 69(1): 131-152.

Eysenbach, G. 2011. "Can tweets predict citations? Metrics of social impact based on Twitter and correlation with traditional metrics of scientific impact." Journal of medical Internet research, 13(4):e123

Garfield, Eugene. 1998. "Long-term vs. short-term impact: Part II. Cumulative impact factors." The Scientist 12(14): 12-13.

Garfield, Eugene. 1998. "Long-term vs. short-term journal impact: does it matter." Scientist 12(3): 11-12.

Glanzel, W. 2003. "Bibliometrics as a Research Field A course on theory and application of bibliometric indicators."

Glänzel, W. 2012. "The role of core documents in bibliometric network analysis and their relation with h-type indices." Scientometrics, 93(1): 113-123.

Glänzel, W., & Thijs, B. 2011a. "Research in emerging fields: who takes the lead?." ISSI Newsletter, 7(4): 85-95.

Glänzel, W., & Thijs, B. 2011b. "Using 'core documents' for the representation of clusters and topics." Scientometrics, 88(1): 297-309.

Glänzel, W., & Thijs, B. 2012. "Using 'core documents' for detecting and labelling new emerging topics." Scientometrics, 91(2): 399-416.

Glänzel, Wolfgang, and Henk F. Moed. 2002. "Journal impact measures in bibliometric research." Scientometrics 53(2): 171-193.

Glenisson, P., Gla¨nzel, W., Janssens, F., & de Moor, B. 2005. "Combining full text and bibliometric information in mapping scientific disciplines."

Information Processing & Management, 41(6): 1548-1572.

Guerrero Bote, Vicente P., Borja González-Pereira, and Félix de Moya-Anegón. 2009. "The SJR indicator: A new indicator of journals' scientific prestige." ArXiv.

Guerrero-Bote, V. P., & Moya-Anegón, F. (2014). "Relationship between Downloads and Citation and the influence of language." Scientometrics, 101(2): 1043-1065.

Guo, H., Weingart, S., & Börner, K. 2011. "Mixed-indicators model for identifying emerging research areas." Scientometrics, 89(1): 421-435.

Hammarfelt, B. 2014. "Using altmetrics for assessing research impact in the humanities." Scientometrics, 101(2): 1419-1430.

Haustein, S., Peters, I., Bar-Ilan, J., Priem, J., Shema, H., & Terliesner, J. 2013. "Coverage and adoption of altmetrics sources in the bibliometric community." In 14th International Society of Scientometrics and Informetrics Conference (pp. 1-12). Digital Libraries. Retrieved from http://arxiv.org/abs/1304.7300.

He, Chunpei, and Miranda Lee Pao. 1986. "A discipline-specific journal selection algorithm." Information Processing & Management 22(5): 405-416.

Henning, V. 2010. "The top 10 journal articles published in 2009 by readership on Mendeley." Retrieved from http://www.mendeley.com/blog/academic-features/the-top-10-journal-articles-published-in-2009-byreadership-on-mendeley/.

Hirsch, J. E. 2005. "An index to quantify an individual's scientific research output," Proceedings of the National Academy of Sciences, 102(46): 16569-16572.

Hirst, Graeme. 1978. "Discipline impact factors: A method for determining core journal lists." Journal of the American Society for Information Science 29(4): 171-172.

Hsieh, P. N., & Chang, P. L. 2009. "An assessment of world-wide research productivity in production and operations management." International Journal of Production Economics, 120(2): 540-551.

Iwami, S., Mori, J., Sakata, I., & Kajikawa, Y. 2014. "Detection method of emerging leading papers using time transition." Scientometrics, 101(2):

1515-1533.

Jo, Y., Lagoze, C. & Giles, C. L. 2007, “Detecting research topics via the correlation between graphs and texts.” In KDD-2007 Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining(pp. 370-379). San Jose.

Kajikawa, Y., Yoshikawa, J., Takeda, Y., & Matsushima, K. 2008. “Tracking emerging technologies in energy research: Toward a roadmap for sustainable energy.” Technological Forecasting and Social Change, 75(6): 771-782.

Kleinberg, J. 2003. “Bursty and hierarchical structure in streams.” Data Mining and Knowledge Discovery, 7(4): 373-397.

Lamirel, J. C., Ta A. P., Attik, M. (2008). Novel labeling strategies for hierarchical representation of multidimensional data analysis results. In A. Gammerman (Ed.), Proceedings of the 26th IASTED International Conference on Artificial Intelligence and Applications, Austria (pp. 169-174), 11-13 February 2008. Innsbruck, Austria: ACTA Press.

Lee, W. H. 2008. “How to identify emerging research fields using scientometrics: An example in the field of Information Security.” Scientometrics, 76(3): 503-525.

Lendi, S., & Huggett, S. 2012. “Usage: an alternative way to evaluate research.” Research Trends, 28.

Li, X., Thelwall, M., & Giustini, D. 2012. “Validating online reference managers for scholarly impact measurement.” Scientometrics, 91(2): 461-471.

Liu, Jean. 2014. Bringing Altmetrics to Institutions, 2014 ALA Annual Conference, <http://libraryconnect.elsevier.com/sites/default/files/Jean_Liu_June_2014.pdf> [cited 2015. 8. 28].

Liu, X., Yu, S., Janssens, F., Gla¨nzel, W., Moreau, Y., & De Moor, B. (2010). Weighted hybrid clustering by combining text mining and bibliometrics on large-scale journal database. Journal of the American Society for Information Science and Technology, 61(6), 1105-1119.

Liu, X., Yu, S., Janssens, F., Gla¨nzel, W., Moreau, Y., & De Moor, B. 2010. “Weighted hybrid clustering by combining text mining and bibliometrics on large-scale journal database.” Journal of the American Society for Information Science and Technology, 61(6): 1105-1119.

Lucio-Arias, D., & Leydesdorff, L. 2007. “Knowledge emergence in scientific communication: from “fullerenes” to “nanotubes”.“ Scientometrics, 70(3): 603-632.

McDonald, J. D. 2007. “Understanding journal usage: A statistical analysis of citation and use.” Journal of the American Society for Information Science and Technology, 58(1): 39-50.

Moed, H. F. 2005. “Statistical relationships between downloads and citations at the level of individual documents within a single journal.” Journal of the American Society for Information Science and Technology, 56(10): 1088-1097.

Moed, Henk F. 2010. “Measuring contextual citation impact of scientific journals.” Journal of Informetrics 4(3): 265-277.

Moed, Henk F., and Thed N. Van Leeuwen. 1995. “Improving the Accuracy of Institute for Scientific Information’s Journal Impact Factors.” JASIS 46(6): 461-467.

Mohammadi, E., & Thelwall, M. 2014. “Mendeley readership altmetrics for the social sciences and humanities: Research evaluation and knowledge flows.” Journal of the Association for Information Science and Technology, 65(8): 1627-1638.

Nisonger, Thomas E. 2004. “The benefits and drawbacks of impact factor for journal collection management in libraries.” The Serials Librarian. 47(1-2): 57-75.

Pesch, O. 2012. “Usage Factor for Journals: A New Measure for Scholarly Impact: Edited by Oliver Pesch.” The Serials Librarian, 63(3-4): 261-268.

Piwowar, H. 2013. “Altmetrics: Value all research products.” Nature, 493(7431): 159-159.

Piwowar, H. 2013. “Introduction altmetrics: What, why and where?.” Bulletin of the American Society for Information Science and Technology, 39(4): 8-9.

Priem, J., & Hemminger, B. H. 2010. “Scientometrics 2.0: New metrics of scholarly impact on the social Web.” First Monday, 15(7).

Priem, J., Piwowar, H., & Hemminger, B. H. 2012. “Altmetrics in the wild: Using social media to explore scholarly impact.” ArXiv: 1203.4745v1.

Rafael Ball, Dirk Tunger. 2006. “Bibliometric Analysis-A New Business Area for Information Professionals in Libraries? : Support for Scientific Research by

Perception and Trend Analysis." Scientometrics, 66(3): 561-577.

Rousseau, R. 2006. "A case study: Evolution of JASIS' h-index." Science Focus, 1(1): 16-17.

Schloegl, C., & Gorraiz, J. 2010. "Comparison of citation and usage indicators: the case of oncology journals." Scientometrics, 82(3): 567-580.

Shema, H., Bar Ilan, J., & Thelwall, M. 2014. "Do blog citations correlate with a higher number of future citations? Research blogs as a potential source for alternative metrics" Journal of the Association for Information Science and Technology, 65(5): 1018-1027.

Shepherd, P. T. 2011. "Journal Usage Factor-a promising new metric." Serials, 24(1): 64-68.

Shibata, N., Kajikawa, Y., Takeda, Y., & Matsushima, K. 2008. "Detecting emerging research fronts based on topological measures in citation networks of scientific publications." Technovation, 28(11): 758-775.

Shuai, X., Pepe, A., & Bollen., J. 2012. "How the scientific community reacts to newly submitted preprints: Article downloads Twitter mentions, and citations." Retrieved from http://arxiv.org/abs/1202.2461v1.

Takeda, Y., & Kajikawa, Y. 2009. "Optics: A bibliometric approach to detect emerging research domains and intellectual bases." Scientometrics, 78(3): 543-558.

Thelwall, M., Haustein, S., Larivie're, V., & Sugimoto, C. R. 2013. "Do altmetrics work? Twitter and ten other social web services." PLoS ONE, 8(5): e64841.

van Leeuwen, Thed N., and Henk F. Moed. 2002. "Development and application of journal impact measures in the Dutch science system." Scientometrics 53(2): 249-266.

Waltman, Ludo, et al. 2013. "Some modifications to the SNIP journal impact indicator." Journal of informetrics 7(2): 272-285.

Wouters, P., & Costas, R. 2012. "Users, Narcissism and Control: Tracking the Impact of Scholarly Publications in the 21st Century." Utrecht, Netherlands: SURF Foundation. research-acumen. eu/wp-content/uploads/Users-narcissism-and-control. pdf.

Zahedi, Z., Costas, R., & Wouters, P. 2014. "How well developed are altmetrics? A cross-disciplinary analysis of the presence of 'alternative metrics' in

scientific publications.” Scientometrics, 101(2): 1491-1513.
감주식, 김무웅, 박상대, 현병환. 2012 “바이오 미래유망 연구분야 도출에 관한 연구.” 기술혁신학회지, 15(2): 345-368.
김판준, 이재윤. 2010. “학술지 영향력 측정을 위한 h-지수의 응용에 관한 연구.” 정보관리학회지 27(1): 269-287.
이우형. 2005. “과학기술정보 분석을 통한 신기술 도출 방법론 및 지표 개발.” 과학기술정책연구원.
이재윤. 2006. “연구성과 측정을 위한 h-지수의 개량에 관한 연구.” 정보관리학회지 23(3): 167-186.
이준영 외. 2013. “고피인용 논문을 활용한 유망기술 발굴.” 한국전자통신학회 논문지 8(11): 1655-1664.
정의섭, 여운동, 조대연, 서일원. 2008. “과학계량학적 정보분석을 통한 건설교통분야의 유망연구영역도출.” 한국콘텐츠학회논문지, 8(2): 231-238.
조재인. 2015. “Altmetrics 를 통한 연구의 영향력 평가에 관한 연구.” 한국도서관·정보학회지, 46(1): 65-81.
한국과학기술정보연구원. 2013. “과학기술 이슈변화 탐지를 위한 키워드 경로 분석 방법론 개발”
한국과학기술정보연구원. 2014. “미래기술 탐색을 위한 정보분석체제 구축.”

Altmetric 홈페이지 http://www.altmetric.com/
PLOS Article-Level Metrics 홈페이지 http://article-level-metrics.plos.org/
Plum Analytics 홈페이지 http://plumanalytics.com/