

# 신기술 탐색을 위한 Bibliometric 분석 방법

## Bibliometric analysis for emerging technology exploring

이우형 · 손성혁 · 윤문섭

과학기술정책연구원 신기술분석연구센터

Woo Hyoung Lee · Sung Hyuk Son · Moon Seob Youn

Center for Emerging-Technology Assessment, STEPI

### Abstract

The aim of this study is to map the intellectual structure of the field of Organic Light Emitting Diode(OLED). Co-word analysis was employed to reveal patterns and trends in the OLED field by measuring the association strengths of terms representative of relevant publications or other texts produced in OLED field. Data were collected from INSPEC. Important keywords were extracted from author keywords. These author keywords were further standardized. In order to trace the dynamic changes of the OLED field, present the technology mapping. The results show that the OLED field has some established research theme and it also changes rapidly to embrace new themes.

### 1. 서론

#### 1.1 연구배경

오늘날, 지적이거나 정책적인 이유에 있어서, 과학분야나 사회과학분야에서의 개념, 아이디어, 그리고 문제들 간의 연계성을 도식화할 수 있도록 하는 것은 매우 중요하다. 이러한 도식화를 위해 몇 가지 방법이 시도되어졌다. 과학연구와 과학정책에서 사용되어진 전통적인 방법은, 상대적으로 소수 전문가들의 견해를 구하는 방법이였다(peer review)(Law & Whittaker, 1992). bibliometrics analysis은 정량적 측면에서 이러한 작업을 수행하기 위한 또 다른 방법이다. 또한, bibliometrics analysis은 사회적이고 정책적인 질문들에서 과학과 기술을 개발하고 지식의 발전을 목표로 하고 있다(van Raan, 1997). 저자나 저널의 co-citation analysis과 같은 전통적인 서지분석기술들은 과학논문에 내포되어있는 인용자료들에 대한 분석을 근간으로 이루어진다. 이러한 종류의 분석들은 흥미로운 결과를 가져

다주는 반면, 논문을 이용한 연구주제의 실제적인 내용에 대한 즉각적인 그림을 공급해 주지는 못한다. 주어진 주제 상에서 출간되어진 논문내에서 키워들의 동시출현을 측정하거나 분석하는 Co-word 분석은 이런류의 분석적 문제를 간결하게 처리할 수 있는 잠재력을 가지고 있다(Callon et al., 1991).

Co-word 분석방법은 데이터에 포함되어져 있는 기본적인 정보를 유지하면서 데이터를 독특한 시각적 방법으로 단순화시켜서 보여준다. 이것은 과학적인 개념이나 아이디어 그리고 지식 등의 중요한 제공자로서 자연어를 근간으로 한다(van Raan & Tijssen, 1993). 많은 연구자들이 다양한 분야에서 개념의 연결관계를 찾아내기 위한 중요한 방법으로 Co-word 분석기법을 사용하고 있는데, 예를 들면, Software engineering(Coulter et al., 1998), polymer chemistry (Callon et al., 1991), scientometrics (Courtial, 1994), 신경망 연구 (Noyons & van Raan, 1998a; van Raan & Tijssen, 1993), 생물학적 안전성 (Cambrosio et al., 1993), 산성화 연구(Law & Whittaker, 1992), 특허 (Courtial et al., 1993), 광학공학 (Noyons & van Raan, 1994), Bioelectronics (Hinze, 1994), 약학 (Rikken et al., 1995), 생물학 (Rip & Courtial, 1984; Looze & Lemarie, 1997), 농축물질물리학 (Bhattacharya & Basu, 1998) 등이 있다.

#### 1.2 연구목적

본 연구에서는 유기EL 분야의 새로운 연구동향 및 연구변화가 문헌의 서지적 현상속에 숨겨져 있다는 가설아래, 유기EL 분야의 지적구조 및 변화를 계량정보학적 기법인 Co-word 분석을 통하여 규명해 봄으로써 유기EL 분야의 발전방향과 학문적 위상을 정립해 보고자함에 그 목적이 있으며, 아울러 향후 유기EL 분야의 연구활동에 보충적 참고자료

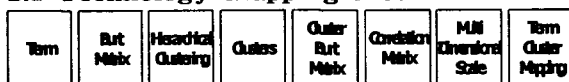
로 활용될 수 있는데 하는데 그 부차적 목적이 있다.

## 2. 연구방법

본 연구에서는 크게 두 가지 분석을 실시하였다. 첫째는 Co-word 분석 방법을 활용하여 해당 주제영역의 지적구조를 밝히는 연구 중<sup>1)</sup> Ding(2000), Rann & Velde(1991) 등의 연구에서 사용한 방법을 적용하여 유기EL 분야의 구조를 Technology Mapping 분석 방법을 통해 분석하였다. 분석대상 DB는 1969년에서 2003년까지 INSPEC DB에 저장된 유기EL 분야의 논문들을 조사하여 인용데이터를 수집하였다. 또한 분석을 위해 SPSS 11.0을 사용하여 Hierarchical Clustering(Ward Method), Correlation Coefficient Index, MDS(Multidimensional Scale) 방법을 통해 유기EL 분야의 지적구조를 분석하였다.

둘째는 각종 지표를 이용해 유기 EL 분야에 대한 심층 분석을 실시하였다. 즉, 혁신지표(Innovation Indicators), 기술기회분석(Technology Opportunity Analysis) 등을 통해 유기 EL 분야의 기술적 수준, 국내 유기 EL 분야의 포지셔닝 등을 살펴보았다. 이러한 분석방법은 GIT(Georgia Institute of Technology)의 TPAC(Technology Policy & Assessment Center)에서 개발한 것으로 특정 분야의 기술에 대한 기회를 분석하는데 유용한 도구이다.

### 2.1 Technology Mapping Method



<그림 1> 본 연구의 방법론

#### · Term

INSPEC DB를 이용하여 “유기 EL”과 관련 있는 Key-word를 추출하였다. 추출결과 797개의 Key-word를 추출하였다. 이들 Key-word 중 동시출현빈도수가 3개 이하인 것들을 제외하여 총 225개의 Key-word를 가지고 분석을 실시하였다.

#### · Burt Matrix

225개의 Key-word를 가지고 상관계수를 이용하여 225\*225 상관행렬을 만든다.

1) 지금까지 소개된 Co-word 분석을 이용한 연구방법은 다양하다. 그 중 대표적인 방법은 Ding et al.(1998), Noyons & Rann(1998), Rann & Velde(1991), Porter et al.(1999) 등이 주장한 방법이 있다.

#### · Hierarchical Clustering

클러스터링을 하기 위한 방법론으로 여기에는 대표적인 방법이 PCA(Principal Components Analysis)(TOA, 1999)방법과 계층적 클러스터링 방법 중 하나인 Ward 방법(Ding, 2001)이 있다. 본 연구에서는 Hierarchical Clustering을 이용하여 클러스터링 하였다. 분석을 위해서 계층적 클러스터링 방법 중 Ward Method를 이용하였다.

#### · Clusters

Hierarchical Clustering 분석 결과 20개의 클러스터를 완성하였다. 그리고 각 클러스터의 명칭<sup>2)</sup>은 전문가의 조언을 얻어 작성하였다.

#### · Cluster Burt Matrix

이렇게 구해진 앞서 20개의 클러스터를 가지고 각 클러스터에 속한 문헌들의 동시출현빈도를 구한다.

#### · Correlation Matrix

이 동시출현빈도수를 Pearson Correlation Coefficient를 이용한 20\*20 상관행렬을 만든다. 또한 최종 Term Cluster Mapping을 완성하여 각 클러스터간의 연결 강도를 표시하기 위해 Proximity Index(Rann & Velde, 1991)<sup>3)</sup>를 구한다.

#### · Multidimensional Scale & Term Cluster Mapping

이렇게 형성된 Pearson Correlation Coefficient의 20\*20 상관행렬을 MDS(Multidimensional Scale)을 사용하여 2차원으로 표현한다. 그리고 0.5 이상의 Proximity Index값을 가진 클러스터들을 MDS 지도에 선으로 표현한다.

2) 각 클러스터의 명칭을 부여하는 방법은 여러 가지가 있다. 즉, Ding et al.(1998)과 Van Rann(1991)은 클러스터내에 있는 Key-word의 빈도수가 가장 높은 것을 클러스터의 명칭으로 사용하였다. 그러나 본 연구에서는 먼저 가장 빈도수가 많은 것들을 명칭으로 부여한 후 전문가의 도움을 받아 적절한 단어로 수정 보완하였다.

3) Proximity Index는 근접지수로서, 각 term간에 유사도를 나타내는 지수이다.

$$P(ij) = \frac{C_{ij}}{C_i C_j} * N$$

Cij : 키워드 쌍(Mi, Mj)이 있는 문헌의 수,  
Ci : 기사세트에서의 키워드 Mi의 출현빈도  
Cj : 기사세트에서의 키워드 Mj의 출현빈도

### 3. 분석결과

앞서 제시한 방법론을 바탕으로 유기EL에 대한 Co-word 분석을 실시한 결과는 다음과 같다.

#### 3.1 Hierarchical Clustering

<표 1> Hierarchical Clustering 결과

	Cluster Name	키워드 수
1	organic compounds	8
2	display panel	8
3	application	7
4	semiconductors	6
5	integration and circuits	19
6	thin/thick film process	15
7	electronic structure	11
8	optical properties	13
9	carrier injection	9
10	stability	10
11	elements	23
12	semiconductor devices	14
13	luminescence	9
14	electrical properties	16
15	polymer	6
16	optical measurement	11
17	thick/thin film process	11
18	optical circuits	11
19	chemical reaction	8
20	thermal property measurement	10
	합계	225

주) 동시출현빈도수가 4개 이상인 것들을 대상으로 분석함

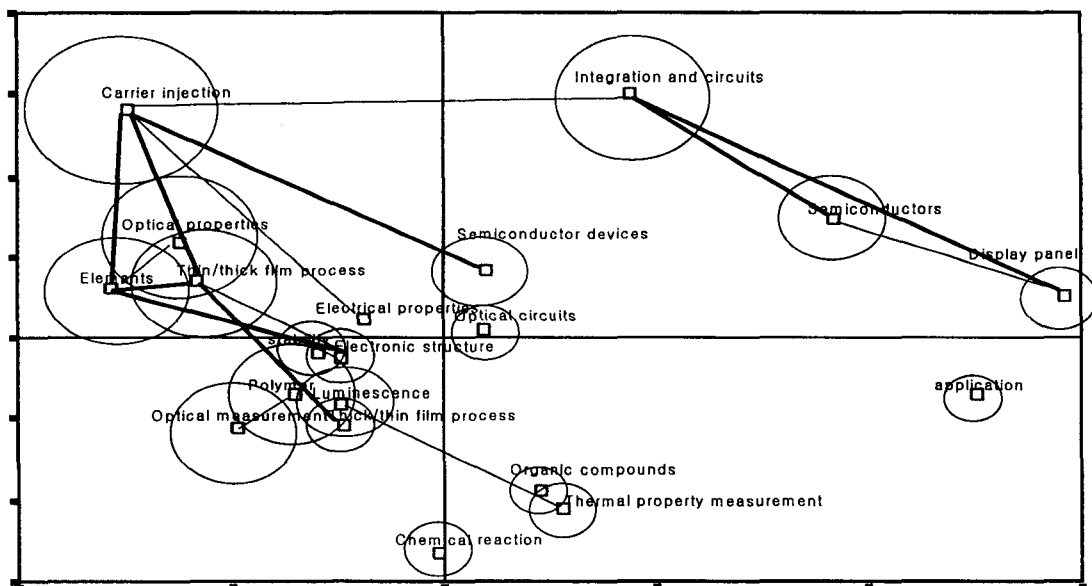
계층적 클러스터링 분석 결과 유기 EL의 Keyword를 20개의 요인으로 묶을 수 있음을 확인했다. 이러한 요인분석 결과를 가지고 Cluster Mapping을 실시하였다(<그림 2> 참조).

#### 4. 결론

본 연구는 같은 영역이나, 하위영역으로부터 얻어지고, 서로 다른 기간동안에 개략적 같은 양으로 나뉘어 질 수 있는 많은 양의 언어 자료들 내에서, 패턴이나 확인된 경향들을 알아내기 위해 실용 가능한 접근법으로서 Co-word 분석의 타당성을 보여준다. 그러므로, IR의 발생에서의 본 측정은 IR영역과 이와 관련된 하위영역으로의 연구가 나아갈 방향을 지적해 준다. 또한, Co-word 분석은 과거에 결과를 통한 논리적 설명시 시각화 방법의 중요성문제 때문에 충분히 이해되지 않는 키워드들의 내부관계와 IR의 하위 분야들을 성공적으로 시각화 시켜서 보여준다.

Co-word 분석은 과학과 정보의 시각화되어진 지도를 작성하는 새로운 방법을 제시해 준다. Co-word는 통계적인 결과보다 더 나은 결과를 제공해 준다. 우리는 IR 영역내의 검색자들에게 Co-word 지도의 시각적 효과를 제공하는 것을 목적으로 하였고, 이는 상당히 유용한 결과를 가져왔다.

전체적으로, 이러한 연구는 Co-word 분석에 대한 신뢰감을 증가시켜주었다. Law와 Whittaker(1992)가 지적한 것처럼 Co-word분석의 장점을 살펴봄으로 적절한 주장을 할 수 있을 것이다. 이것은 정확히 말해서 문장에 따



<그림 2> LOED 분야의 Technology Mapping

라 다를 수 있지만, 이러한 사실의 장점이 학문적인 문장의 미묘한 차이에 따른 민감도의 수준을 문제제기할 수 있다.

## References

- Callon, M., Courtial, J. P. & Laville, F., "Co-word analysis as tool for describing the network of interactions between basic and technological research : The case of polymer chemistry", *Scientometrics*, Vol.22, No.1, pp.153-203, 1991
- Bhattacharya, S. & Basu, R. K., "Mapping a research area at the micro level using co-word analysis", *Scientometrics*, Vol.43, No.3, pp.359-372, 1998
- Cambrosio, A., Limoges, C. Courtial, J. P. & Laville, F., "Historical scientometrics? Mapping over 70 years of biological safety research with co-word analysis", *Scientometrics*, Vol.27, No.2, pp.119-143, 1993
- Coulter, N., Monarch, I. & Konda, S., "Software engineering as seen through its research literature : A study in co-word analysis", *Journal of the American Society for Information Science*, Vol.49, No.13, pp.1206-1223, 1998
- Courtial, J. P., "A coword analysis of scientometrics", *Scientometrics*, Vol.31, No.3, pp.251-260, 1994
- Hinz, S., "Bibliographical cartography of an emerging interdisciplinary discipline : the case of bioelectronics", *Scientometrics*, Vol.29, No.3, pp.353-376, 1994
- Law, J. & Whittaker, J., "Mapping acidification research : A test of the co-word method", *Scientometrics*, Vol.23, No.3, pp.7-17, 1992
- Looze, M. D. & Lemarie. J., "Corpus relevance through co-word analysis : An application to plant proteins", *Scientometrics*, Vol.39, No.3, pp.267-280, 1997
- Noyons, E. C. M. & Van Rann, A. F. J., "Bibliometric cartography of scientific and technological development of an R&D field", *Scientometrics*, Vol.30, No.1, pp.157-173, 1994
- Noyons, E. C. M. & Van Rann, A. F. J., "Advanced mapping of science and technology", *Scientometrics*, Vol.41, No.1-2, pp.61-76, 1998
- Rikken, P., Kiers, H. A. L. & Vos, R., "Mapping the dynamics of adverse drug reactions in sunsequent time periods using INDSCAL", *Scientometrics*, Vol.33, No.3, pp.367-380, 1995
- Rip, A. & Courtial, J. P., "Co-word maps of biotechnology : An example of cognitive scientometrics", *Scientometrics*, Vol.6, pp.381-400, 1984
- Van Rann, A. F. J. & Tijssen, R. J. W., "The neural net of neural network research", *Scientometrics*, Vol.26, No.1, pp.169-192, 1993