

텍스트마이닝을 활용한 전통 장류 언론자료 분석

유창환¹, 이병서^{2*}

¹세계김치연구소 산업지원연구단, ²농진중양회 한국농업개발원

Analysis of Media data for Traditional Pastes using Text Mining

Chang-Hwan Yu¹, Byong-Seo Lee^{2*}

¹Industrial Solution research group, World Institute of Kimchi

²Korea Agricultural Development Institute, Nongjin Association

요약 본 연구에서는 2017년부터 2022년 8월까지 5년 8개월간 보도된 전통 장류 관련 신문기사 1,130건을 수집하고 해당 기사들에서 핵심 키워드를 추출하여 분석하였다.

우선 1,130건의 기사에서 출현 빈도 및 가중치가 높은 전통 장류 관련 상위 10개 핵심 키워드를 도출하였다. 도출된 키워드를 분석하여 현재까지 전통 장류 관련 신문 기사의 특징을 살펴보고 도출된 10개 키워드 간 관계를 연관분석기법으로 수치화 및 시각화하였다. 분석 결과 '전통 장류' 키워드를 가장 큰 허브로, '순창군', '청국장'과 '우수성'을 하위 허브로 키워드 그룹이 형성되는 구조였다. 이는 전통 장류 관련 기사는 다양한 주제 중에서 '순창군', '청국장' 및 '우수성'이 중심임을 보여준다.

다음으로 중복 기사를 제외한 628건 신문 기사 내 핵심 키워드를 추출하는 분석을 하였다. 주요 키워드는 전통 장류의 지역, 품목, 형용사에 관련된 키워드들이 추출되었다. 중복 기사를 제외한 주요 키워드는 순창군, 강진군, 중학교가 상위에 있었다. 지역 키워드는 주로 지자체의 정책 지원과 연관되었음을 유추할 수 있으며, 전통 장류의 전체 기사와 중복 기사를 제외한 키워드의 우선순위는 차이점이 있는 것으로 나타났다.

본 연구는 신문 기사를 직접 분석하는 기존의 방법에서 벗어나 텍스트마이닝 기법을 활용하여 텍스트 데이터를 분석하였다는 데 의의가 있다. 기존 내용분석 방법과 달리 텍스트마이닝을 사용하여 텍스트 통계량을 계산하고 이를 연관분석기법과 연계함으로써 기사의 내용을 키워드 형태로 파악하고 특징적인 메시지를 도출한 실험적인 연구이다. 분석을 통해 나타난 결과를 요약하면, 전통 장류 관련 기사는 특정 지역 중심의 추진사업에 주목하고 있어 전통 장류의 문화적 가치와 역사성, 산업발전 방안 등을 대중과 공유하려는 노력이 필요하다고 판단된다.

Abstract In this study, we examined 1,130 newspaper reports related to traditional pastes among traditional foods during a 68 months duration (Jan 2017 to Aug 2022) to analyze keywords and extract core words from the articles.

The top 10 keywords referring to traditional sauces were extracted, depending on their frequency of appearance and weightage in these articles. The keywords were analyzed to examine the characteristics of the newspaper articles reporting traditional sauces. The association between the 10 keywords was subsequently quantified and graphically visualized using the association analysis. Results revealed that the largest group formed contained the keyword 'traditional pastes', with Sunchang-gun, Cheonggukjang, and excellence as the major sub-hubs. This indicates that amongst the various factors discussed, articles on traditional sauces focus more on Sunchang-gun, Cheonggukjang, and excellence.

The next analysis extracted keywords from 628 newspaper articles, excluding duplicated ones. Primary keywords were selected based on the region, item, and adjective of traditional sauces. The major keywords determined were Sunchang-gun, Gangjin-gun, and middle school. Thus, we inferred that regional keywords are well-associated with the supporting policy of local governments, and there are differences in keyword trends when considering all articles and excluding duplicated articles.

This study is significant in that text mining is used to analyze the press data instead of applying direct analysis prevalently used. Contrary to conventional research, this is an experimental study wherein text mining is applied to calculate text statistics, followed by combining with association analysis to identify the article contents in the form of keywords and uncover characteristic messages. In summary, we determined that the articles related to traditional pastes focus on region-specific support and promotion projects rather than on the sauce industry or cultural values. Although newspaper articles tend to concentrate on social issues with the property of the newspaper considered, it is necessary to continuously share and publicize the cultural values, historicity, and development plan of the sauces.

Keywords : Traditional Sauce, Newspaper Article, Text Mining, Wordcloud, Social Network Analysis

본 연구는 농촌진흥청 농업과학기술개발사업(PJ016688) 재원으로 세계김치연구소 산업지원연구단(KE2203-1-1)과 수행되었음.

*Corresponding Author : Byong-Seo Lee(KADI)

E-mail: leebso31@naver.com

Received October 4, 2022

Revised October 26, 2022

Accepted December 7, 2022

Published December 31, 2022

1. 서론

1.1 연구의 배경

한국 정부는 '97년 유통개혁 대책의 일환으로 '산지 가공산업육성'을 통한 농수산물의 수급 조절과 가격안정, 농의소득 증대, 농촌경제 활성화 도모 및 우리 농산물에 가공기법을 접목시킨 전통 식품 개발·계승·발전을 추구하고 있다. 또한, 전통 식품의 정의를 "국산 농수산물을 주원료로 제조·가공되고 예로부터 전승되어 오는 우리 고유의 맛·향 및 색깔을 나타내는 식품"이라고 고시되어 있으며, 전통 식품 품목은 14종류(과자류, 죽류, 식육제품류, 당류, 김치절임류, 두부류, 식용유지류, 면류, 다류, 음료, 조미식품류, 주류, 인삼류, 기타식품류)로 나누어져 있다[1].

전통 식품의 개발과 활성화는 원재료의 수요창출과 농업인의 소득 증대를 도모하고 가공을 통해 유통기간을 연장하여 부가가치를 제고시킬 수 있다.

소비자 가구의 식품 소비 구조는 '신선 식품비', '가공 식품비', '외식비'로 나눌 수 있는데 최근 생활 여건의 변화로 가공식품 소비 비중은 29.6%로 성장 잠재력이 큰 것으로 보인다[2]. 하지만, 식품 소비에서 가공식품의 비중이 증가하면서 수입 원료사용량이 늘어 국산 농산물의 사용이 제한되는 양상을 보인다. 그러나 전통 식품은 국산 원료를 주로 사용하기 때문에 일반 가공식품과 차별성이 있으며, 현재 WTO 체제의 수입 개방으로 우리 농산물 수요 감소와 가격하락 문제가 대두되는 상황에서 전통 식품이 가지는 기여도는 주목할 필요가 있다.

'21년 국내 장류 시장규모는 생산량 72만 2,366톤, 생산액 9,027억원으로 5년간 평균 4.9% 증가하였다. 하지만, 장류 산업은 성장세가 둔화되면서 현재까지 약 1조원대의 규모를 벗어나지 못하고 있으며, 몇몇 장류 제품은 오히려 생산량이 감소하고 있다[3].

이러한 상황에서 최근 전통 장류와 관련하여 농업과의 연계를 위한 전통식품 관련 시설 지원, 장류산업진흥법 제정, 장 발효기술 및 소재화·소스화 연구개발, 수출시장 개척 등 다양한 이슈들이 언론에서 거론되고 있는데, 이러한 광범위한 자료들을 대상으로 빅데이터 분석을 통해 전통식품의 가치와 시사점을 파악할 필요가 있다.

전통 장류 관련 선행연구를 살펴보면, Kim[4]은 한·중·일의 대두 발효식품의 특성과 국내 장류 산업, 수출 현황 등을 정리, 분석하고 국제식품규격위원회(CODEX)의 장류 규격화 기준과 조건을 조사하여 우리 두류 발효 식품의 비전을 제시하였고 Lee[5]는 장류의 발전 현황과

다양한 브랜드 장류 상품의 이용실태를 조사하여 장류 산업 발전 방향의 기초 자료를 정립하고, 현대인들의 장류에 대한 인식 및 기호성과 관심도를 고찰하였다.

Ku[6]는 우리나라 장류를 중심으로 발효식품의 변화와 현황 연구를 위해 발효식품 관련 논문, 인터넷 웹사이트 자료 등을 수집 정리하고, 식품 고문헌과 국역 사업으로 추진된 농식품 고문헌의 편역본 내용을 추가로 정리하였으며, 지역적 특성 연구를 위해 순창군의 장류 산업을 대상으로 결과들을 분석 및 종합하여 기술하였다.

지금까지 전통 장류의 연구는 장류 산업 현황, 발효식품 및 소비자 구매 특성, 제품개발, 수출확산 등의 연구가 주를 이루고 있으며, 시대적 배경에서 나타나는 근본적인 문제나 이슈를 다룬 연구는 거의 없는 실정이다.

빅데이터 활용 연구 분야에서 보면, Kim et al.[7]은 IT 기술의 급격한 발달로 인해 방대한 양의 데이터가 빠른 속도로 생성·공유되는 것이 가능해지고 스마트 기기의 대중화에 힘입어 텍스트, 이미지, 동영상 등 비정형 데이터가 사회연결망을 통해 급속도로 유통되고 있으며, 이러한 비정형데이터를 분석하여 의미 있는 정보를 생산하는 방법은 빅데이터 분석 기술과 함께 발전하고 있는데 그중 의사 표현 및 정보전달의 대표적인 수단인 텍스트를 분석하는 텍스트마이닝(Text Mining)은 다양한 목적에 활용되고 있다고 하였다. Metzler et al.[8]은 텍스트 형태의 비정형데이터로 이루어진 정보를 사용하는 모든 분야를 아우를 정도로 그 활용 분야가 매우 다양하며, 특정 기사(Article), 원문(Source) 파악 및 범주화 연구로 새로운 유형의 시도들이 활발히 이루어지고 있다고 하였다.

Lee et al.[9]의 연구에서는 특허문서의 기술 동향 패턴을 파악하여 특정 제품과 기술에 관한 진화 패턴을 발견하고 Kim et al.[10]은 텍스트마이닝 기법을 활용하여 신문사별 기사 내용 및 논조의 차이를 분석하였다.

1.2 연구의 목적

최근 연구 방법론으로 빅데이터 분석이 주목받고 있는데, 이는 실시간에 가까운 속도로 생성되는 데이터의 기하급수적인 양적 증가 및 사용가능한 데이터 형태의 다양화를 해결하여 데이터의 가치를 만들어낼 수 있기 때문이다[[11]. 현재까지, 대부분 문서와 주제에 잠재되어 있는 핵심 용어들을 파악한 연구들이 많지만, 전통 식품 관련 이슈들의 메시지를 파악한 연구는 찾아보기 힘들다.

본 연구는 전통 식품 중 '전통 장류'에 대한 빅데이터를 분석하여 기존 이슈 도출 및 새로운 시각을 제시하는

데 타 연구들과 차별성이 있다. 이에 따라 전통 장류 기사 DB를 바탕으로 텍스트마이닝 기법을 활용하여 키워드 분석 및 연구 방향을 제시하고자 한다.

논문의 구성은 다음과 같다. 제2장에서는 한국의 장류 생산 및 소비현황, 전통 장류 관련 언론 기사 자료 등을 파악하고, 제3장에서는 분석 방법 및 자료에 대한 특징을 정리하여 계량적 분석모형을 설명한다. 제4장에서는 한글 형태소 분석을 위한 KoNLP 패키지를 활용하여 TF-IDF(Term Frequency-Inverse Document Frequency), 워드클라우드(Wordcloud), 자연어 사회연결망 분석(Social Network Analysis) 등의 분석 결과를 정리하고, 마지막 5장에서는 연구 결과를 요약하고 시사점을 제시한다.

2. 국내 장류 시장동향 및 언론 기사 현황

2.1 국내 장류 생산 및 소비현황

최근 소비자의 생활양식 변화와 가정 제조 장류의 소비가 감소 되고 대기업의 장류 산업 참여가 확대되면서 시판제품의 소비가 증가하고 있다<Table 1>.

'21년 국내의 장류 시장 규모를 살펴보면 생산량 72만 2,366톤, 생산액 9,027억원이며, 판매량과 판매액은 각각 57만 8,449톤, 1조 1,989억원으로 나타났다. 그중 '한식 장류'의 판매 점유율은 35%(4,195억원)로 품목별로는 '고추장' 22.4%(2,683억원), '청국장' 6.7%(801억원), '한식 된장' 2.2%(262억원), '한식 간장' 2.2%(264억원), '한식 메주' 1.5%(184억) 순으로 나타났다. 또한, 수출량과 수출액은 각각 4만 5,086톤, 8,566만 달러이며, '고추장', '혼합 장', '혼합간장' 순으로 높게 나타났다.

장류의 최근 5년간 판매 점유율을 살펴보면, '고추장', '혼합 장', '혼합간장' 순으로 높은 비중을 차지하고 있으며, '개량 메주', '한식 간장', '청국장'은 지속적으로 판매액이 증가하고 있다. 전년 대비 증가율에서는 '한식 된장', '개량간장류'의 감소가 크고 '메주'의 판매가 증가한 것으로 보아 가공식품 구매보다 가공재료를 구매하여 직접 제조하는 트렌드가 형성되고 있는 것으로 보인다.

Table 1. '21 Korea sauces Production and Sales Status
(Unit : Ton, 1 Million Won, \$)

| Items | Production volume | production value | domestic sales volume | domestic sales | share (%) | export volume | export value |
|---|-------------------|------------------|-----------------------|----------------|-----------|---------------|--------------|
| Korean traditional fermented soybean lump | 2,896 | 13,285 | 2,217 | 18,422 | 1.54 | - | - |
| improved fermented soybean lump | 2,360 | 10,527 | 2,073 | 13,842 | 1.15 | - | - |
| Korean traditional soy sauce | 9,214 | 18,421 | 5,044 | 26,447 | 2.21 | 389 | 823 |
| fermented soy sauce | 69,051 | 74,452 | 57,114 | 143,824 | 12 | 2,477 | 4,018 |
| acid hydrolyzed soy sauce | 62,299 | 35,647 | 52,044 | 35,013 | 2.92 | 1,717 | 988 |
| mixing soy sauce | 168,357 | 118,691 | 150,138 | 192,436 | 16.05 | 7,504 | 7,094 |
| Korean traditional Soybean Paste | 2,667 | 29,191 | 2,205 | 26,231 | 2.19 | 22 | 289 |
| Soybean Paste | 119,814 | 115,534 | 73,693 | 128,525 | 10.72 | 4,912 | 7,631 |
| Red Chili Paste | 148,359 | 228,734 | 108,668 | 268,294 | 22.38 | 20,587 | 47,876 |
| chun jang | 18,734 | 22,709 | 16,554 | 26,201 | 2.19 | 721 | 1,171 |
| Rich Soybean Paste Stew | 13,097 | 65,709 | 13,709 | 80,092 | 6.68 | 130 | 402 |
| mixed soy paste | 104,727 | 166,241 | 94,330 | 234,912 | 19.59 | 6,580 | 15,207 |
| Other sauces | 791 | 3,602 | 660 | 4,669 | 0.39 | 47 | 162 |
| total | 722,366 | 902,743 | 578,449 | 1,198,908 | 100 | 45,086 | 85,661 |

Source : Food and other production performance, MFDS, 2021

Table 2. Korea sauces Sales increase/decrease rate
(Unit : 1 Million Won, %)

| Items | '17 | '18 | '19 | '20 | '21 | year-on-year increase rate | GA GR |
|---|---------|-----------|-----------|-----------|-----------|----------------------------|-------|
| Korean traditional fermented soybean lump | 17,787 | 16,667 | 16,939 | 16,971 | 18,422 | 8.55 | 0.9 |
| improved fermented soybean lump | 5,996 | 7,414 | 8,631 | 8,513 | 13,842 | 62.6 | 23.3 |
| Korean traditional soy sauce | 20,632 | 22,421 | 23,436 | 25,580 | 26,447 | 3.39 | 6.4 |
| fermented soy sauce | 113 | 120,971 | 118,430 | 149,659 | 143,824 | -3.9 | 6.3 |
| acid hydrolyzed soy sauce | 35,444 | 40,218 | 39,905 | 40,219 | 35,013 | -12.94 | -0.3 |
| mixing soy sauce | 147,833 | 178,420 | 178,512 | 190,076 | 192,436 | 1.24 | 6.8 |
| Korean traditional Soybean Paste | 20,654 | 23,595 | 23,483 | 30,838 | 26,231 | -14.94 | 6.2 |
| Soybean Paste | 129,961 | 123,016 | 122,429 | 123,247 | 128,525 | 4.28 | -0.3 |
| Red Chili Paste | 219,401 | 226,445 | 220,639 | 250,404 | 268,294 | 7.14 | 5.2 |
| chun jang | 27,900 | 24,156 | 25,338 | 27,664 | 26,201 | -5.29 | -1.6 |
| Rich Soybean Paste Stew | 56,400 | 57,565 | 67,663 | 73,805 | 80,092 | 8.52 | 9.2 |
| mixed soy paste | 188,061 | 198,580 | 220,872 | 223,644 | 234,912 | 5.04 | 5.7 |
| Other sauces | 7,588 | 5,541 | 5,367 | 4,799 | 4,669 | -2.71 | -11.4 |
| total | 877,770 | 1,045,009 | 1,071,644 | 1,165,419 | 1,198,908 | 2.87 | 4.9 |

Source : Food and other production performance, MFDS, 2021

최근 5년간 장류(4.9%)의 전체 판매액의 연평균증가율은 증가세를 유지하고 있으나, 판매 점유율 대비 '메주'(23.3%)와 '청국장'(9.2%)의 소비가 가장 많이 늘었고 기타 장류(-11.4%)와 '춘장'(-1.6%)의 소비가 가장 저조하게 나타났다(Table 2).

'21년 전체 장류업체 수는 총 2,003개소로 '17년 대비 2.9% 감소하였으며, 매년 지속적으로 줄고 있다. 현

재 지역별 장류업체 수는 '경기', '경북', '전남' 순으로도 단위 위주로 업체들이 많이 형성되어 있고 '울산' 등 광역시의 업체 수 분포는 적지만, 그중 '인천'과 '광주' 업체 수는 5년간 보합세로 조사되었다.

Table 3. Current status of Korea sauces manufacturer by region
(Unit : ea)

| Sort | '17 | '18 | '19 | '20 | '21 |
|-----------|-----|-----|-----|-----|-----|
| Seoul | 29 | 28 | 28 | 23 | 21 |
| Busan | 41 | 43 | 42 | 38 | 39 |
| Daegu | 50 | 48 | 48 | 44 | 41 |
| Incheon | 32 | 33 | 38 | 39 | 36 |
| Gwangju | 23 | 23 | 22 | 22 | 23 |
| Daejeon | 22 | 20 | 18 | 16 | 17 |
| Ulsan | 14 | 12 | 10 | 8 | 9 |
| Sejong | 7 | 9 | 11 | 10 | 11 |
| Gyeonggi | 272 | 257 | 270 | 287 | 299 |
| Gangwon | 149 | 146 | 152 | 148 | 145 |
| Chungbuk | 167 | 168 | 161 | 152 | 148 |
| Chungnam | 176 | 167 | 170 | 171 | 181 |
| Jeonbuk | 266 | 266 | 256 | 257 | 247 |
| Jeonnam | 255 | 266 | 256 | 257 | 247 |
| Gyeongbuk | 316 | 303 | 288 | 280 | 278 |
| Gyeongnam | 229 | 234 | 244 | 239 | 224 |
| Jeju | 14 | 23 | 30 | 26 | 25 |

Source : Food and other production performance, MFDS, 2021

지역별 장류 생산량은 '충남'(175,047톤), '경기'(163,504톤), '전북'(158,633톤) 순이나, 생산액은 '전북'(2,197억원)이 가장 많으며, 국내 판매는 '경기도'(2,945억원)가 가장 높은 24.6%의 점유율을 나타내고 있다.

수출은 '전북'(315억원), '충남'(291억원), '경기'(101억원) 순으로 각각 지역별로 비교해보면 생산·투자·국내·수출 판매지역으로 나누어 설명할 수 있으나, 3지역으로 한정되어 있다.

'20년 우리 식탁에서 볼 수 있는 100가지 다빈도 식품 중 장류 순위는 8위 '쌈장'(15.7%), 17~18위 '된장 국·찌개'(22.8%), 31위 '고추장'(7.7%), 54위 '간장'(5.04) 순으로 조사되었으며, 장류의 품목별 1일 평균 섭취량은 '간장'(6.0g), '고추장'(5.5g), '된장'(4.0g), '쌈장'(2.1g), '청국장'(0.8g) 순으로 나타났다[12]. 하지만, 5년간 양념장류의 1일 평균 섭취량은 전체적으로 오름과 내림이 반복적이거나 최근 감소추세로 전환되었다. 장류의

소비는 식생활 변화에 따라 전체적으로 감소하는 것을 알 수 있는데, 마요네즈, 케첩 등 서양식 소스가 된장과 간장을 대체하고 1인 가구 증가와 여성의 경제활동 참여 확대에 장류의 소비 및 섭취량은 지속적으로 감소할 것으로 보인다.

Table 4. Rice Consumption in the Industrial Sector
(Unit : Ton, 1 Million Won, \$)

| Sort | Production volume | Production value | domestic sales volume | domestic sales | share (%) | export volume | export value |
|-----------|-------------------|------------------|-----------------------|----------------|-----------|---------------|--------------|
| Seoul | 32 | 210 | 31 | 247 | 0.02 | - | - |
| Busan | 21,307 | 25,044 | 19,615 | 41,728 | 3.48 | 882 | 858 |
| Daegu | 31,118 | 30,377 | 23,609 | 32,669 | 2.72 | 65 | 302 |
| Incheon | 2,592 | 3,695 | 2,548 | 4,651 | 0.39 | - | - |
| Gwangju | 446 | 1,192 | 431 | 1,975 | 0.16 | - | - |
| Daejeon | 46 | 179 | 45 | 234 | 0.02 | - | - |
| Ulsan | 38 | 214 | 34 | 374 | 0.03 | - | - |
| Sejong | 2,042 | 2,852 | 2,003 | 3,245 | 0.27 | - | - |
| Gyeonggi | 163,504 | 178,450 | 146,388 | 294,507 | 24.56 | 9,049 | 10,105 |
| Gangwon | 19,752 | 24,607 | 10,432 | 32,850 | 2.74 | 4 | 36,587 |
| Chungbuk | 21,145 | 59,476 | 18,338 | 63,802 | 5.32 | 2,820 | 5,481 |
| Chungnam | 175,047 | 206,495 | 111,024 | 273,100 | 22.78 | 12,950 | 29,116 |
| Jeonbuk | 158,633 | 219,735 | 126,255 | 277,306 | 23.13 | 15,266 | 31,474 |
| Jeonnam | 35,456 | 46,058 | 31,987 | 41,763 | 3.48 | 2,657 | 6,460 |
| Gyeongbuk | 19,107 | 44,584 | 18,137 | 55,278 | 4.61 | 267 | 604 |
| Gyeongnam | 72,007 | 58,481 | 67,476 | 73,989 | 6.17 | 1,124 | 1,227 |
| Jeju | 96 | 1,094 | 95 | 1,190 | 0.10 | - | - |
| total | 722,368 | 902,743 | 578,448 | 1,198,908 | 100 | 45,084 | 122,214 |

Source : Food and other production performance, MFDS, 2021

Table 5. Korea sauces average daily intake
(unit : g)

| Sort | soy sauce | Red Chili Paste | Soybean Paste | ssamjang | Rich Soybean Paste Stew |
|------|-----------|-----------------|---------------|----------|-------------------------|
| '16 | 6.35 | 5.49 | 4.33 | 2.19 | 0.93 |
| '17 | 7.8 | 6.44 | 3.96 | 2.04 | 0.63 |
| '18 | 6.46 | 6.16 | 4.22 | 1.84 | 0.8 |
| '19 | 6.51 | 6.4 | 4.21 | 2.11 | 0.81 |
| '20 | 6.01 | 5.51 | 3.96 | 2.08 | 0.78 |

Source : National Nutrition Statistics, KKIDI, 2020

2.2 전통 장류 관련 언론 기사 현황

'17년 1월 1일부터 '22년 8월 15일까지 5년 8개월간 언론사(전국일간지(11), 경제일간지(8), 지역일간지(28), 방송사(5), 전문지(2))를 통한 전통 장류 연관 기사는 총 1,130건으로, '17년도(312건)는 “강진 전통 장류 설 선물 인기”, “순창 장류 산업 소스 시장 개척”, “초등·중학교 학교급식 전통 고추장·된장 시범 공급” 등의 기사, '18년도(237건)는 “순창 전통 장류 전문가과정 교육과정”, “학교급식 전통 장류 공급”, “장 담그기 체험” 등의 기사, '19년도(201건)는 “순창군 전통 장류 제조기업 지원”, “로컬푸드 마켓 직거래” 등의 기사, '20년도(131건)는 “강진 전통 장류 인기”, “전통 장류 6차 산업 모델 조성”. “전통 장류 신기술 보급사업” 등의 기사, '21년도(161건)는 “도시민 장독대 분양”, “청국장 패혈증·급성위염 억제 효과” 등의 기사, '22년도(88건)는 “식품진흥원 소스·전통 장류 지원 사업”, “순창군 전통 장류 종균 상품화” 등의 기사들 위주로 보도되었으며, 전통 장류 관련 기사 건수는 '17년 이후 지속적으로 줄어들고 있다. 또한, 대부분의 전통 장류 관련 언론 기사가 지역일간지 중 ‘전북도민일보’(121건), ‘충도일보’(116건), ‘대전일보’(57건)에 많이 보도된 것으로 보아 ‘전북’과 ‘충청’ 지역의 전통 장류 이슈들이 많은 것을 알 수 있다.

Table 6. Korea sauces Number of articles by year and media

| Year | Korea sauces Article | Number of articles by media |
|------------|----------------------|--|
| '17 | 312 | Chonbuk Domin Ilbo(38), JOONGDO(35), Jeonbuk(23), Deajeon(21), other(195) |
| '18 | 237 | JOONGDO(29), Chonbuk Domin Ilbo(24), Jeonbuk(15), Chungcheong Today(14), other(155) |
| '19 | 201 | JOONGDO(29), Chonbuk Domin Ilbo(18), Joong-Bu Maeil(11), ChungCheong(11), other(132) |
| '20 | 131 | Chonbuk Domin Ilbo(12), JOONGDO(9), Chonnam(8), Kwangju Maeil(7), other(95) |
| '21 | 161 | Chonbuk Domin Ilbo(17), ChungCheong(11), other(113) |
| '22. 08.15 | 88 | Chonbuk Domin Ilbo(12), Jeonbuk(9), Chungbuk(6), Deajeon(6), other(55) |

Source : Naver Portal Sites, 2022.08.15.

3. 분석 방법 및 자료

3.1 텍스트마이닝 분석 방법

3.1.1 TF-IDF 분석

텍스트마이닝 연구는 다량의 텍스트 수집 후 불용어 처리, 형태소 분석 등 전처리 과정을 거쳐 2~10개의 음절로 구성된 명사를 추출한 후 특정 문서 내 특정 단어가 얼마나 자주 등장하는지를 의미하는 단어 빈도수(TF)와 전체 문서에서 특정 단어가 얼마나 자주 등장하는지를 의미하는 역문서 빈도(IDF)를 통해 다른 문서에서는 등장하지 않지만, 특정 문서에서만 자주 등장하는 단어를 찾아 문서 내 가중치를 계산하는 TF-IDF 방식을 활용하여 문서 핵심어 추출, 유사도, 중요도를 정하는 핵심 단어를 선정할 수 있다[13].

이와 같은 선택과정은 비정형데이터인 텍스트를 정형 데이터로 변환하여 설명될 수 있다. TF-IDF는 개별문서 (j)에서 나타난 단어 (i)의 출현 빈도수 TF_{ij} 와 단어 (i)를 포함한 문서 수에 전체 문서 수(N)를 나누고 \log 를 취한 $\log(\frac{N}{df_i})$ 값의 곱으로 구할 수 있다.

$$TF-IDF = TF \times IDF = TF_{ij} \times \log(\frac{N}{df_i}) \quad (1)$$

핵심 단어란 전통 장류 관련 기사 본문을 대표할 수 있는 주요 단어를 의미하며, 수집된 텍스트의 특징을 나타내는 단어들이며, 전처리 후 추출된 단어를 전부 사용하는 방법도 있지만, 전체 내용을 요약하여 표현할 수 있다면, 정제된 핵심 단어만 사용해도 무방하다.

핵심 단어 선정 시 주의할 점은 빈도수만 계산하여 출현 빈도가 높은 단어를 핵심 단어로 선정하는 오류에 빠질 수 있다. 그런 이유로 TF-IDF를 이용하여 문서 내 단어 중요도를 고려하였으며, TF-IDF 누적 비율과 연구자의 정성적 판단을 포함하여 누적 점유율 30%를 차지하는 단어를 선정하였다. 또한, 관련 용어 수집에서 가중치를 부여하기 위해 키워드 중심으로 주변 용어들을 포함하는 Snippet set의 용어 즉, “2022년 우수 장류 제품”이란 문장에서 장류가 키워드라면 앞뒤 1개의 단어를 Snippet으로 만들면 “우수 장류 제품”으로 추출한 후 TF, IDF를 구하였다.

3.1.2 워드클라우드(Wordcloud) 생성

연구 대상 자료의 제목과 기사 분석을 위한 기초 작업

인 텍스트 전처리로서 명사 키워드를 추출하고 출현 빈도수를 계산하여 워드클라우드 기법으로 시각화하여 분석하는 것이다. 출현 빈도수에 따라 각각 단어의 크기가 표시되고 색상 또한 다르게 표시된다. 출현 빈도수가 높은 단어는 중요도나 관심도가 높은 것을 의미한다.

R 프로그램에서 빈도수 계산 결과 처리는 다음과 같은 5가지 경우로 구분되며, 5가지 유형 중 (2), (4)번만 수렴하여 각 단어의 빈도수를 직관적으로 시각화한다.

- (1) 빈도수가 높고 중요성이 낮은 단어
- (2) 빈도수가 높고 중요성이 높은 단어
- (3) 빈도수도 낮고 중요성이 낮은 단어
- (4) 빈도수는 낮고 중요성이 높은 단어
- (5) 빈도수는 높으나 부적격 값(조사, 감탄사, 동사, 등)

3.1.3 자연어 사회연결망 분석

사회연결망 분석은 다양한 키워드들의 동시 출현 분석 기법에서 시작되어 단어 간 의미 또는 맥락상 연결 관계 및 해당 자료를 시각화하거나 중요 개념을 네트워크 속의 중심성(Centrality)에 따라 추출하는 방법으로 객체 간의 관계를 분석하는데 효과적이다.

연결망 분석에서는 점(Node) 사이의 관계를 선으로 표현함으로써 데이터만으로 파악하기 어려운 연결 구조를 한눈에 파악할 수 있도록 도와주며, 도출된 네트워크 구조 간의 차이를 비교함에 있어 활용되는 지표는 개별적 노드의 국소적(Local) 특성을 나타내는 연결선 수(Degree)와 중심성(Centrality), 네트워크 전역적(Global) 특성을 나타내는 밀도(Density) 등이 있다. 연결선 수와 관련된 연결 정도는 한 점과 맺고 있는 결점의 숫자로 정의된다[14].

연결망 분석에서 가장 많이 사용되는 중심성(Centrality) 지수는 중심에서 위치하는 정도를 표현하는 지표로써 연결 정도와 중요도 등을 알 수 있으며, 신문 기사에 등장하는 주요 단어의 사용 빈도, 연결 상태, 네트워크 연결 구조를 통해 의도적인 의미를 파악하는데 유용한 방법론으로 활용되고 있다.

지금까지 우리의 식문화에서 전통 식품 관련 정보는 유기적인 관계가 있음에도 불구하고, 현실에서는 사용자들이 직접 관련성을 조사하거나 분석에 대한 접근이 용이하다고 말하기 어렵다. 일반적으로 사용자들은 단순 정보 검색으로만 궁금증을 해소하거나, 특정 정보를 찾아내기까지 많은 시행착오를 거쳐야 하는 어려움이 있다. 따라서 본 연구에서는 전통 장류가 매 순간 노출되고

있는 인터넷 뉴스 기사 분석을 통해 관련 정보를 패키징하여 보여줄 수 있는 기반을 마련하고자 한다.

3.2 언론 기사 분석 자료

본 연구에 이용할 자료는 우리나라 대부분의 언론사(55개)로부터 뉴스를 제공받아 게시하는 네이버(Naver)에서 2017년 1월 1일부터 2022년 8월 15일까지의 '전통 장류' 기사를 크롤링하여 데이터베이스화하였다. 특히, 인터넷 뉴스 포털 사이트의 경우, 주요 언론사의 기사를 취합하여 재공급하는 특성이 있기에 특정 매체의 시각에서 편향되지 않는 이슈를 제공한다는 장점이 있다. 분석 자료는 '전통 장류' 단어가 거론된 전체 언론 기사 1,130건과 중복 기사를 제외한 언론 기사 628건이다. 모든 분석 과정은 2016 엑셀과 분석 패키지 R(4.0.2)을 활용하여 결과를 도출하였다.

4. 분석 결과

4.1 전통 장류의 관심도 분석

전통 장류 주제와 관련하여 일정 기간 작성된 기사량은 사회적 관심도를 나타낸다고 볼 수 있다. 이러한 관점에서 연·월 단위로 수집된 전통 장류 기사량을 확인함으로써 기간별 이슈를 간단하게 알아볼 수 있다.

'17년(306건)부터 '21년 8월(87)까지 5년 8개월간 전통 장류의 기사 건수는 매년 감소추세이다. 각 연도별 기사의 주요 이슈를 살펴보면, '17년도 전통 장류 기사량이 많은 1월(46건)과 9월(39건)의 주요 기사는 '명절선물', '전통 장류 학교급식 추진', '장류 소스화 사업 시작' 등, '18년도 11월(36건)은 '제주 정관장물 전통식품 판매', '해남 농축수산물(전통장류) 국내의 판매', '초·중학교 충남 콩을 이용한 전통 장류 보급' 등, '19년도 8월(24건)은 '영양 고추 HOT 페스티벌', '전라북도 온·오프라인 한가위 큰 장터', '세종시교육청 조리 아카데미 직무연수' 등, '20년도 11월(29건)은 '경북 6차 산업 안테나 숲', '영산대 지역 우수 전통 장류 판로개척 지원', '전남 식품산업연구소 연구센터 종균 된장 제조법 이전' 등, '21년도 3월(23건)은 '순창 청국장 패혈증·급성위염 억제 효과', '충남 초·중·고 무상 급식 전통 장류 공급 사업', '공주시 농·특산물 수출 교육' 등, '22년도 3월(21건)은 '순창군 토종 청국장용 종균 상품화 및 만화로 보는 전통 장류 제조관리', '괴산군 전통 장류 시설·장비 개선 경쟁력 강

화', '우리나라 최초 장류 품종 파주 장단콩' 등의 기사가 많이 보도되었다.

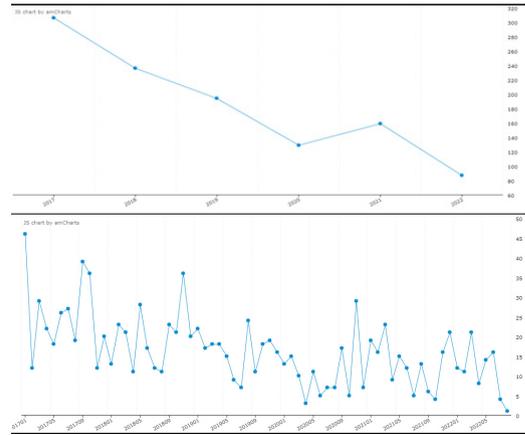


Fig. 1. Korea sauces Keyword Number of articles by year and media

4.2 전통 장류 관련 언론 기사의 키워드 분석

본 연구의 전체 분석 대상인 1,130건의 기사에서 추출된 용어의 빈도수가 10회 이상 출현한 단어 중 분석에서 불필요하고 무의미한 단어나 중의적 단어를 제외하고 토픽랭크 알고리즘에 기반하여 가중치를 반영한 상위 10개의 단어를 선정하였다. <Table 7>과 같이 전통 장류 기사에서 가장 큰 가중치를 차지한 키워드는 순창군, 청국장, 우수성 순이며, 빈도수가 가장 많은 키워드는 강진으로 나타났다.

순창군 기사는 141건으로 '장류 매뉴얼 발간', '발효 테마파크 조성', '전통 장류제조사 교육', '전통 장류 소스화', '전통 장류 기능성 연구 성과', '전통 장류 기업 지원', '발효 종균 보급', '순창장류축제' 등의 내용들을 주로 다루었다.

청국장 기사는 118건이며, '대장염, 뇌 건강, 당뇨 개선 효과', '종균상품화 성공', '패혈증, 급성위염 효과', '파주 장단콩 체험', '장류 축제' 등의 기사에 검색어로 많이 사용되었다.

'전통 장류'에서 우수성 관련 키워드가 많이 들어간 기사는 145건으로 '장류의 효능', '우수한 농·특산물', '식품박람회 참가', '농·특산물 홍보 판매전' 등 전통 장류의 효과와 홍보에 연관되는 형용사로 많이 사용되었다.

Table 7. Korea sauces Ranking of related word weights

| Keyword | Weighted value(%) | frequency |
|---------------|-------------------|-----------|
| Sunchang-gun | 87.56 | 391 |
| cheonggukjang | 44.07 | 362 |
| superiority | 36.22 | 165 |
| Jangdokdae | 34.29 | 199 |
| school lunch | 28.65 | 522 |
| shiitake | 28.22 | 57 |
| middle school | 27.51 | 82 |
| consumer | 24.22 | 405 |
| Gangjin | 21.91 | 565 |
| soy sauce | 21.42 | 53 |

연구 대상 검색 결과 중 분석 기사와 연관성이 높은 키워드를 워드클라우드 기법으로 중요도나 관심도가 높은 단어 순으로 크기와 색을 다르게 표시하여 2차원 평면으로 시각화하였다.



Fig. 2. Word Cloud for Keywords

전통 장류 1,130건 중 중복 기사를 제외한 628건에서 추출된 상위 10개의 키워드는 순창군, 강진군, 중학교 순으로 나타났다.

강진군 기사는 53건으로 '전통 장류 명절선물', '건강 10대 농식품', '브랜드 초록민음', '전남 6차 산업', '찾아가는 직거래장터' 등의 기사가 많이 보도되었다.

중학교 기사는 46건이며, 그중 중학교와 연관된 기사는 '충남도내 학교급식 전통 장류 공급' 기사가 독보적으로 많으며, 공주시, 아산시 등의 기사에서도 같은 맥락으로 많이 사용되었다.

'장독대' 키워드가 들어간 기사는 79건으로 주로 '장담고기 체험', 저장 공간의 내용물이 많으며, 괴산군에서 추진하는 장 담그기의 일환으로 '장독대 분양'이란 키워드의 기사가 많이 보도되었다. '청국장' 또한, 118건의 기사가 보도되었는데 건강, 종군 상품화, 전문가 양성 교육, 기능 및 효능, 전통 식품 등의 주제에서 많이 사용되

었는데 기본적으로 전통 장류에서 발효나 장류로 연결되는 내용에서는 '청국장' 단어가 가장 많이 사용되는 것을 알 수 있다.

Table 8. Ranking of word weights related to the exclusion of duplicate articles from Korea sauces

| Keyword | Weighted value(%) | frequency |
|---------------|-------------------|-----------|
| Sunchang-gun | 32.08 | 108 |
| Gangjin-gun | 25.71 | 107 |
| middle school | 24.44 | 60 |
| school lunch | 21.47 | 344 |
| Jangdokdae | 21.47 | 132 |
| cheonggukjang | 20.67 | 161 |
| soy sauce | 19.37 | 47 |
| shiitake | 16.88 | 30 |
| superiority | 16.72 | 88 |
| MAFRA | 15.56 | 40 |

중복 기사를 제외하여 분석한 결과도 '순창군' 키워드의 가중치 값이 가장 높으나, 빈도수는 '학교급식'이 가장 많은 것으로 나타났다. 또한, '순창군' 키워드를 제외한 '강진군', '중학교', '학교급식', '장독대', '청국장' 키워드의 가중치 값 차이도 크지 않은 것으로 보아 전통 장류의 연관 키워드로 많이 거론된 것을 알 수 있다.



Fig. 3. WordCloud for keywords excluding duplicate articles

4.3 전통 장류 관련 언론 기사 사회연결망 분석

사회 연결망 분석을 위해 전통 장류 기사 본문의 형태소를 분석하여 명사 상당어구를 추출한 후 명사 상당어구에 개체명 분석 알고리즘을 적용하였다. 개체명 가중치는 전통 장류 기사 2건으로 부여한 후 정확도 상위 100건의 분석 뉴스에서 추출된 개체명(인물, 지역, 기관,

키워드) 사이의 연결 관계를 색깔별로 나누어 네트워크 형태로 시각화하였으며, 전통 장류와 연관된 인물, 지역, 기관, 키워드 개체명별로 각각 비중이 높은 키워드들의 연관성 분석을 하였다.

먼저 [인물]은 이00 충청남도 농산물유통과장(13건)이 가장 많이 거론되었으며, 연관 키워드는 영양교사, 디딤돌, 초·중학생, 학부모, 농산물유통과장이다.

[지역] 보은군(7건)의 연관 키워드는 장외면, 산외면, 교육생, 생활지원팀, 보은군농업기술센터 안00 지도사, 박00 소장, 충북 3번째 전통 장류 식품명인 조00 명인, 청주시 문화동, 이00 지사, 임00 보은 부군수, 이00 충북도 농업정책과장이다.

[기관] 농촌진흥청(8건)의 연관 키워드는 공주, 청양, 연구사, 농산물가공연구회원, 비즈니스, 농림축산식품부, 진천군 한식 양념 소스 4종 개발 만디식품 권00 대표이다.

[키워드]는 학부모(15건), 농산물유통과장(13건), 코로나19(13건) 순으로 가장 많은 비중을 차지하였으며, 학부모의 연관 키워드는 이00 농산물유통과장, 영양교사, 디딤돌, 교직원, 수혜학생, 체육건강과장, 공주시, 국립농산물품질관리원, 조리사, 여00 충남농협 본부장, 김00 교육감, 시스템이다. 농산물유통과장의 연관 키워드는 이00, 초·중학생, 학부모, 영양교사, 디딤돌, 레시피 대회, 시스템, 교직원, 흥성이며, 코로나19의 연관 키워드는 김00 공동위원장, 양주시, 괴산군 청천면, 진주시, 임00 농림축산과장, 감물면, 당진시 김00 농업정책과장이다.

<Fig. 5>의 개체명(인물, 지역, 기관, 키워드)별 빈도는 점의 크기와 색깔로 구분하였으며, 점에 이어진 선이 연관 키워드로 볼 수 있다. 즉, 전통 장류 기사에서 출현 빈도가 높은 개체명을 알아보고 연관 키워드별로 교차한 결과를 사회 연결망 분석을 통해 쉽게 파악할 수 있다.



Fig. 4. Word Cloud for the object names(Person, Area, Institution, Keywords) Related to Korea Sauces

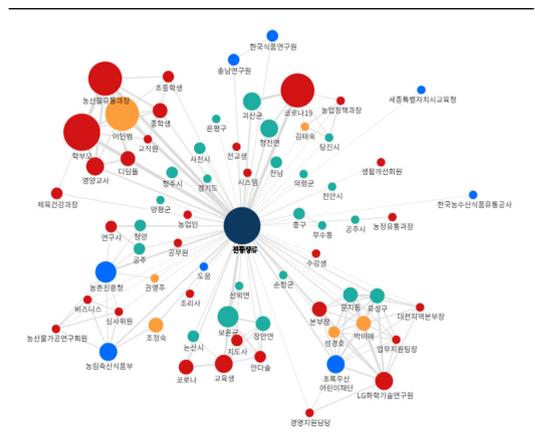


Fig. 5. Social Network Analysis of object names(Person, Area, Institution, Keywords) Related to Korea Sauces

5. 결론

최근 소비자의 생활양식 변화와 가정 제조 장류의 소비가 감소되면서 대기업의 장류 제품의 소비가 증가하고 있는 실정이다. '21년 국내의 장류 시장 규모는 생산량 72만 2,366톤, 생산액 9,027억원이며, 판매량과 판매액은 각각 57만 8,449톤, 1조 1,989억원으로 나타났다. 그중 '한식 장류'의 판매 점유율은 35%(4,195억원)로 품목별로 '고추장' 22.4%(2,683억원), '청국장' 6.7%(801억원), '한식 된장' 2.2%(262억원), '한식 간장' 2.2% (264억원), '한식 메주' 1.5%(184억) 순으로 나타났다. 또한, 수출량과 수출액은 각각 4만 5,086톤, 8,566만 달러이며, '고추장', '혼합장', '혼합간장' 순으로 높게 나타났다.

최근 전통 장류와 관련하여 농업과 연계를 위한 전통 식품 관련 시설 지원, 장류산업진흥법 제정, 장 발효기술 및 소재화·소스와 연구 개발, 수출시장 개척 등 다양한 토끼들이 언론에서 보도되고 있는데 이런 광범위한 자료들을 대상으로 빅데이터 분석을 통해 전통 장류의 가치와 시사점을 파악하였다.

본 연구에서는 '전통 장류'에 관련한 핵심 용어의 분석을 위해 네이버포털사이트에서 2017년부터 2022년 8월 15일까지 보도된 언론 기사를 크롤링하여 1,130건을 수집하였으며, 해당 기사들에서 핵심 키워드를 추출하는 분석을 수행하였다. 우선 1,130건의 기사에서 가장 빈도가 높고, 가중치가 높은 전통 장류 관련 상위 10개의 키워드를 도출하였다. 도출된 키워드를 대상으로 전통 장류 관련 언론 기사의 특징을 살펴보고 10개의 핵심키워

드 간 관계를 연관관계 분석을 통해 수치화 및 시각화하였다. 연관관계 분석 결과 전통 장류 키워드를 가장 큰 허브로 '순창군'과 '청국장', '우수성' 순의 키워드 그룹이 형성되는 구조를 보였다. 순창군 기사(141건)에서는 '장류 매뉴얼 발간', '발효테마파크 조성', '전통 장류 제조사 교육', '전통 장류 소스화', '전통 장류 기능성 연구 성과', '전통 장류 기업 지원', '발효 중균 보급', '순창 장류 축제' 등의 내용들을 주로 다루었다.

청국장 기사(118건)에서는 '대장염, 뇌 건강, 당뇨 개선 효과', '중균상품화 성공', '폐혈증, 급성위염 효과', '파주 장단콩 체험', '장류 축제' 등의 기사에 키워드가 많이 사용되었다.

전통 장류에서 우수성 키워드가 많이 들어간 기사는 145건으로 '장류의 효능', '우수한 농·특산물', '식품박람회 참가', '농·특산물 홍보 판매전' 등 전통 장류의 효과와 홍보 연관 형용사로 많이 사용되었다. 이는 지금까지 전통 장류 관련 기사는 다양한 주제를 다루었지만, 분석 결과 '순창군', '청국장' 및 '우수성' 키워드가 가장 많이 출현하고 있음을 보여준다.

또한, 중복 기사를 제외한 628건 기사들의 키워드를 분석하였다. 전체 기사의 경우 전통 장류의 지역, 품목, 형용사 관련 키워드들이 추출되었다면, 중복 기사를 제거한 후 주요 키워드는 순창군, 강진군, 중학교 키워드가 상위 에 랭크 되었다. 강진군 기사(53건)는 '전통 장류 명절선물', '건강 10대 농식품', 브랜드 초록민음, '전남 6차 산업', '찾아가는 직거래장터' 등의 기사가 많이 보도되었으며, 중학교 기사(46건)는 초·중학교와 연관된 '충남도내 학교급식 전통 장류 공급' 기사가 압도적으로 많으며, 공주시, 아산시 등의 기사에서도 같은 맥락으로 많이 사용되었다.

중복 기사를 제외한 경우의 분석 결과도 '순창군' 키워드의 가중치 값이 가장높았으나, 빈도수는 '학교급식'이 가장 많은 것으로 나타났다. 또한, '순창군' 키워드를 제외한 '강진군', '중학교', '학교급식', '장독대', '청국장' 키워드의 가중치 값 차이도 크지 않은 것으로 보아 전통 장류의 연관 키워드로 많이 거론된 것을 알 수 있다.

지역명 키워드는 지자체의 정책 지원들과 연관되었음을 유추할 수 있으며, 전통 장류 관련 전체 기사와 중복 기사를 제외한 키워드 동향은 차이가 있는 것으로 나타났다.

전통 장류 검색 개체명(인물, 지역, 기관, 키워드)의 연관성을 분석한 결과 [인물]은 이00 충청남도 농산물유통과장(13건)이 가장 많이 거론되었으며, 연관 키워드는 영양교사, 디딤돌, 초·중학생, 학부모, 농산물유통과장이다.

[지역]은 보은군(7건)의 연관 키워드는 장안면, 산외면, 교육생, 생활지원팀, 보은군농업기술센터 안00 지도사, 박00 소장, 충북 3번째 전통 장류 식품명인 조00 명인, 청주시 문화동, 이00 지사, 임00 보은 부군수, 이00 충북도 농업정책과장이다.

[기관]에서는 농촌진흥청(8건)의 연관 키워드는 공주, 청양, 연구사, 농산물가공연구회원, 비즈니스, 농림축산식품부, 진천군 한식 양념 소스 4종 개발 만디식품 권00 대표이다.

[키워드]는 학부모(15건), 농산물유통과장(13건), 코로나19(13건) 순으로 가장 많은 비중을 차지하였으며, 학부모의 연관 키워드는 이00 농산물유통과장, 영양교사, 디딤돌, 교직원, 수혜학생, 체육건강과장, 공주시, 국립농산물품질관리원, 조리사, 여00 충남농협 본부장, 김00 교육감, 시스템이다. 농산물유통과장의 연관 키워드는 이00, 초·중학생, 학부모, 영양교사, 디딤돌, 레시피 대회, 시스템, 교직원, 홍성이며, 코로나19의 연관 키워드는 김00 공동위원장, 양주시, 괴산군 청천면, 진주시, 임00 농림축산국장, 감물면, 당진시 김00 농업정책과장이다.

분석을 통해 나타난 결과를 요약하면 전통 장류와 관련된 기사는 장류 산업이나 문화적 가치보다 상대적으로 지역에 한정된 지원 및 추진사업에 주목하고 있음을 알 수 있다. 물론 언론 기사 특성상 사회적 이슈에 집중하는 경향이 있지만, 전통 장류의 문화적 가치와 역사성, 발전 방안 등을 지속적으로 대중과 공유하려는 노력이 필요하다고 판단된다.

본 연구는 지금까지 연구되어왔던 언론 기사의 직접 분석 방법에서 벗어나 텍스트마이닝 기법을 활용하여 다량의 데이터를 분석하였다는데 의의가 있다. 기존의 일반적인 내용분석 방법과 달리 텍스트마이닝과 연관관계 분석기법을 사용하여 5년 8개월간 보도된 전통 장류 관련 기사들에서 언어의 통계량을 계산하고 이를 연관관계 분석기법과 연계함으로써 기사의 내용을 키워드 형태로 파악하고 특징적인 메시지를 도출한 실험적인 연구이다. 본 연구 결과에서는 전통 장류 관련 주요 이슈들을 파악할 수 있고 언론사별로 어떤 기사를 주로 다루는지 확인할 수 있다. 향후 지속적인 연구를 통해 전통 식품에 대한 키워드를 파악한다면, 대중의 인식과 국가 또는 지자체 주도의 전통 식품 발전 시책 간의 간극을 줄이는데 도움이 될 것으로 보인다. 정부는 한정된 예산과 인력하에서 계승, 개발, 발전이라는 전통 식품 보호 정책은 때론 구체적인 시행에서 어려움이 있을 수도 있다. 키워드 분석에서와 같이 전통 장류는 특정 지자체에 초점을 맞춘

정책이 대부분이며, 중앙정부의 직접적인 지원이나 개발 관련 기사는 드물다. 향후 텍스트마이닝 분석을 통한 정량적 접근법은 전통 식품 발전 방향을 수립하는데 새로운 시각을 제공할 수 있으며, 더 나아가 SNS 데이터를 통한 소비자 인식 분석까지 확대한다면 전통 장류 산업 성장에도 도움이 될 것으로 사료 된다.

References

[1] Korean Law Information Center, "Agricultural Products Processing Industry Promotion Act", 2005. <https://www.law.go.kr/LSW/lsinfoP>

[2] Korea Rural Economic Institute, "Characteristics of Processed Food Consumption Expenditure by Household in 2021", 2021. <https://www.krei.re.kr/foodInfo>

[3] Ministry of Food and Drug Safety, "Food and other production performance in 2021", 2021. https://www.mfds.go.kr/brd/m_374/list.do

[4] M. H. Kim, Character Anlysis of the fermented soybean in Korea, China, and Japan And the current status for export of soybean sauce on Korea, SoMaster's thesis at Sookmyung Women's University Graduate School of Traditional Arts, 2002. http://www.riss.kr/search/detail/DetailView.do?p_mat_type

[5] J. W. Lee, Management of Korean Traditional Soybean Sauces by Korean Families and Soybean Food Industry, Master's thesis, Hankyung University Graduate School of Industry, Korean, 2004. <https://scienceon.kisti.re.kr/srch/selectPORSrchArticle>

[6] B. K. Ku, A research of the regional distinctiveness of Sunchang's fermented foods, Master's thesis, Jeonju University Graduate School of General Studies, Korean, 2018. <https://scienceon.kisti.re.kr>

[7] N. G. Kim, D. H. Lee, H. H. Choi, X. S. Wong, Investigations on Techniques and Applications of Text Analytics. The Journal of Korea Institute of Communications and Information Sciences, 42(2), 471-492, 2017. <https://scienceon.kisti.re.kr/>

[8] Metzler, D., Bernstein, Y., Crofit, W. B., Moffat, A. and Zobel, J. "Similarity measures for tracking information flow". In proceedings of the 14th ACM International Conference on Information and Knowledge Management, 517-524, 2005. DOI: <https://dl.acm.org/doi/10.1145/1099554.1099695>

[9] S. J. Lee, S. H. Lee, H. J. Seol, Y. T. Park, Using patent information for designing new product and technology: keyword based technology roadmapping. R&D Management, 38(2), 169-188, 2008. <https://scienceon.kisti.re.kr/srch/>

[10] M. A. Kam, M. Song, A Study on Differences of

Contents and Tones of Arguments among Newspapers Using Text Mining Analysis. Journal of Intelligence and Information Systems, 18(3), 53-77, 2012.

<https://scienceon.kisti.re.kr/>

[11] J. S. Park, C. S. Kim, G. Y. Kwahk, Investigation of Reaearch Trend in Hotel Domain Using Text Mining and Social Network Analysis. Journal of Tourism and Leisure Research, 28(9), 209-226, 2016. <https://www.kci.go.kr/kciportal/ci>

[12] Korea Health Industry Development Institute, "National Nutrition Statistics in 2020", 2020. <https://www.khidi.or.kr/nutristat>

[13] H. S. Bang, K. H. S. Moon, A Study on the Methodology to express the main topics of text in time series using text mining. Journal of Korean Data&Information Science Society, 30(6), 1259-1276, 2019. <https://scienceon.kisti.re.kr/srch>

[14] I. A. Lee, H. J. Kim, Analyzing the study trends of 'Sense of Place' using text mining techniques. Journal of the Korean Biblia Society for Library and Information Science, 30(2), 189-209, 2019. <https://scienceon.kisti.re.kr/>

유 창 환(Chang-Hwan Yu)

[정회원]



- 2011년 2월 : 충북대학교 농업경제학과 (경제학석사)
- 2015년 6월 : 중국 서북농림과학기술대학교 농업경제관리학과(농업경제관리학박사)
- 2018년 10월 ~ 현재 : 세계김치연구회 산업지원연구단 선임연구원

<관심 분야>

농업경제, 산림·자원경제, 환경경제

이 병 서(Byong-Seo Lee)

[정회원]



- 1986년 2월 : 고려대학교 농업경제학과 (경제학석사)
- 1994년 8월 : 고려대학교 농업경제학과 (경제학박사)
- 2022년 1월 ~ 현재 : (사) 농진중앙회 한국농업개발원 연구위원

<관심 분야>

농업경제, 농업경영, 농산물유통 및 가격