

## 토픽모델링을 활용한 무역분야 연구동향 분석

이지훈  
세종대학교 경영학부 부교수

김정숙  
세종대학교 경제학과 조교수

# A Study on the Research Trends in Int'l Trade Using Topic modeling

Jee-Hoon Lee<sup>a</sup>, Jung-Suk Kim<sup>b</sup>

<sup>a</sup>School of Business, Sejong University, South Korea

<sup>b</sup>Department of Economics, Sejong University, South Korea

Received 23 May 2020, Revised 26 June 2020, Accepted 28 June 2020

### Abstract

This study examines the research trends and knowledge structure of international trade studies using topic modeling method, which is one of the main methodologies of text mining. We collected and analyzed English abstracts of 1,868 papers of three Korean major journals in the area of international trade from 2003 to 2019. We used the Latent Dirichlet Allocation(LDA), an unsupervised machine learning algorithm to extract the latent topics from the large quantity of research abstracts. 20 topics are identified without any prior human judgement. The topics reveal topographical maps of research in international trade and are representative and meaningful in the sense that most of them correspond to previously established sub-topics in trade studies. Then we conducted a regression analysis on the document-topic distributions generated by LDA to identify hot and cold topics. We discovered 2 hot topics(internationalization capacity and performance of export companies, economic effect of trade) and 2 cold topics(exchange rate and current account, trade finance). Trade studies are characterized as a interdisciplinary study of three agendas(i.e. international economy, International Business, trade practice), and 20 topics identified can be grouped into these 3 agendas. From the estimated results of the study, we find that the Korean government's active pursuit of FTA and consequent necessity of capacity building in Korean export firms lie behind the popularity of topic selection by the Korean researchers in the area of int'l trade.

**Keywords:** LDA, Research Trend, Trade Studies, Topic Modeling

**JEL Classifications:** F00, F10

<sup>a</sup> First Author, E-mail: [petra@sejong.ac.kr](mailto:petra@sejong.ac.kr)

<sup>b</sup> Corresponding Author, E-mail: [js\\_kim@sejong.ac.kr](mailto:js_kim@sejong.ac.kr)

© 2020 The Korea Trade Research Institute. All rights reserved.

## I. 서론

지식이 어느 정도의 성숙도에 도달하면 학자들이 기존 문헌 자체에 관심을 갖는 것이 일반적이다. 어떤 주제들이 기존 논문에서 다루졌고, 어떤 주제들이 인기가 있는지는 연구자에게 중요한 정보가 된다.

특히 무역학은 여러 학문의 융합, 통섭 학문 성격을 가지며, 학문으로서의 정체성에 대한 의문이 종종 제기되고 있어(Shim Chong-Seok, 2009) 연구의 지형도와 동향을 파악하는 작업은 각별히 중요한 의미를 갖는다. 무역학 분야에서 연구동향 분석이 계속 이뤄져온 이유이다.

그러나 기존 연구동향 분석에는 몇 가지 문제가 있다. 첫째, 연구 주제들을 주관적으로 분류하는 데서 비롯되는 선택 편향의 문제이다. 기존 연구들은 연구자들이 논문의 초록이나 주제어를 바탕으로 몇 가지 연구 범주와 주제들을 뽑아내고 그에 해당하는 논문의 수를 비교 분석하는 빈도 분석이 주류를 이룬다. 그러나 이는 연구자의 주관적 판단에 따라 분류 방법이 달라지고, 때로 분류에 오류가 생길 수도 있다.

둘째, 미리 결정된 연구 범주는 모든 연구 주제를 포괄하지 않을 수 있으며, 특히 연구자들이 새로운 분야의 연구에 대해 알지 못하거나 킨센서스를 형성하지 못할 때 그럴 수 있다. 무역학 관련 논문 중 한국연구재단의 학문 분류 체계 상 '달리 분류되지 않은 무역'으로 분류되는 논문이 많았던 것도(La Kong-Woo, 2011), 이런 문제에 기인한 것이라 볼 수 있다.

셋째, 하나의 논문은 두 가지 이상의 주제를 포함하는 경우가 많기 때문에 한 논문을 한 주제로 이름 붙이는 것은 적절하지 않다. (Lee Hak-Yeon, Seo Han-Bin and Geum Young-Jung, 2018)

넷째, 기존 수작업 방식으로는 방대한 양의 문헌 정보를 처리하는 데 지나치게 많은 시간과 노력이 요구된다는 점이다. 빅데이터 시대를 맞아 전자 도서관의 이용 가능성이 크게 높아져 학자들이 문헌 정보를 쉽게 얻을 수 있게 되었으나, 이는 학자들에게 정보의 과부하라는 새로운 딜레마를 안겨주었다. (Fang, et al., 2018)

본 연구에서 이용한 토픽모델링 접근법은 이런 문제들에 대한 유망한 해결책 중 하나이다.

이 접근법은 대규모의 비정형 문서 모음에 들어 있는 잠재적 주제들을 기계적으로 발견해내는 알고리즘이다. 문서에 사전적으로 라벨을 붙일 필요가 없으므로 인간의 주관적 판단과 독립적으로 분석이 이뤄질 수 있다. 각각의 논문은 여러 주제의 집합으로 간주되며, 알고리즘에 의해 그 확률분포가 제시된다. 따라서 기존의 블랙박스식 방법론과 비교했을 때 정확하고 심층적인 정보를 제공하며, 범주 분류의 정확성을 제고한다. 또한 과학적이고 반복 가능한 알고리즘이므로 이를 소프트웨어화해 대량의 정보를 매우 짧은 시간에 처리할 수 있다.

이러한 이점들로 인해 토픽모델링은 텍스트 마이닝의 주요 방법론 중 하나로 자리잡았으며, 다양한 학문 분야에서 주제를 식별하는 데 활발하게 이용되고 있다.

본 연구는 토픽모델링 분석 기법을 활용해 무역 및 국제통상 분야 연구의 연구동향과 지식구조를 파악하고자 하였다. 이를 위해 2003년부터 2019년까지 국내 무역학 주요 학술지인 무역학회지, 국제통상연구, Journal of Korea Trade에 수록된 논문 1868편의 초록 정보를 수집해 분석했다.

텍스트 마이닝을 비롯한 구조적 분석은 철저한 문헌 검색과 반복 가능하고 과학적이며 투명한 과정을 통해 선택 편향을 최소화한다는 점에서 기존의 서술적 내용 분석과 구별된다. (Vrabel, 2015) Kao and Poteet (2007)는 텍스트 마이닝을 비 구조화된 텍스트로부터 흥미롭고 일상적이지 않은 지식을 추출하는 것이라고 정의했다. 텍스트 마이닝을 활용하면 방대한 논문 정보를 가공되지 않은 상태로 분석해 연구 주제를 분류하고 주요 단어간 연관관계를 찾아내는 등 다양한 분석이 가능하다.

본 논문은 텍스트 마이닝 기법 중 문헌에 숨겨져 있는 주제들을 찾아내는 통계 추론 모델인 토픽모델링을 활용해 국내 무역학 연구 주제의 양상과 변화를 분석한다. 특히 잠재디리클레할당(Latent Dirichlet Allocation, LDA) 알고리즘을 적용해 분석한다.

무역학 분야 연구 동향 분석에서 양적 분석은 많지 않으며, 양적 분석의 경우 대부분 사회연결망 분석 기법을 이용했다. 토픽모델링 기

법을 활용한 양적 분석은 본 논문이 최초이다.

본 연구는 약 20년 간 무역 분야 연구의 주제와 동향을 식별하고 정량화하고 이해하기 위한 과학적 시도로서, 연구자와 저널 편집자, 저널 발행기관이 보다 정확한 정보에 입각해 판단을 내리는데 도움을 줄 것이다.

본 연구의 목표는 다음을 포함한다. (1)2003~2019년 기간 동안 국내 무역학 분야 연구 동향 및 특징 파악 (2)잠재적인 연구 주제 발굴 (3)무역학 분야 연구의 미래 방향 제시.

정량적 방법으로 학술, 출판물 자료를 분석한 연구를 정보계량학(scientometrics)이나 계량서지학(bibliometrics)라고 부른다. 본 연구도 그 연장선에 있다.

## II. 문헌 연구

### 1. 무역학 분야 연구 동향 분석 문헌 연구

그간 무역학 분야 연구동향 분석은 서술적 방법에 의한 질적 분석이 주를 이루었다. 예를 들어 Lee Chun-Su and Lee Jang-Rho (2002)는 국내 전자무역 논문들을 여러 사전적 기준으로 분류하고 빈도 분석을 통해 시사점을 도출했다. Yoon Hee-Young (2015)은 FTA에 관한 논문 340편을 연구자가 사전에 정한 여러 기준으로 분류하고 빈도 분석을 통해 연구 흐름과 방향을 분석했다

이밖에 FTA 시대에 경영 분야의 대응을 다룬 논문들을 분석한 Kim Moon-Hee (2017), 무역리스크 관리 분야 연구동향을 분석한 Jung Hong-Ju, Sung Soo-Nam and Choi Yoo-mi (2012), 중국 전자상거래 관련 연구동향을 분석한 Mi Liqun, Kim Wan-Min, Yun Kwang-Woon (2013), '통상정보연구'에 게재된 논문들의 연구 분야와 핵심 키워드, 저자 특성을 분석한 Lee Ho-Hyung (2017), '통상정보연구'에 게재된 논문 중 지역통상정보로 분류되는 논문들의 연구 동향을 분석한 Choi Won-Seok and Hong Seung-Lin (2018), 중소기업의 국제화 관련 연

구동향을 분석한 Jung Min-Ji, Yang Hee-Soon and Chung Jae-Eun (2014) 등이 주로 빈도 분석과 서술적 방법을 통해 연구동향을 분석했다.

무역학 분야 연구동향 분석 중에서 정보계량학적 접근 방법을 택한 양적 연구는 많지 않으며, 주로 사회연결망 분석 방법론을 채택하고 있다.

Kang Dong-Joon and Lee Kil-Nam (2015)은 사회연결망 분석을 활용해 1975~2015년 한국무역학회지에 수록된 논문 공저자 네트워크를 구성하고, 중심성 분석 등을 통해 그 속성과 시기별 변화를 분석했다. Kim Sung-Kuk and Pak Jee-Moon (2015)은 사회연결망 분석을 이용해 5개 무역 관련 학회지의 논문 공저자 네트워크를 구성하고 중심성 분석을 수행했다. Yang Kun-Woo (2017)는 사회연결망 분석 기법을 적용해 '통상정보연구' 학술지에 게재된 논문의 주제어 네트워크를 구성한 뒤 연결도와 매개중심성을 기준으로 중심 키워드를 파악하고, 클러스터링 분석을 통해 유사성이 있는 키워드들을 그룹화했다.

Jhang Se-Eun and Lee Su-Ho (2016)는 해운경제 국제학술지에서 수집된 논문들의 저자 키워드를 가지고 사회연결망 분석을 수행, 여러 중심성 지표를 도출하고 기간 별 변화를 관찰했다. Tian, Geng, and Sarkis (2018)은 국제무역에 체화된 흐름(embodied flows) 분야 논문들을 대상으로 동시 인용 및 사회연결망 분석을 수행했다.

### 2. 토픽모델링을 이용한 연구 동향 분석 문헌 연구

LDA가 발견한 주제는 인간 전문가들이 파악한 주제와 매우 부합하며, 이는 이 방법론의 타당성과 효과를 잘 보여준다.(Dunbar and Weber, 2014)

Sun and Yin (2017)은 1990~2015년 세계 운송 분야 연구에 LDA 모델을 적용해 50개의 주제를 추출했는데, 대부분 기존의 운송 연구 하위 분야와 대응했다. Park Ja-Hyun and Song Min (2013)은 LDA 알고리즘을 이용해 42년간

의 국내 문헌정보학 연구동향을 분석한 결과, 기계학습을 통해 자동 분류된 주제들을 문헌정보학 주제분류표 상의 주제와 대응시킬 수 있었다. 이들 연구는 사람의 주관적 판단 없이 논문 정보 자체만으로 전문가와 비슷한 수준의 주제 분류를 할 수 있음을 보여준다.

LDA를 사용하여 연구 주제를 발견한 뒤에는 문서(논문)에 대한 평균 할당(문서-주제 분포)이 시간이 지남에 따라 증가하고 있는지 감소하고 있는지를 조사함으로써 핫토픽과 콜드토픽을 식별해낼 수 있다(Griffiths and Steyvers, 2004).

Kim Chang-Sik, Choi Su-Jung and Kwahk Kee-Young (2017)은 이 방법으로 국내 정보시스템분야 연구에서 핫토픽과 콜드토픽을 구분했다. Na Sang-Tae, Kim-Ja-Hee, Jung Min-Ho and Ahn Joo-Eon (2016)는 시뮬레이션 관련 국내 학술 논문 데이터로부터 분야별 주제를 도출하고 기간 세분화를 통해 그 변화를 파악했다.

### Ⅲ. 연구 방법, 데이터 수집 및 처리

#### 1. 연구 방법

토픽모델링은 텍스트 데이터를 자동적으로 분석해 문서 내의 숨어 있는 의미구조를 파악하는 기계학습 기반의 텍스트마이닝 기법이다. 문서 집합을 토픽(주제)들의 확률적 분포(혼합체)로 표현하고, 각 토픽은 단어들의 확률적 분포(혼합체)로 표현함으로써 문헌의 구조를 파악한다.

토픽모델링은 ‘감독되지 않은(unsupervised)’ 기계학습으로도 불리는데, 인간에 의해 사전 분류된 태그나 훈련 데이터가 필요하지 않기 때문이다. 관찰된 변수인 단어를 통해 문헌의 주제와 같은 보이지 않는 변수를 추론하며, 결과적으로 전체 문헌 집합의 주제들과 각 문헌이 각 주제에 포함될 확률, 각 단어가 각 주제에 포함될 확률들을 알아낼 수 있다(Park Ja-Hyun and Song Min, 2013). 문헌에 숨겨져

있는 주제들을 찾아내는데 유용해 최근 학술 동향 분석의 기법으로 널리 활용되고 있다.

토픽모델링의 여러 알고리즘 중 잠재디리클레 할당(Latent Dirichlet Allocation, LDA)은 대표적인 확률론적 토픽모델로 널리 쓰이고 있다. 이 모델은 Blei, Ng, and Jordan (2003)이 고안한 것으로, 주제를 자동으로 찾는 토픽 모델링 과정에 디리클레 통계 분포를 접목한 것이다. LDA의 가정을, 본고의 분석 대상인 논문과 연관시켜 표현하면 다음과 같다.

1. 각 논문은 여러 개의 주제를 포함한다.
2. 각 주제는 여러 개의 단어를 포함한다.
3. 각 논문에 사용된 단어 하나 하나는 어떤 주제에 포함된다.

LDA는 또 연구자가 문헌을 작성하는 과정이 다음과 같은 생성 모델에 의해 이뤄진다고 가정한다.

1. 논문에 어떤 주제들이 들어갈지를 미리 정한다. (논문의 주제 분포 결정)
2. 주제들 중에서 하나를 고른다. (주제 선정)
3. 해당 주제를 구성하는 단어들 중에서 하나를 고른다. (단어 선정)
4. 그 단어를 논문에 입력한다.

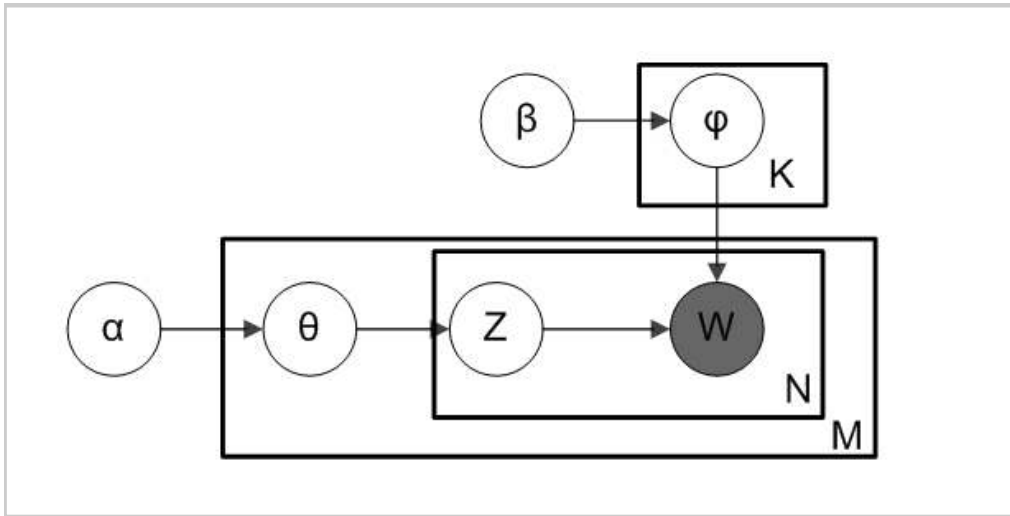
즉 연구자는 논문을 쓰기 전에 1을 결정하며, 논문에 한 단어 한 단어 입력할 때마다 2~4의 과정을 반복한다.

그런데 우리가 실제로 관찰할 수 있는 것은 논문의 주제 분포와 주제별 단어 분포가 아니라 논문에 들어 있는 단어들뿐이다. 논문들에 어떤 주제가 포함되어 있는지를 알기 위해서는 관찰된 단어들로부터 논문별 주제 분포와 주제별 단어 분포를 추론하는 통계적 작업이 필요하다.

LDA는 정확한 확률 분포를 모르는 상황에서 분석 대상 논문에 포함된 단어들을 관측해 가면서 확률 분포를 개선해 나간다. 이를 베이지스 추론 과정이라 한다.

주제가 k가지가 있을 때 이중 하나의 주제를 고르는 행위는 다항분포에 해당한다. 어떤 주

Fig. 1. Document Generation Process of LDA



Source: Blei, Ng, and Jordan (2003)

제에 포함되는 단어가  $v$ 개가 있다고 할 때  $v$ 개의 단어 중 하나를 고르는 행위 역시 다항분포이다. 우리가 관측하는 단어는 다항분포의 결과가 두 번 중첩된 것이다.

베이즈 추론 과정은 어떤 사건이 발생할 확률을 가정할 뒤 새로운 관측을 할 때마다 그 사건이 발생할 확률을 더 정확하게 추론해 나간다. 앞의 확률을 사전 확률, 뒤의 확률을 사후 확률이라고 한다.

수학자들은 특정 관측 사건에 대해 어떤 사전 확률 분포를 쓰면 계산이 매우 간편해진다는 사실을 발견해 정리했다. 이 특정 사건과 짝을 이루는 사전 확률 분포를 켈레 사전 분포 (conjugate prior)라고 하는데, 다항 분포의 켈레 사전 분포는 디리클레 분포이다.

LDA 즉 잠재디리클레할당이란 말에서 ‘잠재’는 관찰된 단어들로부터 숨어 있는, 즉 잠재적인 분포를 찾아낸다는 의미이다. ‘디리클레’는 디리클레 분포에서 나온 말이며, ‘할당’은 각 단어를 결정할 때 다항 분포로 주제를 할당하고 그 주제에서 단어를 할당한다는 의미이다.

Blei, Ng, and Jordan (2003)은 LDA 모형을 (Fig. 1)과 같이 표현했다. 문서가 어떻게 생성되는지, 그리고 주요 모수들이 문서 생성에 어

떻게 관여하는지 과정을 요약한 그림이다.

회색으로 칠해진  $W$ 는 우리가 관측할 수 있는 단어를 뜻하고, 상자 모양으로 둘러싸인 것은 그것이 반복된다는 것을 의미한다.  $M$ 은 전체 문헌의 개수(우리의 경우 분석 대상 논문 1898개)를 나타내고,  $K$ 는 주제의 개수,  $N$ 은 특정 문헌(우리의 경우 분석 대상 논문 중 하나)에 속한 단어의 개수를 나타낸다.

$\theta$ 는 문헌의 주제 분포를 나타낸다. 예를 들어 본 연구의 분석 대상 논문들을 3개의 주제로 분류했다고 하자(예를 들어 국제경제학, 국제경영학, 무역실무.) 각 논문은 이 세 가지 주제의 집합으로 표현할 수 있다. 이를테면 논문 1은 국제경제학과 관련된 내용이 50%를 차지하고, 국제경영학은 30%, 무역실무는 20%를 구성한다. 논문 1의 주제 분포를 벡터로 표현하면 다음과 같다.

논문 1의 주제 분포: (0.5, 0.3, 0.2)

반면 논문 2는 국제경제학 관련 내용이 10%, 국제경영학이 80%, 무역실무가 10%를 구성한다. 논문 2의 주제 분포는 다음과 같이 표현된다.

논문 2의 주제 분포: (0.1, 0.8, 0.1)

하나의 논문은 두 가지 이상의 주제를 포함하는 경우가 많다. 따라서 한 논문을 한 주제로 이름 붙이는 것은 적절하지 않으며, 이처럼 여러 주제들의 분포로 파악하는 것은 보다 현실에 적합한 가정이라 할 수 있다.

주제 분포는 문헌(논문)마다 다들테니 문헌의 개수 M개만큼의 주제 분포가 있을 것이다. 본 연구의 경우 분석 대상 논문의 개수 1898개만큼의 문헌-주제 분포가 있을 것이다.

$\varphi$ 는 주제의 단어 분포를 나타낸다. 즉 각 주제는 전체 문헌에 포함된 모든 단어와 그에 대응하는 가능도의 집합으로 표현할 수 있다. 우리의 분석 대상인 모든 논문에 출현하는 단어의 수가 10개 밖에 안 된다고 가정하자. 예를 들어 FTA, 관세, 정부, 정책, 중재, 신용장, 원산지 규정, 브랜드, 중소기업, 소비자라고 하자.

FTA, 관세, 정부, 정책이 주제 1 즉 국제경제학 주제에 포함될 확률이 크고, 다른 단어들은 확률이 낮다면 주제 1의 단어 분포는 다음과 같이 표현할 수 있을 것이다.

주제 1의 단어 분포: (0.3, 0.2, 0.2, 0.05, 0.01, 0.02, 0.05, 0.09, 0.08)

각 숫자는 10개의 단어 각각이 주제 1에 포함될 확률을 나타낸다. 첫 번째~네 번째 단어 즉 FTA, 관세, 정부, 정책의 확률이 높은 것으로 표현돼 있다. 한편 여기서 10개 단어에 해당하는 확률을 다 합하면 1이 된다.

주제 3, 즉 무역실무 주제의 단어 분포는 5~7번째 단어, 즉 중재, 신용장, 원산지 규정의 확률이 높은 다음과 같은 벡터로 표현할 수 있을 것이다.

주제 3의 단어 분포: (0.04, 0.01, 0.02, 0.05, 0.3, 0.2, 0.2, 0.09, 0.04, 0.05)

단어 분포는 주제마다 다들테니 주제의 개수 K개만큼의 단어 분포가 생긴다. 즉 앞서 든 예처럼 주제가 3개라면 3개의 주제-단어 분포가 생긴다.

$\alpha$ 와  $\beta$ 는 초매개변수(hyper parameter)이다. 모델에서 사용자가 직접 세팅해 주는 값으로, 전체 말뭉치(corpus, 본 연구의 경우 분석 대상인 무역 관련 논문 전체)에 하나가 세팅된다. LDA는  $\theta$ 와  $\varphi$ 가 초매개변수  $\alpha$ 와  $\beta$ 를 따르는 디리클레 분포라고 가정한다. 한편  $z$ 는 해당 단어가 속하는 주제의 번호를 나타낸다.

LDA 분석을 통해 우리가 얻고자 하는 값은  $z$ 와  $\theta$ ,  $\varphi$ 이다. LDA는 실제 문헌 내의 단어들, 즉  $W$ 를 관측해 나가면서 각 단어마다 임의의 주제를 차례로 부여한다.(즉 임의의  $z$  값을 정한다.) 그리고 이 결과에 따라  $\theta$ ,  $\varphi$ 의 디리클레 분포를 업데이트한다. 이런 과정을 반복해 가능한 모든 경우의  $z$ 값 중 가장 가능도가 높은  $z$ 값을 찾아내고,  $\theta$ ,  $\varphi$ 를 추정한다. 반복적인 시뮬레이션 과정이라 할 수 있다.

LDA를 실시할 때 사용자는 사전에  $\alpha$ 와  $\beta$ , 주제의 개수  $K$ , 그리고 시뮬레이션 횟수를 선택해야 한다.  $\alpha$ 와  $\beta$ 를 어떻게 선택하느냐에 따라 LDA의 분석 결과가 달라진다.  $\alpha$ 가 작으면 하나의 문서가 특정 하나의 주제로 대부분 채워지는 형태의 분포가 되고,  $\alpha$ 가 크면 하나의 문서가 다양한 주제로 구성되게 된다. 또  $\beta$ 가 클수록 주제들 간의 유사성이 높게 되고,  $\beta$ 가 작을수록 주제들이 서로 명확히 구분된다(Kwak Min-Ho, Min Hye-ree and Kim Mee-reem, 2019). 적절한  $\alpha$ 와  $\beta$  값에 대해서는 선행 연구마다 결론이 다르다.

LDA 분석은 연구자가 정한 개수만큼 주제를 분류해 내지만, 주제명을 정해주지는 않는다. 주제명은 연구자가 LDA 분석 결과 나타난 주제별 단어 분포와 각 주제로 분류된 문서들을 보면서 직접 결정해야 한다.

LDA를 통해 발견한 잠재적 주제들은 단어 자체의 분포 이외에는 어떤 정보도 사용하지 않고 완전히 감독되지 않은 방식으로 발견된다. 이는 주제의 범주들이 논문 저자가 사용하는 단어의 수준에서 추상적인 내용의 실질적 차이를 반영하고 있음을 암시한다. 이 결과는 해당 분야 연구의 영역 혹은 분야의 분류 체계로 사용될 수 있는데, 그러한 잠재적 주제들이 일반적으로 기존에 학계에서 일반적으로 받아들여지는 연구 영역 분류 체계와 매우 부합하

**Table 1.** Number of Papers in Each Journal

Year	Journal of Korea Trade	Journal of Int' l Trade and Industry Studies	Korea Trade Review	Total
2003	17	26	80	123
2004	16	22	72	110
2005	13	18	57	88
2006	28	34	69	131
2007	24	28	86	138
2008	17	26	69	112
2009	21	25	76	122
2010	14	23	69	106
2011	16	24	75	115
2012	18	30	77	125
2013	9	18	83	110
2014	16	20	70	106
2015	22	21	64	107
2016	22	18	49	89
2017	21	17	66	104
2018	22	19	56	97
2019	-	18	97	115
Total	296	387	1,215	1,898

기 때문이다(Griffiths and Steyvers, 2004).

## 2. 데이터 수집 및 처리

본 연구는 무역학회지, 국제통상연구, Journal of Korea Trade의 3개 국내 학술지를 분석 대상으로 선정했다. 분석 기간은 3개 학술지의 영문 초록과 주제어 정보가 모두 입수 가능한 2003년부터 2019년까지로 정했다. 영문으로 발행되는 Journal of Korea Trade와의 일관성을 유지하기 위해 3개 학술지의 서지 정보 중 영어로 제공된 것을 분석 대상으로 했다.

사회연결망 분석 프로그램인 Netminer의 확장 프로그램인 Biblio Data Collector를 이용해 기초 데이터를 수집하였다. 최종적으로 수집된 논문은 1,898편이며, 선정된 학술지 별, 연도별로 발표된 논문의 편수는 <Table 1>과 같다.

토픽모델링은 이들 논문의 제목, 초록, 주제어에 포함된 모든 단어를 활용해 분석하였다. 특히 초록은 정보(단어)의 양 측면에서 압도적이다. 초록은 논문 전체의 압축된 표현이고 일반적으로 연구 주제에 대한 핵심 단어들을 충분히 포함하고 있기 때문에 논문의 대용물로 쓸 수 있다(Griffiths and Steyvers, 2004).

토픽모델링 분석에 활용된 단어 수는 총 11,774개이다. 분석에 앞서 단어들을 정제하는 작업을 거쳤다. 대문자와 소문자가 서로 다르게 표시된다든지 하이픈(-)이나 가운데점(·)이 포함돼 서로 다른 단어로 인식되는 경우에도 분석에 영향을 주지 않는 범위 내에서 표준화해 일관성을 확보할 수 있게 했다. 정제 과정을 거친 후 최종적으로 분석 대상으로 추출된 단어 수는 8,624개이다.

**Table 2.** An Example of Ex-post Paper-topic distribution

1 <sup>st</sup> Topic	1 <sup>st</sup> Prob	2 <sup>nd</sup> Topic	2 <sup>nd</sup> Prob	3 <sup>rd</sup> Topic	3 <sup>rd</sup> Prob	...	20 <sup>th</sup> Topic	20 <sup>th</sup> Prob
Topic 10	0.418	Topic 7	0.127	Topic 1	0.073	...	Topic 15	0.015

**Table 3.** An Example of Ex-post Topic-word Distribution

1 <sup>st</sup> Keyword	1 <sup>st</sup> Prob	2 <sup>nd</sup> Keyword	2 <sup>nd</sup> Prob	3 <sup>rd</sup> Keyword	3 <sup>rd</sup> Prob	...	8624 <sup>th</sup> Keyword	8624 <sup>th</sup> Prob
Strategy	0.038	Subsidiary	0.037	Knowledge	0.035	...	1990s	0.000

## IV. 분석 결과

### 1. 주제의 발견

토픽모델링의 여러 알고리즘 중 잠재디리클레할당(Latent Dirichlet Allocation, LDA)은 가장 대표적인 확률론적 토픽모델로 널리 쓰이고 있으며, 본 논문도 이를 활용해 분석했다. LDA 알고리즘은 크게 블록 기반 변분 알고리즘과 깃스 표집의 둘로 나눌 수 있다. 본 연구는 깃스 표집을 선택했다. 본 연구에 활용한 소프트웨어는 Netminer 4이다.

토픽 수는 10, 15, 20, 25, 30개로 바꿔 가며 LDA 분석을 실시해 가장 의미 있는 결과가 나온 20개로 결정했다. 시물레이션 횟수는 1000 회로 정했다.  $\alpha$ 와  $\beta$  역시 연구자가 정해야 한다. 본 연구는 최적  $\alpha$ 가 0.01~0.1이고,  $\beta$ 는 0.01임을 보인 Zhao, Chen, and Zen(2015)의 연구에 입각해  $\alpha$ 는 0.01,  $\beta$ 는 0.01로 정했다.

LDA 모델을 실행하면 두 가지 유형의 사후적 확률분포를 얻을 수 있다. 각 논문의 주제 분포와 각 주제의 단어 분포이다.

예를 들어 'A Study on the factors determining the performance of Korean subsidiaries in China'란 논문에 대해 LDA 분석은 <Table 2>와 같은 확률분포를 추정해 내놓았다.

이 논문은 주제 10과 관련된 내용이 0.418(41.8%)을 차지해 가장 많고, 주제 7 관련이 12.7%로 그 다음으로 많다. 주제 15 관련은

1.5%에 불과해 가장 적다. (확률을 다 더하면 1이 된다.) 본 연구는 이 경우 이 논문을 주제 10으로 분류했다. 즉 가장 확률이 높은 주제 (<Table 2>에서 1st Topic)를 해당 논문의 주제로 분류했다. LDA는 이같은 확률 분포를 논문의 개수(1898개) 만큼 추정한다.

이번에는 각 주제의 단어 분포를 살펴보자. 예를 들어 주제 10에 대해 LDA 분석은 <Table 3>와 같은 확률분포를 추정해 내놓는다.

즉 주제 10에 'strategy'란 단어가 포함될 확률은 0.038(3.8%)로 가장 높고, 그 다음이 'subsidiary'로 3.7%이다. '1990s'란 단어가 포함될 확률은 0으로 가장 낮다. LDA는 각각의 주제에 대해 단어의 개수(8624개) 모두에 대해 확률을 계산한다. 이를 다 더하면 1이 된다. 한편 LDA는 이같은 확률 분포를 주제의 개수(20개) 만큼 추정한다.

본 연구는 이같은 방식으로 무역분야 20개의 주제를 추출하였으며 그 결과가 <Table 4>에 요약돼 있다.

LDA는 논문들을 비슷한 주제들끼리 묶어줄 뿐 주제의 이름을 알려주지는 않으며, 주제명은 연구자가 정해야 한다. 본 연구의 공동 연구자들은 각 주제 별 상위 단어들과 각 주제로 분류된 논문 중 확률이 높은 것들의 내용을 분석하면서 토론을 통해 주제명을 결정했다.

<Table 4>에는 각 주제의 상위 단어 10개와 해당 주제에 포함된 논문의 개수도 포함돼 있다.

본 연구를 통해 도출된 20개의 주제들은 무



역 및 국제통상 분야에서 기존에 확립된 연구 분야와 매우 잘 대응하는 것으로 나타났다. 예를 들어 주제 1은 주로 공급사슬 관리, 전략적 제휴와 관련이 있고, 주제 2는 국제물품매매계약(CISG), 입증 책임 등 국제상사계약과 관련된 이슈를 주로 다루고 있다.

하나의 주제가 여러 주제를 포괄하는 경우도 있다. 주제 5는 전자상거래, 무역보험, 리스크 관리에 관한 내용을 모두 포함하고 있다. 주제 8은 국가 이미지와 한국 제품 마케팅, 브랜드 관련 내용이 포함돼 있다. 주제 16은 FTA 원산지 규정 관련 논문들이 적합도가 가장 높게 나타나며, CGE 모형 등을 이용해 FTA의 경제적 효과를 분석한 논문들이 뒤를 잇는다.

무역학은 학제간 연구의 성격이 강하며, 크게 국제경제, 국제경영, 무역상무의 집합으로 볼 수 있는데, 본 연구에서 도출된 주제들도 이에 대응한다. 즉 <Table 5>에서 주제 3, 7, 9, 11, 13, 14, 15, 16, 18은 국제경제에, 주제 1, 4, 6, 8, 10, 17, 19는 국제경영에, 주제 2, 5, 12, 20은 무역상무에 대응한다. 무역상무 분야 주제가 4개로 상대적으로 적은 것은, 본 연구의 분석 대상 저널 3개 중에 무역상무를 중점적으로 다루는 저널이 포함되지 않은 것과 관련이 있는 것으로 보인다.

각 주제는 무역학이나 유관 학과를 개설한 대학의 강의 커리큘럼과도 잘 대응한다. 예를 들어 공급사슬관리(주제 1), 다국적 기업 전략과 지식 이전(주제 10)은 국제경영전략 과목에서, 국가 이미지와 한국 제품 마케팅/브랜드(주제 8)은 국제마케팅 과목에서 다루어지는 내용이다. 또 무역 분쟁과 국제상사 중재(주제 12) 상사중재론 과목에 대응한다. <Table 5>에 각 주제에 대응하는 대분류와 대학 교과목을 표시했다.

## 2. 트렌드 분석

LDA 분석 결과를 토대로 시간이 갈수록 활발하게 연구되는 주제(핫토픽)와 점차 연구되지 않는 주제(콜드토픽)를 구분할 수 있다. 이는 이 분석의 가장 매력적인 응용 중 하나이다 (Griffiths and Steyvers, 2004).

본 연구는 각 연도를 독립변수, 연도별로 각 주제의 논문별 평균 점유율을 종속변수로 하여 선형 회귀 분석을 실시하여 계수가 유의한 양의 값이면 핫토픽, 음의 값이면 콜드토픽으로 분류했다.

논문별 평균 점유율을 구하는 과정은 다음과 같다. 예를 들어 2019년에 발표된 논문이 A와 B 2개이고, 주제는 3개이며, 각 논문의 주제 분포가 아래와 같다고 하자.

A논문: (주제1=0.3, 주제2=0.5, 주제3=0.2)

B논문: (주제1=0.1, 주제2=0.4, 주제3=0.5)

이때 2019년에 주제1의 논문별 평균점유율은  $(0.3+0.1)/2=0.2$ 가 된다. 또 주제2의 평균점유율은  $(0.5+0.4)/2=0.45$ 이다. 이때 분모 2는 논문의 개수다.

t년도에 주제 j의 논문별 평균점유율을  $\theta_{j,t}$ 라고 하고 할 때 우리는 회귀식 (1)을 추정했다.

$$\theta_{j,t} = a_j + b_{jt} + \epsilon_j \quad (1)$$

우리가 관심 있는 것은 계수  $b_{jt}$ 의 부호이다. 이 값이 양수(음수)라면 전체 논문 집합에서 해당 주제의 비중이 상승(하락)함을 의미해 핫토픽(콜드토픽)으로 분류했다. 그 결과를 <Table 6>에 정리했다. 분석 결과, 5% 유의수준에서 핫토픽과 콜드토픽이 각각 2개씩 도출됐다. 주제 6(수출기업의 국제화 역량과 성과)과 주제 9(무역의 경제적 효과)가 핫토픽, 주제 14(환율과 경상수지)와 주제 20(무역 금융)이 콜드토픽으로 분류됐다.

10% 유의수준에서는 주제 3(국제경제협력과 통합)과 주제 13(무역구조와 국제경쟁력), 주제 16(FTA 원산지 규정)이 핫토픽, 주제 5(전자상거래와 무역 리스크 관리)와 주제 10(다국적 기업 전략과 지식이전), 주제 15(통상정책)가 콜드토픽에 각각 추가됐다.

국제경제, 국제경영, 무역상무로 구분한 대분류 기준으로는 5% 유의수준에서 무역상무가 콜드토픽으로 분류됐다.

**Table 4.** Identified topics from LDA

No.	Topic	Top 10 Stemmed Words	No. of Papers
1	Supply Chain Management	factor company logistics study business performance research supply system chain	91
2	Int' l Commercial Contract	contract good party law CISG term rule study buyer case	115
3	East Asia Int' l Economic Cooperation	trade paper country model Korea good purpose Koreas region author	75
4	Foreign Direct Investment	firm FDI investment effect country determinant paper result market productivity	155
5	E-commerce and Risk Management	trade risk system study management company business Korea transaction trading	89
6	Internationalization Capacity and Performance of Export Companies	performance firm study SME effect export market innovation relationship capability	108
7	International Trade Issues	China Korea country Japan market import US study tariff product	70
8	National Image and Korean Product Marketing/Brand	consumer product study brand quality effect service intention image exporter	99
9	Economic Effect of Trade	export effect country trade result analysis model datum study import	126
10	Strategy and Knowledge Transfer of MNCs	strategy subsidiary knowledge internationalization study firm performance relationship market MNC	97
11	International Development Cooperation	service trade development sector program paper study ODA cooperation education	73
12	Trade Dispute and Arbitration	arbitration dispute law case party award court agreement rule system	68
13	Trade Structure and International Competitiveness	industry export competitiveness Korea product change sector manufacturing index production	100
14	Exchange Rate and Current Account	exchange rate rate market stock crisis volatility model price result effect	143
15	Trade Policy	policy negotiation government paper market subsidy food trade incentive industry	67
16	Rules of Origin in FTAs	FTA Korea origin effect trade EU tariff agreement country rule	105
17	Organizational Management of MNCs	group study effect performance company commitment relationship difference distance employee	55
18	WTO and Multilateral Trade Norms	measure WTO system trade regulation US custom protection agreement law	101
19	Int' l Logistics	port study container efficiency oil transport shipping cargo energy Busan	78
20	Trade Finance	bank credit payment document letter bill transaction study guarantee lading	83

**Table 5.** Topic Classification and Corresponding University Curriculum

No.	Topic	Classification	Corresponding University Course
1	Supply Chain Management	Int'l Business	International Logistics Management
2	Int' l Commercial Contract	Int' l Trade Practice	Contracts of International Trade
3	East Asia Int' l Economic Cooperation	Int' l Economics	Studies on the Asia-Pacific Economy
4	Foreign Direct Investment	Int'l Business	Theory of Foreign Direct Investment
5	E-commerce and Risk Management	Int' l Trade Practice	Global Electronic Commerce, Marine Insurance Practice
6	Internationalization Capacity and Performance of Export Companies	Int'l Business	International Business
7	International Trade Issues	Int' l Economics	Theory of Int' l Commerce
8	National Image and Korean Product Marketing/Brand	Int'l Business	International Marketing
9	Economic Effect of Trade	Int' l Economics	International Trade Theory
10	Strategy and Knowledge Transfer of MNCs	Int'l Business	International Business
11	International Development Cooperation	Int' l Economics	International Economic Organization
12	Trade Dispute and Arbitration	Int' l Trade Practice	International Commercial Arbitration
13	Trade Structure and International Competitiveness	Int' l Economics	International Trade Theory
14	Exchange Rate and Current Account	Int' l Economics	Foreign Exchange and Finance
15	Trade Policy	Int' l Economics	Trade Policy
16	Rules of Origin in FTAs	Int' l Economics	Theory of Int' l Trade Practice
17	Organizational Management of MNCs	Int'l Business	International Business
18	WTO and Multilateral Trade Norms	Int' l Economics	International Economic Organization
19	Int' l Logistics	Int'l Business	International Transportation
20	Trade Finance	Int' l Trade Practice	International Trade Payment Systems

**Table 6. Slope of the Percentage Trend of Topics**

(Unit: Slope(×1000))

No.	Topic	Hot(p=0.05)	Cold(p=0.05)	Neutral(p=0.05)
1	Supply Chain Management			-0.3903
2	Int' l Commercial Contract			-0.4771
3	East Asia Int' l Economic Cooperation			1.6099 *
4	Foreign Direct Investment			0.6120
5	E-commerce and Risk Management			-1.2473 *
6	Internationalization Capacity and Performance of Export Companies	1.9351 **		
7	International Trade Issues			0.0888
8	National Image and Korean Product Marketing/Brand			0.9532
9	Economic Effect of Trade	2.2889 ***		
10	Strategy and Knowledge Transfer of MNCs			-1.5078 *
11	International Development Cooperation			0.3781
12	Trade Dispute and Arbitration			-1.0128
13	Trade Structure and International Competitiveness			1.2443 *
14	Exchange Rate and Current Account		-2.6984 ***	
15	Trade Policy			-0.6976 *
16	Rules of Origin in FTAs			0.8960 *
17	Organizational Management of MNCs			0.5778
18	WTO and Multilateral Trade Norms			-1.0034
19	Int' l Logistics			-0.3567
20	Trade Finance		-1.1924 **	
	International Economics			2.1064
	International Business			1.8232
	International Trade Practice		-3.9296 **	

Notes: 1. \*p&lt;0.1, \*\*p&lt;0.05, \*\*\*p&lt;0.01

2. Slope is multiplied by 1000

## V. 결론

여러 학문의 융합, 통섭 학문 성격을 갖는 무역학 분야에서 연구의 지형도와 동향을 파악하는 작업은 각별히 중요한 의미를 갖는다.

무역학 분야 연구 동향을 분석한 선행 연구들은 질적 분석이 주류를 이루었으나, 본 연구는 서지 정보 빅데이터에 입각한 양적 분석이며, 특히 인간의 사전적, 주관적 판단 없이 데이터 자체만으로 무역 분야 연구들의 잠재적 주제들을 도출하고 그 흐름을 분석했다는 점에 의미가 있다.

본 연구는 기계학습 기술 진보의 바탕 위에서 토픽모델링의 대표적 방법론인 LDA를 사용하여 3개 주요 무역 저널의 연구 주제와 그 진화 동향을 파악했다. LDA를 통해 20개의 주제를 추출하였는 바, 각 주제는 기존에 무역 분야에 널리 통용되는 연구 분야 및 대학 커리큘럼과 밀접하게 대응하였다.

또한 문서별 주제 분포 행렬 정보의 연도별 변화를 회귀분석으로 파악해 유의수준 0.05 수준에서 2개의 핫토픽과 2개의 콜드토픽을 추출했다. 주제 6(수출기업의 국제화 역량과 성과)과 주제 9(무역의 경제적 효과)가 핫토픽으로 선정되었는데 이는 2003년 이후 한국 정부가 추진해 온 자유무역협정과 이에 따른 기업의 국제화 역량 강화 사업의 필요성이 연구자들의 관심을 끈 것으로 해석된다. 반면 전통적인 국제경제 이슈, 예를 들어 주제 14(환율과 경상수지)와 주제 20(무역 금융)은 통상 이슈에 밀려 콜드토픽으로 분류됐다.

10% 유의수준에서는 주제 3(국제경제협력과 통합)과 주제 13(무역구조와 국제경쟁력), 주제 16(FTA 원산지 규정)이 핫토픽으로 선정되었는데 이 또한 자유무역협정에 이어 각종 메가-FTA, 예를 들어 TPP, RCEP, CJK FTA, CPTPP 등에 대한 연구자들의 관심을 반영하고 있다고 볼 수 있다. 국제경제, 국제경영, 무역상무로 구분한 대분류 기준으로는 5% 유의수준에서

무역상무가 콜드토픽으로 분류됐다. 이는 무역상무에 대한 관심이 줄어든 것이라기 보다는 FTA와 우리 수출기업의 국제화 역량 향상에 대한 관심을 반영하는 것으로 판단된다.

본 연구는 연구자들이 무역 분야의 현재 연구 상황과 미래 동향을 파악하는데 도움을 주어 어느 분야에 집중해야 하는지에 대한 지침을 제공한다. LDA는 사전 지식이 필요 없는 무감독 기계학습 알고리즘이다. 감독되지 않는다는 LDA의 특성은 인간의 개입 없이 데이터 그 자체에서 통찰력 있는 지식을 추출하는 데 매우 적합하다는 것을 의미한다. 이 논문은 LDA를 이용하여 서지 정보를 분석하는 흐름을 쉽게 설명함으로써 무역 이외의 분야 연구 주제 분석에 대해서도 참고자료를 제공한다.

LDA 알고리즘은 빅데이터와 기계학습의 획기적 진보의 산물로서 본고에서 연구한 문서 모형화 외에 유전자 분석, 화상 분류, 자동 화성 분석 등 다양한 분야에 활발히 응용되고 있다. 예를 들어 유전자 정보 분석 시 LDA는 문서 대신에 유전자를 사용하고, 단어 대신에 유전자 특징을 사용하고, 주제 대신 유전자 기능을 사용한다. 무역학 분야에서도 상상력을 발휘한다면 연구 동향 분석 외에도 다양하게 응용될 수 있을 것이다.

하지만 본 연구에는 두 가지 명백한 한계가 있다. 첫째 LDA 모델 자체의 한계이다. LDA는 문서가 서로 독립적이라고 가정하지만, 선행 논문이 후행 논문들에 큰 영향을 미칠 수 있다. 또 LDA는 주제를 유추할 때 문서 내 단어의 위치는 고려하지 않으나, 위치 정보는 문서 내용에 큰 영향을 미칠 수 있다. 둘째 본 연구는 많은 무역 저널 중 국내 무역 저널 3개만 분석 대상으로 삼았다는 한계가 있다. 앞으로 해외 무역 저널들을 분석해 본 연구와 비교한다면 한국의 연구 동향과 세계적 동향의 유사점과 차이점 등 의미 있는 시사점을 도출할 수 있을 것이다.

## References

- Blei, D. M., A. Y. Ng and M. I. Jordan (2003), "Latent Dirichlet Allocation", *Journal of machine Learning research*, 3(Jan), 993-1022.
- Choi, Won-Seok and Seung-Lin Hong (2018), "Analysis and Suggestion on Research Trend of Regional Commerce Information", *International Commerce and Information Review*, 20(2), 23-40.
- Dunbar, A. E. and D. P. Weber (2014), "What influences Accounting Research? A Citations-based Analysis", *Issues in Accounting Education*, 29(1), 1-60. <https://doi.org/10.2308/iace-50603>
- Griffiths, T. L. and M. Steyvers (2004), "Finding Scientific Topics", *Proceedings of the National academy of Sciences*, 101(suppl 1), 5228-5235.
- Fang, D., H. Yang, B. Gao, and X. Li (2018). "Discovering Research Topics from Library Electronic References using Latent Dirichlet Allocation, *Library Hi Tech*.
- Jhang, Se-Eun and Su-Ho Lee (2016), "A Study of Themes and Trends in Research of Global Maritime Economics through Keyword Network Analysis", *Journal of Korea Port Economic Association*, 32(1), 79-95.
- Jung, Hong-Joo, Soo-Nam Sung and Yoo-Mi Choi (2012), "Research Trend and Future Research Area on Risk Management in Foreign Trade", *Korea Trade Review*, 37(3), 69-109.
- Jung, Min-Ji, Hee-Soon Yang and Jae-Eun Chung (2014), "Analysis of Research Streams on Korean SMEs' Internationalization - Based on the Review of KCI Listed Journals", *Korea Trade Review*, 39(3), 145-183.
- Kang, Dong-Joon and Kil-Nam Lee (2015), "A Study on Co-author Networks of Journal of Korea Trade Research Association using Social Network Analysis", *Korea Trade Review*, 40(5), 1-23.
- Kao, A. and S. R. Poteet (2007), *Natural Language Processing and Text Mining*, MVY, MA: Springer Science & Business Media.
- Kim, Chang-Sik, Su-Jung Choi and Kee-Young Kwahk (2017), "Investigation of Research Trends in Information Systems Domain Using Topic Modeling and Time Series Regression Analysis", *Journal of Digital Contents Society*, 18(6), 1143-1150. <https://doi.org/10.9728/dcs.2017.18.6.1143>
- Kim, Moon-Hee (2017), "The Meta-analysis of Correspondence to Management Areas in the Era of FTAs: Focused on Research Papers from 2000 to 2016", *Korean Corporation Management Review*, 71, 117-132.
- Kim, Sung-Kuk and Jee-Moon Pak (2015), "A Study on the Centrality of Co-author Social Network in Korea Trade Research Community", *The Korean Research Institute of International Commerce and Law*, 67, 233-253.
- Kwak, Min-Ho, Hye-ree Min and Mee-reem Kim (2019), "Analysis of Students Open-Ended Course Evaluation Using Topic Modeling", *Asian Journal of Education*, 20(2), 491-522.
- La, Kong-Woo (2011), "A Study for redefining Trade Training Courses and Academic Classification". *Korea Trade Review*, 36(1), 249-273.
- Lee, Chun-Su and Jang-Rho Lee (2002), "A Classification and Analysis of Korean Internet Trade Research Papers-Focused on Korean graduation Theses and Papers" *International Commerce and Information Review*, 4(1), 149-172.
- Lee, Hak-Yeon, Han-Bin Seo and Young-Jung Geum (2018), "Uncovering the Topic Landscape of Product-service System Research: From Sustainability to Value Creation", *Sustainability*, 10(4),

911.

- Lee, Ho-Hyung (2017), "Meta Analysis of International Commerce and Information Review", *International Commerce and Information Review*, 19(2), 259-273.
- Liquan Mi, Wan-Min Kim and Kwang-Woon Yun (2013), "A Study on Research Directions and Implications of E-Commerce in China", *Journal of Korea Research Association of International Commerce*, 13(1), 45-67.
- Na, Sang-Tae, Ja-Hee Kim, Min-Ho Jung and Joo-Eon Ahn (2016), "Trend Analysis using Topic Modeling for Simulation Studies", *Journal of the Korea Society for Simulation*, 25(3), 107-116. 10.9709/JKSS.2016.25.3.107
- Park, Ja-Hyun and Min Song (2013), "A Study on the Research Trends in Library & Information Science in Korea using Topic Modeling", *Journal of the Korean Society for Information Management*, 30(1), 7-32. <https://doi.org/10.3743/KOSIM.2013.30.1.007>
- Shim, Chong-Seok (2009), "A Critical Study on the Reformation of Identity and Researching Field for the Foreign Trade's", *Management Education Research*, 54, 1-24.
- Sun, L. and Y. Yin. (2017), "Discovering Themes and Trends in Transportation Research using Topic Modeling", *Transportation Research Part C: Emerging Technologies*, 77, 49-66. <https://doi.org/10.1016/j.trc.2017.01.013>
- Tian, X and Y. Geng and J. Sarkis and S. Zhong (2018), "Trends and Features of Embodied Flows associated with International Trade based on Bibliometric Analysis", *Conservation and Recycling*, 131, 148-157. <https://doi.org/10.1016/j.resconrec.2018.01.002>
- Vrabel, M. (2015), "Preferred Reporting Items for Systematic Reviews and Meta-analyses", *Oncology nursing forum*. 42(5), 552-554.
- Yang, Kun-Woo (2017), "Research Trend Analysis of 'International Commerce and Information Review' Using SNA-based Keyword Network Analysis", *International Commerce and Information Review*, 19(1), 23-42.
- Yoon, Hee-young (2015), "Review of FTA Research Trends and Recommendations for Future Research Direction", *E-Trade Review*, 13(3), 71-90.