

텍스트마이닝 기법을 이용한 국내 농식품유통 연구동향 분석

유리나*, 황수철
(사)농정연구센터

A Trend Analysis of Agricultural and Food Marketing Studies Using Text-mining Technique

Li-Na Yoo*, Su-Chul Hwang

Center for Food, Agricultural & Rural Policy

요약 이 연구는 1984~2015년간 국내 농식품 유통분야 연구동향을 파악하기 위해 텍스트마이닝 기법을 이용한 분석 결과이다. 텍스트마이닝은 빅데이터 분석방법의 일환으로, 많은 정보를 객관적으로 처리하여 연구주제 분류와 트렌드 분석에 이용할 수 있다. 실제분석에는 빈도분석, 토픽분석, 연관성분석을 수행하였다. 자료는 농업부문 4개 학술지 수록논문과 연구보고서 중 농식품 유통 관련 연구 제목을 이용하였다. 그 결과, 농식품 유통분야의 논문 1,126건은 6개 주제로 분류되었다. 2000년대를 기점으로 이전에는 도매와 산지연구가 활발했던 반면 이후에는 소비, 식품, 수출입 연구가 활발한 것으로 나타났다. 또한 로컬푸드와 학교급식 영역의 연구가 증가했다. 농산물 수급연구는 정책 연구보고서에서만 주기적으로 이루어졌으며, 학술논문에서는 2000년대 이후 관심주제에서 멀어지는 경향을 보였다. 2010년대 이후로는 특히 소비연구가 주류를 이루었고, 크게 소비트렌드와 소비자 행동에 관한 다양한 연구가 이루어졌다. 이 결과를 바탕으로 더 정확한 연구동향 분석을 하기위해서는, 정밀한 주제 분류기법으로 방법론을 보완하고 이용 자료를 키워드와 논문초록으로 확대함으로써 구체적인 결과를 도출해야 할 것이다.

Abstract This study analyzed trends in agricultural and food marketing studies from 1984 to 2015 using text-mining techniques. Text-mining is a part of Big-data analysis, which is an effective tool to objectively process large amounts of information based on categorization and trend analysis. In the present study, frequency analysis, topic analysis and association rules were conducted. Titles of agricultural and food marketing studies in four journals and reports were used for placing the analysis. The results showed that 1,126 total theses related to agricultural and food marketing could be categorized into six subjects. There were significant changes in research trends before and after the 2000s. While research before 2000s focused on farm and wholesale level marketing, research after the 2000s mainly covered consumption, (processed)food, exports and imports. Local food and school meals are new subjects that are increasingly being studied. Issues regarding agricultural supply and demand were the only subjects investigated in policy research studies. Interest in agricultural supply and demand was lost after the 2000s. A number of studies after the 2010s analyzed consumption, primarily consumption trends and consumer behavior.

Keywords : Agricultural and Food Marketing, Big-Data, Marketing and Distribution, Text-mining, Trend Analysis

1. 서론

1.1 연구의 배경 및 목적

최근 기술 및 사회 환경의 변화와 경제 불확실성이 커

본 논문은 농촌진흥청 연구사업(PJ011889) 지원에 의해 수행되었음.

*Corresponding Author : Li-Na Yoo (Center for Food, Agricultural & Rural Policy)

Tel: +82-44-862-9209 email: linayoo29@daum.net

Received July 6, 2017

Revised (1st August 9, 2017, 2nd August 28, 2017)

Accepted October 13, 2017

Published October 31, 2017

지면서 미래변화 예측에 기반한 대응전략수립이 중요해지고 있다. 농식품 유통 연구영역에서도 인구사회구조의 변화와 스마트기술의 발전 등 메가트렌드에 대응한 중장기 연구개발 전략이 필요한 상황이다.

미래 연구개발 전략을 수립하기 위해서는 먼저 지금까지의 농식품 유통 연구를 체계적으로 종합하여 현재 국내 농식품 유통 연구가 어떤 영역에서 어느 수준으로 이루어졌는지 파악해야 한다. 기존연구는 유통단계, 경로, 상품, 기술, 정책 등 다양한 측면에서 진행되었으나 연구동향을 파악하는 방법은 연구자의 주관적 판단에 근거한 문헌조사나 전문가 의견에 의존한 델파이기법과 같은 정성적 방법으로만 이루어졌다.

정성적 연구방법에 따른 시간소요와 주관성의 한계를 보완하기 위해 최근 텍스트마이닝과 같은 정량적 연구방법이 활용되고 있다. 텍스트마이닝(Text-Mining)이란 자연어로 구성된 비정형 텍스트 데이터에서 숨겨진 패턴 또는 관계를 추출하여 의미있고 활용가치가 높은 정보 또는 지식을 찾아내는 분석기법으로 자연어처리기술을 기반으로 한다(김현정의, 2015)[1]. 수치화된 정형 데이터로부터 일정 패턴과 의미를 찾아내는 데이터 마이닝(Data Mining)과 달리 비정형 텍스트데이터를 기반으로 한다는 점에서 논문, 신문기사, SNS 등 대량의 텍스트자료를 처리할 때 다양하게 활용되고 있다.

최근 국내에서도 텍스트마이닝 기법을 활용한 분석이 시도되고 있다. 임시영외(2014)는 897편의 국내의 논문 키워드와 초록을 이용해 U-City와 Smart City 관련 논문들의 키워드를 분석하고 연관분석을 통해 각 군급의 특징을 비교하였다. 김현정의(2015)는 텍스트마이닝 기법 중 하나인 토픽분석을 이용하여 전체 논문집합을 주제별로 구분하고 트렌드분석을 시도하였다. 유사하게 강범일외(2013)는 언론사별 기사를 토픽분석을 통해 주제를 추출한 후 매체별 주제구성의 차이를 살펴보았다. 권미경(2014), 김명미(2015)는 논문집합으로부터 키워드를 추출한 뒤 키워드별 빈도분석과 네트워크분석을 수행함으로써 핵심 연구키워드들 간 연관관계를 분석하였다.[1~5] 이들은 모두 텍스트마이닝 기법을 이용해 논문 텍스트를 분석한 것들로서, 연구동향 분석을 목적으로 한다.

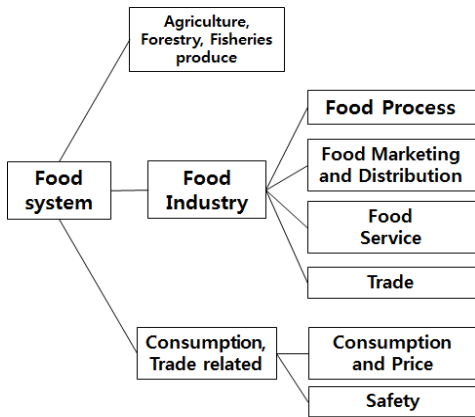
그러나 아직 농업분야에서는 빅데이터 기반 분석은 물론이고 정성적 방법에 의한 연구동향 분석연구도 거의 없다. 오성삼·이문영(1994)이 한국농업교육 연구동향

을 분석하고, 홍승길외(2014)가 유기농업 기술연구동향을 분석한 연구 정도가 있을 뿐이다[6~7]. 또한 특정 주제인 농산물 전자상거래 연구동향을 분석한 박성희외(2014)가 있다[8]. 해당 연구들은 모두 특정 범위의 선행연구에 대해 주제를 파악하고 통계분석을 시도한 사례이다.

본 연구는 국내 농식품 유통분야 연구동향 분석을 목적으로 하는 작업이다. 농식품 유통연구 전체라는 연구 범위에서 최초이며, 빅데이터 분석 방법론을 적용했다는 점에서 최초의 시도이다. 현재까지 이루어진 국내 농식품 유통영역 연구자료를 최대한 수집하여, 농식품 유통 연구가 시기별로 어떤 주제를 가지고 이루어졌는지 텍스트마이닝 기법으로 분석한다. 그리고 텍스트마이닝 방식에 의한 결과가 연구자의 판단에 의거한 정성적 분석결과와 어떤 차이를 보이는지 비교하기 위해 한국식품유통학회 발간 전문학술지인 ‘식품유통연구’의 논문을 대상으로 분석한다.

1.2 농식품 유통연구의 범위

종래 농식품 유통연구는 크게 유통단계연구, 유통경로연구, 상품연구, 유통기술의 경제성 분석연구, 유통정책연구 등으로 다양하게 분화되어 진행되었다. 그러나 최근 유통단계 및 경로 파괴는 물론 소매영역을 중심으로 융합된 영역이 등장하면서 소비자의 농산물 소비패턴 변화에 따른 산지가공 및 직거래방식도 달라지고 있다. 포화상태에 있는 농식품 유통의 발전전략을 모색하기 위해서는 농산물 생산부터 소비, 외식에 이르는 종합적 인식 틀이 필요하며 유통단계나 경로, 상품, 기술 등의 영역이 ‘푸드시스템(Food System)’내 복합적 요소임을 인식해야 한다. 현대화로 인해 농수산업이 최종소비에서 차지하는 비율이 감소하고 식품산업 비중이 지속적으로 성장하면서 “식품소비로 귀결되는 모든 흐름의 합계”(토코야마·에가이츠,2005:4)인 푸드시스템 개념은 더욱 중요해지고 있다. 따라서 본 연구에서는 농식품유통 연구를 푸드시스템 관점에 입각해 생산~도소매유통 및 제조가공~소매,외식, 학교급식 등 유통 전반을 포함하는 것으로 정의한다[Fig.1].



Data: Jin, Hyun-Jung · Hwang, Su-Chul(2010)

Fig. 1. Scope of the Food System

2. 분석방법

2.1 데이터 수집

국내 농식품 유통부문 연구현황은 학술논문과 연구보고서를 통해 파악할 수 있다. 분석을 위해 『식품유통연구』, 『농촌경제』, 『농업경제연구』, 『농업경영·정책연구』, 『한국축산경영학회지』의 게재논문과 한국농촌경제연구원 연구보고서, 그리고 농식품부, 농촌진흥청의 연구용역보고서 등을 수집하였다[Table1][9~12]. 농산물, 식품, 농식품, 산지, 도매, 농산물 소비, 농식품 소비를 키워드로 1962~2015년간 3824건의 국내 석·박사 학위논문도 수집하였으나 검색키워드의 타당성 및 키워드 선정기준의 모호성 때문에 이번 분석대상에는 포함하지 않았다.

실제 텍스트마이닝 분석에는 논문자료 3,672건 중 유통연구에 해당하는 1,126건을 이용하였다. 전체 수록논문을 대상으로 하지 않은 이유는, 유통연구에 해당하지 않는 키워드들이 결과에 영향을 미쳐 그룹별 주제를 뽑아내는 데 혼선을 주기 때문이다.

그리고 시기별 분석을 위해 『식품유통연구』의 발간 상황을 고려해 자료이용 기간을 1984~2015년으로 한정했다. 연구보고서의 경우 유통분야 연구가 많지않아 토픽분석 대상에 넣지 않고 정책연구와 학술논문 간 동향 비교에만 활용하였다.

Table 1. Collected Data for Statistical Analysis

	Data Source	Period	Count	
Theses	Korean Journal of Food Marketing Economics	1984~2015	608	KISS
	Journal of Rural Development	1978~2015	1,362	
	Korean Journal of Agricultural Management and Policy/Journal of Korean Livestock Management Society	2000~2015/ 1985~1999	698/ 171	
	Korean Journal of Agricultural Economics	1958~2015	833	
Reports	KREI	2000~2015	647	NKIS
	MAFRA	1999~2015	30	PRISM
	RDA	2000~2015	67	PRISM, ATIS

2.2 분석절차

분석절차는 크게 형태소 분석 및 전처리과정과 빈도 분석, 연관성분석, 토픽분석 등의 실제분석으로 이루어진다. 모든 분석에는 R 통계패키지를 이용하였다.

형태소분석 순서는 다음과 같다. 먼저 텍스트 문치(Corpus)를 가져온 뒤 띄어쓰기 기준으로 문자열을 나눈다. 나눈 문자열 중 2회 이상 반복되는 단어만 사용하고 나머지는 삭제한다. 그리고 난 뒤 R의 한글 자연어 분석 패키지인 KoNLP를 이용하여 보통명사 및 고유명사 품사만 추출한다. 마지막으로 논문제목에 빈번히 사용되나 결과도출에는 불필요한 단어인 ‘분석’, ‘연구’, ‘-을 중심으로’ 를 제거하였다.

실제분석에는 LDA 알고리즘 기반의 토픽분석을 수행하였고, 부차적으로 빈도분석 및 연관성분석을 실시했다.

2.3 토픽분석

토픽분석은 텍스트 집합 내에서 연관성이 높은 단어들을 기반으로 유사한 내용의 주제끼리 군집화하는 분석 방법으로, 단어가 모여 토픽이 되고 이 토픽들이 모여 전체 텍스트 집합을 이룬다는 개념에 기초한다. 일반적인 군집화 방법으로는 전통적인 ‘계층적 군집분석(Hierarchical Agglomerative Clustering)’이 있으며 장바구니 분석에 주로 활용되고 있다. 그러나 이러한 방법은 개별문서가 특정 주제에만 해당되는 방식으로 분류하기

때문에 하나의 텍스트가 여러 주제를 포함했을 경우는 고려하지 못한다는 한계가 있다. 특히 본 연구와 같이 논문주제를 분류하는 작업에서는 논문건수와 주제군의 1:1 매칭이 쉽지 않다. 예를 들면, “로컬푸드 직매장의 소비자 구매요인 및 유형별 차이 연구”라는 주제는 로컬푸드 및 학교급식과 같은 신유통경로 주제군에도 포함되지만 소비 단계 주제군에도 포함시킬 수 있다. 이를 극복하기 위해 최근에는 Blei et al.(2003)이 제안한 LDA(Latent Dirichlet Allocation)알고리즘에 입각한 토픽분석 연구가 이루어지고 있다[13]. LDA는 하나의 텍스트가 여러 주제를 다룰 수 있음을 고려하여 주제를 분류한다는 강점을 가져 남춘호(2016), 진설아외(2013) 등의 선행연구에서 적용한 바 있다[14~15]. 따라서 전체 분석단어 수와 각 주제군별 단어 수의 합이 일치하지 않는다. 전체 텍스트 집합 내에서 동시출현빈도가 높은 단어들을 기준으로 유사한 주제를 가진 텍스트를 분류하기 때문에 각 토픽군은 키워드 집합으로 표현된다. 따라서 각 토픽군에 대한 이름은 분류한 키워드의 집합을 보고 연구자가 결정하며, 토픽의 개수(k)도 연구자가 지정해줘야 한다는 특징을 갖는다.

실제분석에서 LDA 모형의 사후분포의 근사를 다루기 위해 붐피 깃스 샘플러(Gibbs Sampler) 알고리즘을 이용하였다. R에서는 통계분석을 용이하게 돕는 다양한 패키지를 제공하는데, 여기서는 붐피 깃스 샘플러 알고리즘을 구현한 "LDA" 패키지를 이용해 분석하였다.

2.4 빈도분석 및 연관성분석

빈도분석과 연관성분석은 말 그대로 텍스트 집합 내 자주 언급된 핵심 키워드를 찾아내고 그 핵심키워드와 연관관계를 갖는 키워드, 즉 한 문장 내에 동시출현빈도가 높은 단어를 찾는 방법이다. 이 분석을 위해서는 텍스트 집합을 단순 단어의 집합으로 인식해서는 안 되며, 문서를 기술하는 표현을 문서로부터 추출하고, 이로부터 분류를 예측하는 알고리즘을 만들어야한다. 이러한 목적으로 가장 많이 사용되는 기법은 단어-문서 행렬(TermDocumentMatrix)로 텍스트집합을 표현해 주는 것이다. 본 연구에서도 단어-문서 행렬을 만들어 준 뒤, 여기에서 자주 출현하는 단어를 찾아주고 빈도수가 높게 나온 단어와 연관성이 있는 단어를 찾아 무엇에 관한 연구가 이루어졌는지 검토하였다.

3. 분석결과

3.1 4개 농업 학회지 논문 분석결과

첫 번째 단계로 1984~2015년간 학술지 논문에 나타난 핵심 키워드를 파악하기 위해 빈도분석을 실시했다. 총 1,126 논문건수를 8,186건 단어로 분리한 뒤 가장 빈도수가 높은 단어 1~10위를 시기별로 추출하였다 [Table2]. 전 기간 가장 높은 비중을 보인 키워드는 ‘소비’로 나타났다. 1980~1990년대까지는 ‘도매시장’이 중요 키워드였으나 2000년 이후는 ‘소비’가 핵심 키워드로 부상했다. 특히 2010년 이후에는 ‘도매시장’ 빈도수가 5회로 급격히 떨어져 관심도가 크게 낮아졌음을 알 수 있다. ‘식품’ 키워드는 1990년대 이후로 꾸준히 증가하고 있다.

Table 2. Result of the Frequency Analysis on four Major Agricultural and Marketing Journals

	1984~2015	1980s	1990s	2000s	2010s
Total	8186	799	1655	3584	2148
1st	consumption(243)	marketing and distribution(40)	marketing and distribution(81)	sonsumption(106) market(106)	consumption(97)
2nd	agrifishery products(218)	agrifishery products(30)	agrifishery products(64)	agrifishery products(93)	food(57)
3rd	marketing and distribution(208)	wholesale market(17)	market(55)	marketing and distribution(78)	market(36)
4th	market(199)	supply and demand(18)	consumption(28)	wholesale market(56)	consumer(35)
5th	food(139)	consumption(12)	wholesale market(25)	food(55)	agrifishery products(31)
6th	price(109)	price(12)	supply and demand(24)	price(47)	purchase(30)
7th	wholesale market(103)	milk(10)	price(24)	brand(46)	price(26)
8th	consumer(73)	import(9)	food(21)	rice(33)	brand(21)
9th	brand(69)	meat(8)	export(21)	farm(32)	export(19)
10th	rice(63)	rice(7)	beef(15)	consumer(30)	rice(19)

Note: numbers in () indicate frequencies

두 번째 단계로 전체 1,126건의 논문을 LDA 모델링을 통해 분류하였다. 주제군 분류에 활용된 키워드는

2,178건으로 논문의 주제를 표현해주는 의미단어로만 구성했다. LDA는 모델링을 수행하기 위한 샘플링의 반복 횟수와 주제 수(k)를 지정해 줘야한다. 이 작업에서는 여러번의 시행착오를 통해 논문 텍스트 집합으로부터 주제해석이 가능한 주제집합을 추출한 것으로 파악되는 값을 얻기 위해 2000번 샘플링을 반복하여 6개 주제를 얻었다.

토픽분석 결과, 깃스 샘플러 알고리즘은 약 250번의 반복 후에 두 가지 로그우도가 수렴하는 것을 알 수 있다[Fig2]. [Table3.1~3.2]에는 논문의 6개 주제군 중 2군과 5군에 대한 확률분포를 제시하였다. 2군과 5군 양쪽에 ‘유통’이 높은 확률값을 보이며 등장한 것을 볼 수 있다. 같은 단어 일지라도 2군에서는 도매시장 관련 단어로 주제를 구성한 반면, 5군에서는 산지유통관련 단어로 주제를 구성하고 있음을 발견할 수 있다.

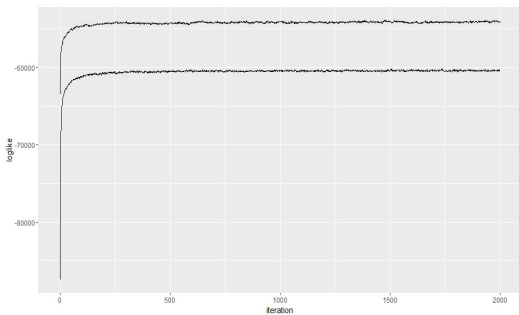


Fig. 2. Two log-likelihoods of LDA model

Table 3-1. Probabilities of Major Keywords in Topic[2]

Keywords	Probability
marketing and distribution	0.5474
wholesale market	0.3719
management	0.2475
improvement	0.1667
system	0.1432

Table 3-2. Probabilities of Major Keywords in Topic[5]

Keywords	Probability
marketing and distribution	0.4851
agricultural products	0.3050
farm	0.2646
sale	0.2811
case	0.1241

1984년~2015년간 4개 학회지의 논문은 식품 구입 및 지출에 관한 연구[3군], 유통일반, 산지출하 연구[5군], 도매시장 연구[2군], 농산물 수급연구[4군], 시장개방에 따른 수입농산물의 영향 및 시장가격에 대한 연구[6군], 소비·소비자, 브랜드 연구[1군] 순으로 주제비중을 보였다.

소비·소비자 연구[1군]은 2010년 이후에 특히 집중도가 높은 주제이나 2010년 이후 기간은 5년밖에 되지 않아 샘플 수 자체가 상대적으로 적어 비중이 낮게 나온 것으로 보인다[Table4].

Table 4. Result of the Topic Analysis on four Major Agricultural and Marketing Journals

Topics	Keywords	Words
[1] consumption, consumer, brand	consumption, consumer, brand, purchase decision, payment, characteristics, understanding, choice, behaviour	1,228
[2] wholesale market management	marketing and distribution, wholesale market, management, improvement, system, fruits and vegetable	1,369
[3] food, purchase	food, processed food, household change, safety, behaviour, purchase, expenditure	1,467
[4] supply and demand of agricultural products	supply and demand model, demand	1,359
[5] marketing and distribution	marketing and distribution, agricultural products, farm, sale, case, APC	1,409
[6] agricultural import and export, agricultural price	price, market, import, export, economy, effect, domestic agricultural products, beef, milk	1,355

세 번째 단계로 6개 주제군에 대한 시기별 연구활동 변화를 파악하기 위해 트렌드 분석을 실시하였다[Fig.3]. 학술지 게재논문은 개인 연구자의 관심사를 반영하기도 하지만 그 해 지원받는 연구과제의 성과와도 관련되어 있기 때문에 연간 변화보다는 큰 추이 변화를 보는 게 바람직할 것이다.

트렌드 분석은 연도별로 분류되어 있는 문서-단어 행렬을 기반으로 전체 문서 중 각 주제군별 핵심 키워드 비중으로 추정하였다. 핵심 키워드는 주제군마다 확률이 높은 5~6개 키워드를 대상으로 하였다.

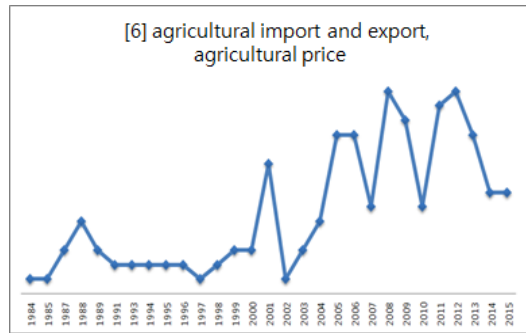
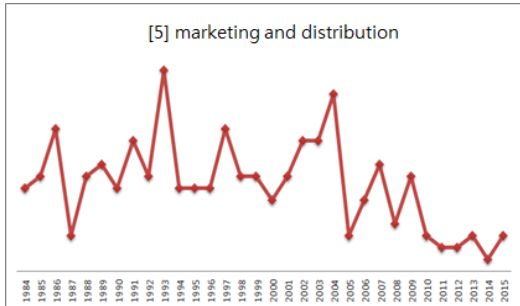
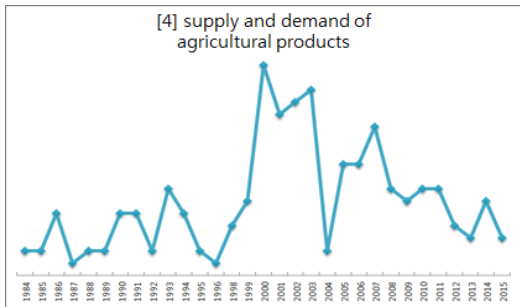
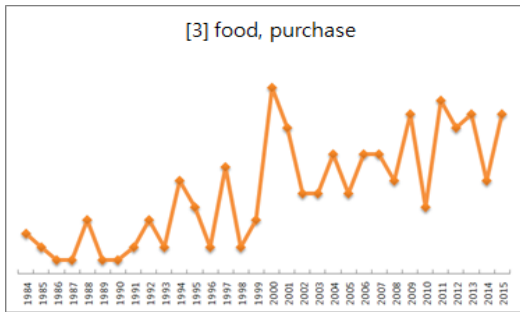
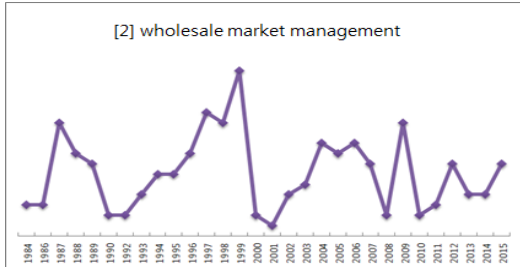
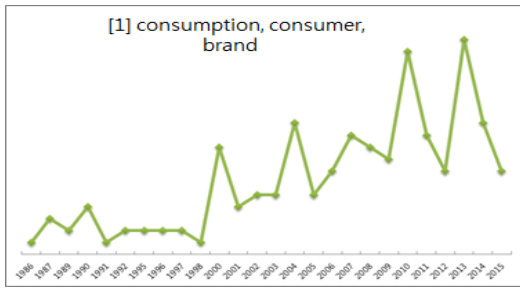


Fig. 3. Trends on Subjects of Agricultural and Food Marketing Studies

전반적인 변화추이는 다음과 같은 특징을 보인다. 2000년대 이후로 연구비중이 증가추세에 있는 주제는 [1군] 소비(자), 구매 주제와 [3군] 식품, [6군] 농산물 수출입 주제임을 알 수 있다. 기존 키워드 빈도분석을 통해 소비와 식품 연구 관심이 높아지고 있음은 확인했으나 추가적으로 2000년대 이후 완전개방시대에 따른 수입농산물과 국산 농산물의 경쟁력강화 문제를 다룬 연구들로 구성된 주제[6군]도 토픽분석을 통해 증가추세에 있음을 확인하였다.

반면 [2군] 도매시장 연구는 1990년대 가장 활발히 이루어졌으나 2000년대 이후로 연차별 등락을 보이며 낮은 관심수준에 머물러 있다. [5군] 유통일반, 산지연구도 점차 연구 관심도가 낮아지고 있다. 주목할 점은 [4군] 농산물 수급 관련 연구의 트렌드이다. 1990년대 말~2000년대 초 수급관련 연구가 활발했으나 이후 꾸준히 관심주제에서 멀어지는 양상을 보이는데, 이는 2000년대 이후 국내 농산물 과잉 및 예측실패 등 다양한 이유로 가격 변동성이 1990년대에 비해 커진 현실과 대조되는 추세이다.

이러한 결과가 나온 이유에 대해서는 면밀한 추가분석이 필요하겠지만, 우선 이용 자료가 학술지 게재논문이라는 점에서 농림축산식품부에서 발주하는 수급관련 연구용역이나 각종 수급안정 토론회 자료 등을 반영하지 못한 측면이 있을 것으로 보인다.

3.2 『식품유통연구』 논문 분석결과

연구자가 직접 주제를 분류하고 연구동향을 파악하는 정성분석 결과와 비교하기 위해 국내 농식품 유통분야 전문 학술지인 한국식품유통학회의 『식품유통연구』 게재논문을 대상으로 토픽분석을 실시했다. 실제 분석에는

1984~2015년간 총 608건의 게재논문 중 외국어로 된 논문과 국제연구 분야에 해당하는 55건을 제외하고, 553건의 논문을 이용했다. 샘플링은 1000회 실시하여 일정 패턴을 갖는 8개의 주제군을 도출했다[Table5].

Table 5. Result of the Topic Analysis on Korean Journal on Food and Marketing Economics

Subjects	Keywords	words
[1] agricultural brand, sales strategy	brand, strategy, common marketing, agricultural coop., purchase, choice, customer, loyalty	509
[2] agricultural produce, farm households	farmers, produce, vegetable, fruits, cabbages, mandarins, agricultural products, competitiveness	501
[3] local food, school meals, export, foreign cases	local food, direct outlets, school meals, export, joining of farmers, japan, U.S	353
[4] wholesale market contracts	wholesale market, contracts, system, auction, 중도매인	523
[5] marketing and distribution	improvement of marketing and distribution system, value chains, structure	465
[6] farm sales, direct sales	farm, agricultural products, sales, logistics, direct sale, packaging, marketing centre	518
[7] agricultural supply and demand, price	supply and demand, imports, forecast, stabilization, fluctuation, price, model, fruits, korean beef, milk	416
[8] consumer, consumption, food purchase	purchase, food, consume, consumer, satisfaction, behaviour, payment	441

추출한 8개 주제군의 비중은 전 기간 [4군] 도매시장 거래, [6군] 산지유통, 산지직거래, [1군] 농산물 브랜드, 판매전략, [2군] 농산물 생산, 농가, [5군] 유통일반, [8군] 소비(자), 식품구매, [7군] 농산물 수급 및 가격, [3군] 로컬푸드, 학교급식, 수출, 해외연구 순으로 나타났다.

[3군]은 어느 군에도 해당하지 않으면서 상대적으로 논문 편수가 적은 ‘기타’ 주제군이다. 실제 [1군]에 해당하는 연구는 [6군] 산지유통에도 포함될 수 있으나, 농산물 브랜드나 농협 마케팅사업 등과 관련 연구가 하나의 큰 주제군을 이루고 있어 별도 그룹으로 분리하였다.

시계열 트렌드를 보면 2005년 이후 도매시장 연구비중이 낮아지고 로컬푸드, 학교급식, 수출 관련 연구에 대한 관심이 높아지고 있음을 알 수 있다. 유통일반, 산지, 브랜드, 생산 관련 연구는 그 트렌드 추이를 파악하기 어려우나 농산물 브랜드와 마케팅 관련 연구는 2000년대 이후 꾸준히 연구가 이루어지고 있다. 소비(자), 식품 관

련 연구는 2000년대 이후 그 관심이 크게 높아졌다.

2000년 이후 새롭게 등장한 로컬푸드, 학교급식 주제군과 집중적인 연구가 이루어지고 있는 소비 및 식품 논문들은 구체적으로 어떤 연구를 하는지 살펴보았다.

먼저 로컬푸드와 학교급식 연구는 건수 자체는 적으나, 로컬푸드 직매장에 대한 연구와 학교급식 유통망 및 식재료 공급에 관한 연구가 진행되었다[Table6].

Table 6. Local Food, School Meal, Export related theses since 2000

Year	Title
2015	a study on consumer purchasing factor and type differences in local food farmer's market
2014	the development strategy of the local food store
2013	an exploratory study on relationship between local food's brand equity, origin effect and brand loyalty in korea
2013	current status and suggestions for improvement of school meal service support centers
2013	shopping motives and customer satisfaction in the local food store
2012	impacts of an increase in the export of paprika on the market equilibrium and social welfare
2011	analysis on the actual conditions about parents' recognition of school meals
2010	a policy direction for promoting the korean agro-food export
2009	analysis on the characteristics of mushroom exports in korea
2007	analysis of the school milk program in japan for expansion of school foodservice's milk consumption
2005	the plan for the food supply system in school food service for the consumption extension of agricultural products
2004	improvement of the food supply system in school food service

두 번째로 소비·식품연구는 크게 ‘소비 트렌드’와 ‘소비자 행동’에 관한 다양한 연구가 이루어졌다. 소비 및 소비자를 연구한 논문 모두 농산물보다는 가공식품을 대상으로 하였다. ‘소비’ 관련해서는 주로 소비패턴, 소비행태, 소비실태 연구가 이루어졌다. ‘소비자’ 관련해서는 상품구입과 소비자 속성이 주요테마였다. 상품구입에 관한 기존연구는 다시 지불의사, 구매의도, 구매행태로 구분할 수 있으며, 소비자 속성에 관한 기존연구는 소비자 선호, 충성도, 인식, 태도 등으로 구분할 수 있다 [Table8].

알고리즘에 의한 정량분석 결과와 정성분석 결과를 비교하기 위해 직접 논문제목과 보고 주제 분류를 시도하였다.

Table 7. Categorization of All Theses in Korean Journal on Food and Marketing Economics(1984~2015)

Category	No. of theses	Proportion(%)
Marketing and Distribution	298	53.9
etc	119	21.5
Consumption	98	17.7
Production	19	3.4
Food Industry	9	1.6
Eating out	7	1.3
Agribusiness	3	0.5
Total	553	100.0

Table 8. Consumption and Food related theses of 2010~2015

Year	Title
2015	An Analysis on Estimation of Willingness to pay for the new products of citrus tea assessed by korean consumers
2015	an analysis of the influence of lifestyle patterns on consumer purchasing behavior for Makgeolli
2015	analysis for factors on buying decision making by female college and graduate students toward Gluten-free food
2015	effect of consumers' doetary lifestyle on the consumption pattern of processed foods
2014	analysis on consumers' perceptions and preferred attributes toward the processed rice foods
2013	analysis for factors on the level of consumers' coenesthesia for food risk and hazard
2013	production, marketing and domestic foreigners' consumption patterns of subtropical vegetables
2013	a study of intention to purchase for organic food in restaurant industry
2013	an analysis of the consumers' acknowledgement level of seo-dong yam
2013	analysis of consumer's risk perception about food safety through the analytic lens of psychometric paradigm
2013	analysis for homemakers' attitude toward the traditional markets
2012	the effects of customers' consciousness about the environment and the environmental agricultural products on their trust, attitude and loyalty
2012	consumer's perception for the medicinal crop and processed product
2012	measuring willingness to pay for safety inspection system of chinese import food
2012	analysis of consumer preference for rice-noodle for consumption promotion
2012	an analysis of consumers' preference for the blending oil
2012	analysis of satisfaction and willings to pay for the high-quality satsuma mandarin
2012	research on consumer preference of egg through conjoint analysis
2011	a study on imported fruit consumption in beuijing and shanghai
2011	a study on the consumers' satisfaction of yangu ligularia fischeri
2011	a structural model for korean apple preference and purchase of malaysian consumers
2011	an analysis of factors influencing the food consumption pattern

2011	the willingness-to-pay price for Makgeolli made of domestic rice
2011	a study on japanese consumers' preference for lilies and export expansion strategies
2011	the study on the selective attributes of restaurant
2011	consumer preference for the quality attributes of fruit juices
2011	a conjoint analysis of consumers' purchasing behavior for mixed cereal using salesclerks in the cereal sector of discount stores
2010	a study on consumer acceptance of food technologies and functional foods
2010	an anlysis and implications on the consumption and consciousness situation of green consumers
2010	a study on the consumption pattern of laver
2010	analysis for consumer agricultural cyber shopping mall purchase behavior on the patterns of lifestyle
2010	consumers' purchasing patterns and preferences of the processed food products made from domestic ingredients
2010	a study of consumers' food choice behavior by comparison of past experience
2010	a study on chinese consumers' korean food consumption behavior based on food-related lifestyle

분류기준 및 절차는 다음과 같다. 첫째, 본 연구에서는 농식품유통연구의 범위를 푸드시스템 관점에서 입각해 정의하였기 때문에 푸드시스템의 영역구분에 대한 국내 농업경제학 교과서를 참고하여 6개의 대분류를 만들었다(한국농업경제학회, 2014; 에가이즈후미오, 2007; 도쿄야마 히로미 에가이즈 후미오, 2006)[16~18]. 둘째, 식품유통연구에 수록된 553편의 논문을 ‘생산’, ‘유통(산지~소매)’, ‘소비’, ‘식품제조(가공)’, ‘외식’, ‘애그리비즈니스(진·후방산업 및 서비스산업)’ 등 6개의 대분류에 따라 구분하고 특정 주제군에 넣기 어려운 성격의 논문은 ‘기타’로 구분하였다. ‘기타’에 포함되는 논문 119편은 시장개방 4건, 한국농업 3건, 가치사슬, 경영효율성, 경영, 선물거래, 정보화, 입지분석 각 2건, 나머지 100건은 모두 각기 다른 주제를 보였다.

셋째, 2명의 평정자(연구진)이 1차 분류한 결과를 검토하고, 평정자간 분류가 일치하지 않는 논문에 대해서는 토론과 합의를 거쳐 2차 분류작업을 하였다(Table7).

실제 학회지 전체 논문에서 수출관련 논문은 총 27건으로 파악된다. 그러나 [Table6]에서 볼 수 있듯이, 3군에 포함된 3건의 수출논문을 제외한 나머지 논문은 마케팅, 조직화, 농산물 브랜드 등의 주제군으로 분류되고 어디에도 속하지 않은 나머지 수출논문이 소위 기타논문에 해당하는 3군으로 분류된 것으로 파악되, 추가 조정작업이 필요하다.

주제분류 중 ‘유통’은 산지~소매에 이르는 유통단계와 관련된 논문이다. 전체 논문의 71.6%가 ‘유통’과 ‘소비’ 관련 논문이며, 분류가 어려운 ‘기타’를 제외하면 대부분의 논문은 이 두 부류에 해당함을 알 수 있다. 전 기간 ‘유통’연구는 다시 도매유통, 기타, 산지유통, 수출, 브랜드, 소매유통, 가격, 수급 순으로 비중이 높았다 [Figure4-1, 4-2]. 분류체계는 다르지만 토픽분석에서 도매시장, 산지유통, 브랜드, 가격 및 수급 순으로 연구 비중이 도출된 것과 맥을 같이 한다.

Marketing and Distribution

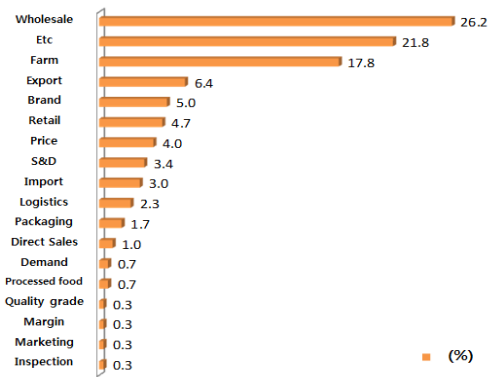


Fig. 4-1. Subcategorization of 'Marketing and Distribution' theses in Korean Journal on Food and Marketing Economics(1984~2015)

Consumption

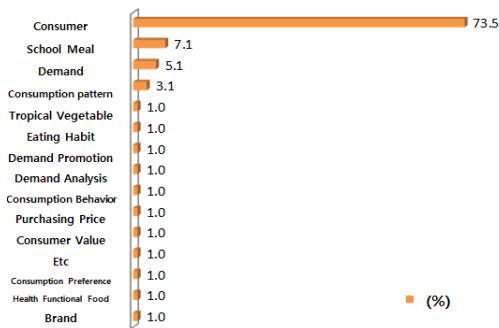


Fig. 4-2. Subcategorization of 'Consumption' theses in Korean Journal on Food and Marketing Economics(1984~2015)

‘유통’연구는 전 기간 도매와 산지 연구가 각각 26.2%, 17.8%로 큰 비중을 차지하지만, 두 영역 모두 2000년대 이후에는 논문 수가 감소하면서 전통적인 유통

단계 연구의 감소세를 이끌고 있다[Figure5-1].

Marketing and Distribution

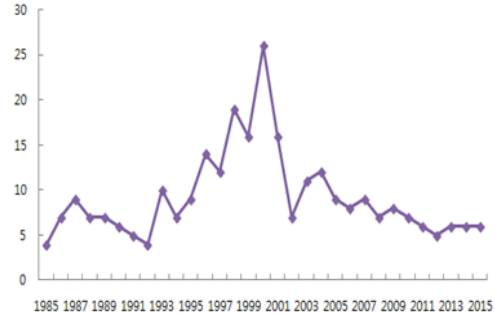


Fig. 5-1. Trend of 'Marketing and Distribution' theses in Korean Journal on Food and Marketing Economics(1984~2015)

Consumption

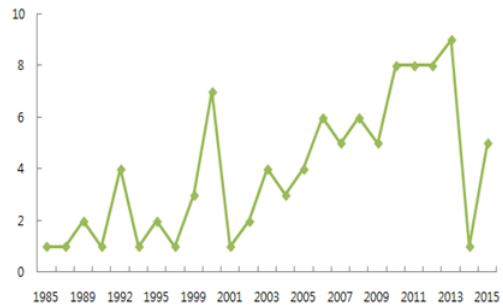


Fig. 5-2. Trend of 'Consumption' theses in Korean Journal on Food and Marketing Economics(1984~2015)

반면 소비연구는 2000년대 이후로 논문 수가 증가 추세에 있음을 확인할 수 있다[Figure5-2]. 전 기간 (1984~2015) 농식품유통연구의 주요 연구영역과 관심 연구주제의 변화흐름은 토픽분석을 통한 정량적 분석과 연구자 판단에 의거한 정성적 분석이 유사하게 나타났다.

그러나 토픽분석에서 소비, 소비자, 식품 연구를 하나의 주제군으로 묶고 학교급식, 로컬푸드, 수출을 일종의 기타 주제군으로 묶은 것과 달리 정성분석에서는 수출을 유통군에 넣고 학교급식을 소비군에 넣었다는 점에서 인식에 차이가 있다. 또한 정성분석에서는 유통단계에 따라 산지·도매·소매로 구분할 수 있었지만, 토픽분석에서는 실제 제목에 사용된 단어를 기준으로 분류하기 때문에 ‘소매’라는 용어를 주제어로 인식하지 못했다.

3.3 정책연구보고서 분석결과

학술논문의 연구주제 동향과 정책연구 동향을 비교하기 위해 한국농촌경제연구원, 농림축산식품부, 농촌진흥청의 연구보고서를 분석했다. 2000년 이후 농림축산식품부와 농촌진흥청의 연구보고서는 정책연구관리시스템(PRISM)과 농촌진흥사업종합관리시스템(ATIS)를 통해 농식품유통 연구로 파악되는 것을 수집할 수밖에 없었으나 한국농촌경제연구원의 연구보고서 및 각종 발간물 정보는 국가정책연구포털(NKIS)을 통해 일괄 수집했다. 국가정책연구포털을 통해 수집한 647건 연구보고서 목록 중 농식품유통에 해당하는 128건만 분석에 이용하였다.

한국농촌경제연구원, 농림축산식품부, 농촌진흥청의 2000~2015년 간 연구보고서 213건을 분석하기 위해 먼저 키워드 빈도를 확인했다. 그 결과 농산물 다음으로 많이 등장한 단어는 '소비(자)'로 나타났다. 다음으로 많이 등장한 '전망'은 2005~2008년간 한국농촌경제연구원에서 농업부문 전망모형 KREI-ASMO와 세계농업 전망모형 Aglink 구축 연구가 진행되어 빈도가 높게 나왔다.

다음으로 정책연구 분석과 관련해서는 키워드 연관성 분석을 통해 두 가지 현상을 검토했다. 첫째, 학술연구 분석에서 가장 뚜렷한 최근 연구양상으로 나타난 소비연구가 정책연구에서도 동일하게 확인되는가. 둘째, 학술연구에서는 2000년대 이후 수급관련 연구 관심이 낮은 것으로 나타났는데 정책연구의 경우는 어떠한가를 확인하였다.

첫째, 수급관련 정책연구는 2년 주기로 이루어졌다. 2000년 이후로 쌀 수급안정 관련 연구가 3회, 한국과 농산물 최대교역국인 중국 내 수급관련 연구가 2건, 과일·채소 수급연구가 3건, 기타 임산물, 우유, 축산 수급연구가 이루어졌다. 이 밖에 수급안정 정책간담회, 토론회, 농업전망 등을 통해 각종 수급논의가 정부정책 중심으로 이루어지고 있다[Table9].

둘째, 소비와 관련해 어떤 정책연구가 이루어졌는지 보기 위해 연관성 분석을 실시하였다. 2005년 전후로는 2000년대 초중반 한미FTA체결 관련, 수입산 쇠고기와 쌀 소비관련 연구가 진행되었다. 2007년 이후로는 친환경, 안전농산물(2007~2008) 소비관련 연구가, 2007~2014년까지는 거의 매년 농식품 혹은 식품소비 관련 연구가 이루어져 학술연구와 마찬가지로 정책연구도 갈수록 소비, 특히 식품소비 영역으로 이동하고 있음을 알 수

있다[Table10].

Table 9. Keywords related to Subjects on 'Supply and Demand'

Keywords	Year
chinese rice	2003
milk	2005
fruits and vegetable	2005
rice	2007
rice	2009
vegetable	2011
beef and pork	2011
rice	2011
forest product	2013
chinese agricultural product	2015
carrot	2015

Table 10. Keywords related to Subjects on 'Consumption' and 'Consumer'

Keywords	Year
beef, consumer attitudes	2002
increase in eating out, agricultural products consumption	2002
agricultural products, consumption patterns	2003
rice, consumer attitudes	2003
U.S., BSE, beef, consumption change	2004
processed crops, consumer response	2005
imported rice, users' intention	2006
eating habits, agricultural products, increasing consumption	2006
imported rice, consumption substitute	2006
food consumption trend	2007
environmentally-friendly agricultural products, consumer confidence	2007
environmentally-friendly agricultural products, increasing consumption	2008
safe food, increasing consumption	2008
agricultural products, consumer attitudes	2008
korea, china and japan, youth, food consumption	2009
korean wheat, consumer awareness	2010
organic processing food, consuming practice	2011
food, consumer policy	2011
environmentally-friendly agri-food, consuming practice	2012
agri-food policy, consumer communication	2013
horticulture industry, stable consumption	2013
cut flowers, consumption	2013
fruits, consuming practice	2014
japan, radiation leak, domestic agri-food consumption	2014
korean gastronomy, rice consumption	2015

4. 요약 및 결론

본고에서는 국내 농식품유통 학술 및 정책연구 자료를 수집·분석하여 연구주제 동향을 분석했다. 분석기법으로는 빅데이터 분석방법의 하나인 텍스트마이닝을 적용하여 토픽분석, 키워드빈도분석 및 연관성분석을 수행했다. 또한 텍스트마이닝 분석결과의 합리성을 판단하기 위해 연구자의 판단에 의거한 분석을 일부 적용하여 결과를 비교했다.

1980년대 중반 이후 농업부문 4개 학회지 논문은 크게 6개 주제로 분류할 수 있다. 2000년대를 기점으로 이전에는 도매와 산지연구가 활발했던 반면 이후에는 소비, 식품, 수출입 연구가 활발했다. 반면 수급연구는 시기에 크게 관계없이 농산물 유통의 중요이슈인데도 학술 논문에서는 2000년대 이후로 연구실적이 저조했다.

『식품유통연구』만을 대상으로 한 결과에서는 토픽 분석으로 8개 주제분류를, 연구자가 직접 분류한 방법으로는 7개 주제분류를 보였다. 7개 주제분류 중 유통, 기타, 소비가 전체의 93.1%를 차지하기 때문에 세부분류는 이들에 대해서만 검토했다. 두 분류결과를 비교한 결과, 전 기간 공통적으로 도매와 산지연구가 가장 많은 것으로 나왔으며 마찬가지로 2000년대 이후 그 비중이 감소하면서 소비연구 비중이 크게 증가한 것으로 나타났다. 그러나 두 방법은 전체적인 분류체계가 달라 기타 연구를 처리하는 방식에서 차이를 보였다.

정책연구보고서도 점차 그 주제가 소비·소비자에 집중하는 경향을 보였다. 그리고 학술연구에서는 미진했던 주기적 수급연구가 정부의 수급안정정책 논의와 함께 이루어지고 있음을 알 수 있다.

텍스트마이닝 기법을 이용한 정량분석은 방대한 양의 정보를 짧은시간 안에 처리하고 분석과정에 최대한 객관성을 확보할 수 있다는 장점을 갖는다. 그리고 미래 계획 수립 등에 필요한 트렌드 분석에 효과적으로 이용할 수 있다.

그러나 몇 가지 측면은 보완이 필요하다. 첫째, 토픽 분석을 통한 주제분류가 완벽하지 않다. 4개 학회지 논문을 분류한 6개 주제 중, 5군 주제 하나로 묶어버린 ‘유통일반’과 ‘산지유통’은 각각의 개별 주제군으로 분리되는 것이 타당하다.

둘째, 분석에 적합한 형태의 자료를 확보하기가 어렵다. 본 연구에서는 각 학회지별 논문관리 상황이 일정하

지 않아 논문제목만을 분석에 활용했다. 그러나 더 신뢰성있고 구체적인 결과를 얻기 위해서는 논문초록과 키워드를 활용해야한다. 자료 처리 과정에서 학회지 게재는 문인데도 키워드가 빠져있는 경우가 많았다. 또한 국제 학술지가 아닌 국내학술지인데 초록과 키워드를 영문으로만 작성하도록 되어있는데, 연구자마다 같은 개념이더라도 사용한 영어단어가 달라 자료로 활용할 수 없었다. 키워드에 한해서는 영문과 한글을 병기하는 방안도 학회지 관리 및 활용도 차원에서 고려해볼직하다. 셋째, 『식품유통연구』 논문의 주제 분류 과정에서 ‘기타’ 비중이 21.5%로 나왔는데, 이 중에는 농식품 유통관련 주제도 있지만 관련이 없는 주제도 많았다. 이 부분은 본 연구의 보완·발전을 위해서 개선되어야 할 부분이지만 학회지의 정체성, 전문 연구물의 축적·관리와도 관련된 사안이다.

넷째, 이 연구에서는 시기별 중점연구 변화에 대한 원인을 규명하지 못하였다. 후속연구를 통해 전문가 인터뷰 및 당시 학회 학술대회, 심포지엄의 연구주제 등을 검토함으로써 원인을 파악할 필요가 있다.

본 연구결과를 토대로 앞으로는 ‘소비’와 같이 최근 관심갖고 있는 연구주제 안에서 유사하거나 다른 세부주제가 어떻게 이루어지고 있는지 빅데이터 방법론으로 분석할 필요가 있다. 또한 동일 주제 군 내에서 얼마나 다양한 방법론으로 연구가 이루어지고 있는지도 분류를 해봄으로써 더 의미 있는 결과를 도출할 수 있을 것이다.

References

- [1] Kim Hyunjung et al., “Text mining-based emerging trend analysis for the aviation industry”, Journal of intelligence and information systems, vol. 21, no. 1, 2015.
- [2] Lim Siyeong et al., “Study on the trends of U-City and smart city researches using text mining technology”, Journal of the Korean Society for Geospatial Information System, vol. 22, no. 3, 2014.
DOI: <https://doi.org/10.7319/kogsis.2014.22.3.087>
- [3] Kang Beomil et al., “A Study on opinion mining of newspaper texts based on topic modeling”, Journal of the Korean Society for Library and Information Science, vol. 47, no. 4, 2013.
DOI: <https://doi.org/10.4275/KSLIS.2013.47.4.315>
- [4] Kwon MiKyung, Analysis on the research trend of human resource development utilizing the keyword network analysis, Master's thesis, Sookmyung women's university, 2014.
- [5] Kim Myungmi, Analysis of research trend on R&D

performance using social network analysis, Master's thesis of the graduate school of Hanyang University, 2015.

- [6] Oh Sungsam and Lee Munyoung, "An analytical study on korean agricultural education research trends", Journal of agricultural resource development, 15, 1990.
- [7] Hong Seunggil et al., "Research trends in organic farming technology by journal article analysis", Korean journal of organic agriculture, vol. 22, no. 4, 2014.
- [8] Park SungHee et al., "Analysis of the Agricultural Electronic Commerce: A Study on Documentary Research and Present Condition Comparison", Korean Journal on Food and Marketing Economics, vol. 31, no. 1, 2014.
- [9] Korean Food Marketing Association, Journal of Food Marketing Economics, 1984-2015.
- [10] Korea Rural Economic Institute, Journal of Rural Development, 1978-2015.
- [11] Korean Agricultural Economics Association, Korean Journal of agricultural economics, 1958-2015.
- [12] Korean Livestock Management Society and Korean Agri-Food Policy Society, Korean Journal of Agricultural Management and Policy, 2000-2015.
- [13] Blei, M, A. Ng, M. Jordan, "Latent Dirichlet Allocations," Journal of Machine Learning Research, Vol. 3, 2003.
- [14] Nahm ChoonHo, "An Illustrative Application of Topic Modeling Method to a Farmer's Diary", Journal of Comparative Culture Research, vol. 22, no. 1, 2016.
- [15] Jin SeulA, "Topic-Network based Topic Shift Detection on Twitter", Korea Society for Information Management, vol. 20, no. 1, 2013.
DOI: <https://doi.org/10.3743/KOSIM.2013.30.1.285>
- [16] Korean Agricultural Economics Association, Agricultural Economics, Yulgok, 2014.
- [17] Fumio Egaitu, Agricultural Economics, Youngbong You(Trans), Jeju University Press, 2007.
- [18] Hiromi Tokoyama and Fumio Egaitu, Economics of Food System, Lee Byungoh and Ko Jongtae(Trans), Kangwon University Press, 2005.

유 리 나(Li-Na Yoo)

[정회원]



- 2013년 8월 : 한국외국어대학교 국제지역대학원 EU학과(경제학 석사)
- 2014년 1월 ~ 현재 : (사)농정연구센터 선임연구원

<관심분야>

농가소득, 로컬푸드, EU농업

황 수 철(Su-Chul Hwang)

[정회원]



- 1996년 2월 : 서울대학교 대학원 경제학과(경제학박사)
- 2007년 1월 ~ 현재 : (사)농정연구센터 소장
- 2014년 1월 ~ 2015년 12월 : 한국식품유통학회 회장

<관심분야>

농업성장, 푸드시스템