

# 빅데이터 기반 시민의견 모니터링 방안 연구 : “경기지역화폐”를 중심으로

안순재\*, 이새미, 유승의  
동아대학교 스마트거버넌스 연구센터 전임연구원

## A Study on Monitoring Method of Citizen Opinion based on Big Data : Focused on Gyeonggi Local Currency (Gyeonggi Money)

Soon-Jae Ahn\*, Sae-Mi Lee, Seung-Ei Ryu  
Full-time researcher, Smart Governance Research Center, Dong-A University

요약 본 연구에서는 비정형적인 대용량의 텍스트 자료로부터 유의미한 정보를 추출하는 빅데이터 분석방법 중 텍스트 마이닝을 이용하여 시행 중인 정책과 제도에 대한 시민의견을 모니터링 할 수 있는지 확인하였다. ‘경기지역화폐’와 관련된 5,108건의 신문기사와 748건의 온라인 카페글을 수집하여 빈도분석, TF-IDF분석, 연관분석, 워드트리 시각화 분석을 수행하였다. 그 결과로 기사에서는 지역화폐의 도입 목적, 제공되는 혜택, 사용방법에 관련된 내용이 많았고 카페 글에서는 지역화폐의 실사용과 관련된 내용 위주로 작성이 되어있음을 확인하였다. 또한 지역화폐 활성화를 위해서 뉴스는 정보전달자로서 지역화폐의 홍보에 관여하고 있었고 카페글은 지역화폐 사용자인 시민들의 의견으로 이루어져 사용과 관련된 실제적인 정보 교환의 장으로 기능하고 있었다. 지역화폐뿐만 아니라 다양한 정책과 제도에 관해서도 SNS와 텍스트 마이닝을 통해 시민들의 의견을 수렴하여 효과적으로 활성화시킬 수 있을 것으로 보인다.

주제어 : 빅데이터, 텍스트 마이닝, 정책평가, 빈도분석, 연관어분석

Abstract Text mining is one of the big data analysis methods that extracts meaningful information from atypical large-scale text data. In this study, text mining was used to monitor citizens' opinions on the policies and systems being implemented. We collected 5,108 newspaper articles and 748 online cafe posts related to 'Gyeonggi Local Currency' and performed frequency analysis, TF-IDF analysis, association analysis, and word tree visualization analysis. As a result, many articles related to the purpose of introducing local currency, the benefits provided, and the method of use. However, the contents related to the actual use of local currency were written in the online cafe posts. In order to revitalize local currency, the news was involved in the promotion of local currency as an informant. Online cafe posts consisted of the opinions of citizens who are local currency users. SNS and text mining are expected to effectively activate various policies as well as local currency.

Key Words : Big Data, Text Mining, Policy Evaluation, Frequency Keyword Analysis, Association Keyword Analysis

\*This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2018S1A3A2075240)

\*Corresponding Author : Soon-Jae Ahn(asj8652@gmail.com)

Received June 2, 2020

Revised June 18, 2020

Accepted July 20, 2020

Published July 28, 2020

## 1. 서론

지역경제를 살리기 위해 지역화폐(Local Currency)를 도입했거나 도입을 검토하는 지자체가 늘고 있으며 지자체마다 개성 넘치는 지역화폐를 발행해 오고 있다[1]. 지역화폐란 특정 지역에서만 통용되는 대안화폐다[1]. 도입 초기 지역화폐는 자본주의 시장경제의 대안으로 나온 '대안화폐'로 인식되었다. 일반적인 화폐와 다른 차원의 거래시스템을 지역 경제에 도입함으로써 인플레이션, 디플레이션 등과 같은 화폐 경제의 부작용을 완화하고자 했다[1]. 지역화폐는 지역 내 전통시장이나 소상공인 점포에서 현금처럼 사용되며, 지역 밖으로는 유통되지 않는 것이 가장 큰 특징이다[1]. 초창기에는 지역사회 공동체 문화를 복원하고자 풀뿌리 단체 중심으로 추진되었지만 지역경제 활성화라는 목표와 함께 지자체의 주도로 발행되고 있는 흐름이다[1]. 민간조직 중심으로 추진해 온 품앗이 화폐보다는 지자체가 행정구역 내에서 소비를 촉진해 지역경제를 활성화하는 상품권이나 할인쿠폰의 개념으로 확산되었고, 점차 신용카드나 모바일 앱 등 다양한 형태로 발전하고 있다[1]. 2015년 892억 원에 그쳤던 지역화폐 발행액은 지난해 3천714억 원으로 늘었고, 2020년에는 2조 원까지 늘어날 전망이다. 현재 지역화폐를 발행 중인 지자체는 70여 곳이며, 2020년에는 120여 곳에 달할 것으로 예상된다. 이는 전국 243개 광역·기초자치단체의 절반가량에 해당한다. 가장 활성화된 곳은 종이·카드·모바일 등 다양한 형태의 지역화폐를 발행하는 경기도다. 성남·시흥·안양·가평 등 31개 시·군에서 2019년에만 4천962억 원 규모를 발행하며, 2019년 상반기부터는 아동수당·청년 배당·산후조리비 등의 각종 복지수당도 지역화폐로 지급하기로 해 주목받고 있다.

지역화폐를 활성화하고 지속유지하기 위해서는 지자체, 지역주민, 지역상권의 대응 방안이 있다[2]. 지자체에서는 첫째, 지역화폐 오프라인 가맹점에 대한 소비자 접근성을 향상해야 한다. 둘째로 정책 수단으로서 지역화폐를 활용할 필요가 있다. 마지막으로 지역화폐에 대한 적극적인 홍보가 필요하다. 지역주민들은 지역경제 발전에 동참한다는 마음으로 지역화폐 사용에 적극적으로 참여하여야 한다. 지역 상권에서는 지역화폐 사용자들의 재방문을 유도할 수 있도록 서비스 및 제품 개선을 위해 지속적으로 노력할 필요가 있다. 이 중에서 특히 지자체는 지역화폐를 발급하고 운영하는 주체로서 지속적인 모니터링을 통해 문제점들을 발견하고 개선사항을 반영하려는 노력이 필요하다.

오늘날은 온라인과 다양한 매체의 발달로 일반 시민, 이해당사자, 이익집단 등 각계각층의 참여자들이 특정 정책에 관한 자신의 의견 및 판단을 공유하고 있다. 기존에는 공공 데이터와 설문 등의 방법을 통해서 시민들의 의견을 모니터링하였지만 이러한 방법만으로는 사용자들이 생각하는 장점이나 문제점을 파악하는 데 한계가 있다. 반면, Social Network Service(SNS)와 같은 온라인 매체는 정형화된 공공 데이터 및 설문에서 다루지 않는 다양한 사회적 이슈들이 신속하게 전파된다. 또한 해당 문제에 관한 일반 시민들의 진솔한 의견 정보를 포함하고 있다[3-5]. 따라서 지금은 여러 의견 및 판단들을 자동화된 평가방법의 결과로 확인할 수 있는 기술이 필요한 시점이다. 또한 자동화된 방식을 이용하여 정부 정책의 효과성을 평가하고 이해당사자들의 관심도와 관련된 쟁점들을 파악하는 연구는 반드시 함께 수반되어야 한다. 텍스트 마이닝은 설문조사, 여론조사 등 시민의 의견정보를 수집하는 데 필요한 시간과 경제적 비용을 획기적으로 줄일 수 있는 방법으로 정책이나 제도의 평가 과정에서 사용되고 있다[6,7]. 홍순구 등(2019)은 부산 감천문화마을의 방문객들의 감성분석을 목적으로 긍정과 부정 평가를 실시하였고 그 원인을 살펴보았다[6]. 텍스트 마이닝을 통해서 해당 관광지에 대한 시민들의 의견을 수렴하고 문제점 및 개선사항을 도출해냄으로써 중간평가가 가능함을 확인한 것이다.

본 연구에서는 현재 시행 중인 지역화폐 중 가장 사용자가 많은 경기지역화폐에 관한 온라인 뉴스 및 카페글을 수집·분석하고, 지역화폐 제도가 활성화되고 지속·유지되기 위한 대응방안들과 비교하여 중간평가 및 개선사항 도출이 가능한지 확인하고자 하였다.

## 2. 관련연구

### 2.1 텍스트 마이닝

텍스트 마이닝(Text mining)은 대규모의 텍스트 자료로부터 유의미한 정보를 추출하는 데이터 마이닝(data mining)의 한 종류이다. 일정 양식과 형태가 갖추어져 있는 정형 데이터를 다루는 것이 아니라 텍스트, 사진, 동영상 등의 비정형 데이터를 다룬다는 점에서 그 중요성이 높아지고 있다. 빅데이터의 데이터 대부분은 비정형 데이터이기 때문이다. 텍스트 마이닝은 비정형 데이터 중에서도 텍스트 데이터를 바탕으로 유의미한 정보를 찾아낸다[8]. 본 논문에서는 뉴스와 카페글을 수집하여 실제

시민들이 지역화폐를 어떻게 생각하는지 확인하기 위해 텍스트 마이닝을 진행하였다. 카페글의 경우 네이버 카페의 수집 기간 내 ‘경기지역화폐’ 관련 검색가능한 모든 글을 텍스트 데이터로 용했다. 언론의 경우 역시 수집 기간 내 모든 ‘경기지역화폐’ 관련 모든 기사의 제목을 텍스트 데이터로 활용했다.

## 2.2 빈도분석(Keyword Frequency Analysis)

빈도분석은 문서 내에서 사용된 단어의 출현 정도로 문서의 기본적인 특징을 알아내는 분석 기법이다. 빈도가 높은 단어 중 핵심어를 추출하기 위해 TF-IDF(Term Frequency-Inverse Document Frequency)를 이용한다. TF-IDF는 텍스트 마이닝에서 이용하는 가중치 모델로서 문서 내부의 단어들 간 상대적 중요도를 평가하기 위한 표현방식으로 고안된 것이다[9,10]. TF-IDF의 값이 큰 단어는 그 단어가 속한 문서의 주제 또는 의미를 판별할 수 있는 가능성이 크다. 따라서 주요 키워드를 추출할 수 있는 척도로 활용할 수 있다. TF-IDF 가중치는 TF(Term Frequency) 값을 IDF(Inverse Document Frequency) 값을 곱한 것으로 아래 다음과 같이 나타난다.

$$TF-IDF = TF * IDF \dots\dots\dots(1)$$

$$IDF = \log (N/DF) + 1 \dots\dots\dots(2)$$

TF 값은 특정 단어가 한 문서 내에서 출현한 빈도수를 의미하고 IDF 값은 문서의 수(N)를 특정 단어가 나타난 문서의 수(DF)로 나눈 것이다.

## 2.3 연관분석(Association Keyword Analysis)

연관분석은 연관어를 추출하여 키워드가 어떤 주제와 연결되어 있는지에 대해 분석하는 방법이다. 연관어란, 서로 상호관계가 있는 단어들을 말한다. 연관분석은 키워드의 집합이나 용어에서 서로의 연관성이나 상호관계를 발견한다[11]. 연관분석은 시각화 방식에 따라 다양한 결과를 알아낼 수 있는데, 본 논문에서는 Heatmap을 사용하여 결과를 시각화하였다. Heatmap은 2차원 배열 데이터의 값을 색의 짙기로 표현하기 때문에, 연관성을 한 눈에 파악하기에 편하다[12,13]. 본 논문에서는 언론과 카페글에서 많이 등장하는 단어를 추출해내고 단어들 간 연관성을 살펴보기 위해 연관분석을 이용했으며, Heatmap을 통해 결과를 시각화했다.

## 2.4 워드트리 시각화

단순히 빈도가 높고 연관성이 높은 단어들을 찾더라도 단어의 빈도만을 가지고 그 단어가 문장 안에서 어떤 의미로 사용되었는지 판단하기는 어렵다. 따라서 빈도가 높은 단어들의 전후로 어떤 단어가 연속적으로 출현했는지 확인하고 어떤 의미로 사용되었는지 판단하기 위하여 워드트리 방법으로 시각화하였다.

# 3. 데이터 수집 및 분석

## 3.1 데이터 수집

SNS의 반응 연구를 위해 카페글을 활용하였다. 다양한 SNS 채널이 존재하지만 특정 지역에서만 사용되는 지역화폐에 관해서는 특정 지역의 인터넷 카페에서 정보가 많이 공유되고 있다. 언론의 반응을 분석하기 위해서 포털사이트 중 ‘네이버’를 이용했다. 거의 모든 언론사의 기사가 모인다는 점에서 네이버에 올라온 기사를 데이터로 사용하였다[14]. 2019년 1월 1일부터 2020년 2월 28일까지 ‘경기지역화폐’가 언급된 모든 신문 기사와 카페글을 크롤링하였다[15]. 최종적으로 총 748개의 신문 기사와 5,108개의 카페글을 수집하였다. 수집된 카페글 중 신문 기사와 지자체의 홍보글을 그대로 옮긴 글들은 삭제하여 총 971개의 카페글을 분석에 사용하였다.

## 3.2 빈도분석

각 매체에서 많이 쓰인 단어들과 중요단어를 추출하기 위하여 빈도분석과 TF-IDF 분석을 진행하였다. 크롤링한 데이터를 형태소 단위로 쪼개어 각 단어의 빈도수를 추출하였다. Table 1는 각 매체에서 많이 쓰인 키워드의 빈도수와 TF-IDF 값을 정리한 표이다. 그 결과 언론의 경우에는 ‘경제, 정책, 기본소득’ 등 지자체의 지역화폐 도입 목적과 관련된 단어들의 빈도와 TF-IDF 값이 높았고 카페글의 경우에는 ‘카드신청, 인센티브, 가맹점’등과 같이 지역화폐의 실사용과 관련된 단어들의 빈도와 TF-IDF 값이 높았다.

Table 1. Keyword Frequency and TF-IDF value

News	Freq	TF-IDF	Online cafe	Freq	TF-IDF
Economy	1102	533.177	Card application	572	391.701
Policy	1056	756.405	Incentive	313	587.779
Card	1036	773.959	Benefits	249	443.144

Small Business	816	517.625	Franchisee	228	461.203
Promotion	711	749.3	Postpartum care cost	177	371.531
Incentive	562	729.189	Sale	173	386.972
Traditional market	545	604.643	Income deduction	159	346.869
Alley business district	508	493.949	Where to use	136	293.799
Franchisee	463	596.198	Available	122	263.555
Youth	445	715.003	Policy	106	286.924
Basic income	436	828.908	Regional economy	94	241.677
Sale	431	638.223	Parturition	92	228.165
Regional economy	241	354.036	Supermarket	78	189.049
Welfare	231	390.555	Postpartum care	65	211.148
Postpartum care cost	225	369.198	Small business	62	174.985
Self-employed	220	414.327	Youth dividends	57	161.369
Basic Youth Income	206	365.529	Department store	54	151.316
Sales	185	284.071	Traditional market	49	156.202
Consumption	183	321.049	Hospital	48	151.623
Municipality	173	369.071	Postpartum care center	41	144.98

### 3.3 연관분석

각 매체의 상위 빈도 20개 단어 간의 연관 관계를 살펴보기 위해 heatmap으로 시각화했다. heatmap은 각 단어들이 얼마나 자주 함께 쓰이는지를 한눈에 살펴보기에 적합한 시각화 함수이다[12,13]. heatmap 안의 값은 각 단어 간의 자카드 유사도 값이다. 자카드 유사도는 두 집합 사이의 겹치는 정도를 나타낸다. 두 데이터의 교집합의 크기를 합집합의 크기로 나눈 계수이다. 두 집합이 동일한 값이면 1이 나오고, 전혀 다르면 0의 값이 나타난다. 두 항목이 겹치는 부분의 절대량만을 고려하는 것이 아닌, 공통부분이 얼마나 많은지를 고려하여 상대적인 값을 유사도로 나타낸다. 따라서 자카드 유사도를 통해서 단어들이 얼마나 동시에 출현하는지 확인할 수 있다. 분석결과는 아래 Fig1, Fig2와 같다.

뉴스의 경우에는 빈도가 높은 단어들은 골고루 함께 사용된 것을 볼 수 있다. 카페글의 경우에는 실제로 어디서 사용할 수 있는지 사용처에 관련된 단어들이 주로 함께 쓰인 것을 볼 수 있다.

### 3.4 워드트리 시각화

각 매체의 상위 빈도 20개를 골라 단어들의 전후로 어떤 단어들이 연속적으로 출현했는지 확인하기 위하여 워드트리로 시각화하였다. 분석결과는 아래 Fig3, Fig4와 같다. 뉴스의 경우에는 ‘지역-경제’, ‘전통시장-소상공인’,

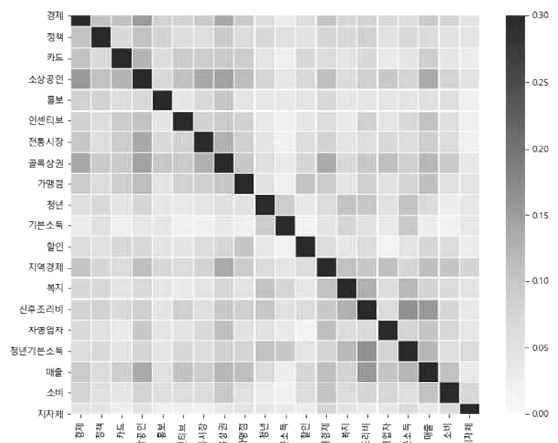


Fig. 1. Association Analysis Heatmap : NEWS

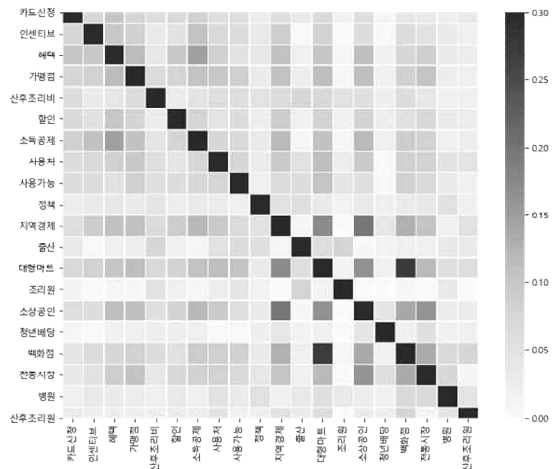


Fig. 2. Association Analysis Heatmap : Cafe text

‘전통시장-골목상권’과 같이 지자체에서 지역화폐를 추진하는 목적과 관련된 단어들이 함께 출현하고 ‘인센티브-지급’, ‘소득공제’, ‘청년배당-산후조리비-지원’ 등과 같이 혜택과 복지수당에 관련된 단어들이 함께 출현함으로써 지역화폐가 어떤 목적으로 출시되었고 어떤 혜택이 제공되는지에 대한 정보전달을 목적으로 작성되었음을 확인할 수 있다. 카페글의 경우에는 ‘소득공제-혜택-제공’, ‘인센티브-제공’, ‘인센티브-지급’, ‘혜택-제공’ 등 사용 혜택과 관련된 단어들이 함께 출현하고 ‘경기지역화폐-사용처’, ‘가맹점-확인’, ‘슈퍼마켓-유흥업소’, ‘대형마트-기업형’, ‘대형마트-백화점’ 등 실제 사용처에 관련된 단어들이 함께 출현하여 혜택과 사용처에 관련된 내용을 궁극해함을 확인할 수 있다.

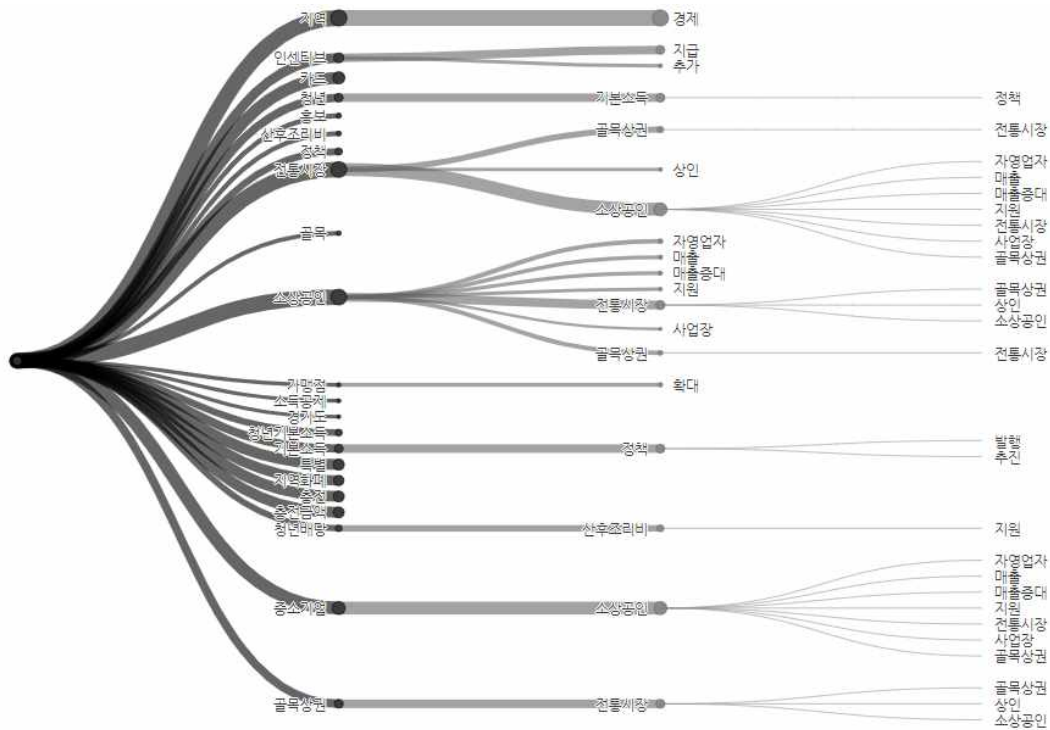


Fig. 3. 1-way Wordtree : News

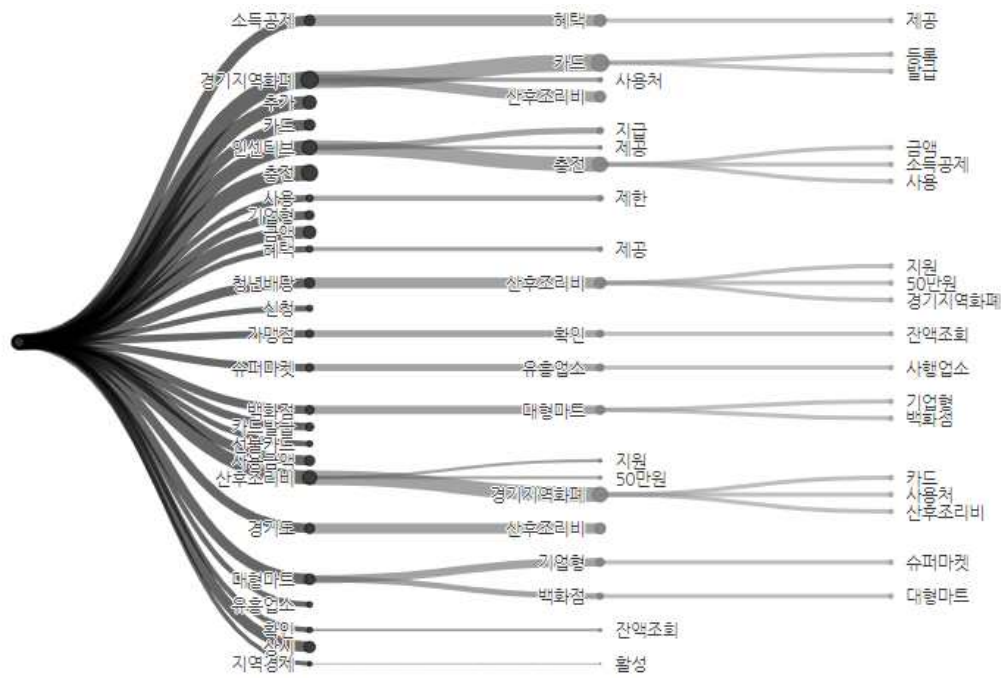


Fig. 4. 1-way Wordtree : Cafe

#### 4. 연구 결과에 대한 논의 및 평가

본 연구에서는 시행 중인 정책과 제도에 대한 중간평가 및 활성화를 위해 언론과 SNS의 데이터를 수집하여

시민의견을 모니터링하는 방안을 제안하고자 하였고 ‘경기지역화폐’를 대상으로 하였다. 연구를 위해 텍스트 마이닝 기법을 사용하였고, 그 중에서도 빈도분석, 연관분석, 워드트리 시각화 방법을 사용하였다. 결과적으로, ‘경

기지역화폐'에 관해서 언론과 SNS에서 주로 다루어지는 주제가 달랐다는 점과 SNS를 통한 시민 의견 모니터링이 가능함을 확인하였다.

빈도분석을 통해 언론과 SNS에서 다루어지는 내용을 확인할 수 있었다. 먼저 언론에서는 지역화폐의 도입 목적, 제공되는 혜택, 사용방법에 관련된 내용이 많았고 SNS에서는 지역화폐의 실사용과 관련된 내용 위주로 작성되어 있었다. 다시 말해, 언론은 지자체에서 시민들에게 알리고자 하는 정보전달의 내용을 주로 다루고 SNS는 실제 시민들이 사용하는 과정에서 궁금한 내용들과 세부적인 혜택에 관련된 질문 답변 글이 많았다.

그리고 연관분석을 통해 빈도수가 높은 단어들의 상관관계를 알아보았다. 언론은 정보전달을 위해 사안과 관련된 단어들만 골고루 사용되고 있지만 SNS의 경우에는 실제 사용처에 관련된 단어들만 주로 사용되는 것으로 보인다. 워드트리시각화를 통해 연관분석과 마찬가지로 시민들이 지역화폐의 혜택과 사용처에 관련된 내용을 궁금해함을 확인할 수 있다.

지역화폐 활성화를 위한 방안으로 지자체는 뉴스를 통하여 적극적인 홍보를 하고 있고 정책 수당으로서 지역화폐를 잘 활용하고 있는 것으로 보인다. 하지만 SNS를 통하여 시민들의 의견을 들여다보면 가맹점에 대한 소비자의 접근성 향상이 제대로 이루어지지 않는 것을 확인할 수 있다. 지역화폐 관련 앱이나 홈페이지에서 제공되는 가맹점 검색기능이 가맹점과 소비처에 대한 시민들의 궁금증을 완전히 해소해 주지 못하는 것이다. 지역주민들은 SNS를 통해 적극적으로 질문을 하고 관심을 보임으로써 지역화폐 사용에 대한 적극적인 참여의지를 보이고 있다. 따라서 지역화폐 제도가 지속적으로 활성화되려면 지자체와 지역상권이 잘 협조하여, 보다 쉬운 가맹점 검색기능을 제공하고 상가의 가맹점포시 스티커 부착, 지역화폐 사용가능처 광고 등을 통하여 소비자의 접근성을 향상시킬 필요가 있다.

지역화폐 사용량 및 발행액의 증가와 지역경제에 끼치는 영향 등을 고려하여 지역화폐 제도가 성공적으로 진행되고 있다고 평가할 수 있다. 하지만 시민들이 지속적으로 사용하지 않는다면 긍정적으로 작용했던 경제효과가 다시 처음으로 돌아가는 문제가 발생할 수도 있다. 계속적인 지역화폐 활성화를 위해서는 꾸준히 시민들의 의견을 수렴하여 제도 개선에 반영해야 하지만 설문조사와 같은 기존의 방법들은 많은 시간과 비용이 소비된다. 시민들의 의견이 다양하게 표출되는 SNS를 통해 의견을 수집하고 텍스트 마이닝 방법을 통해 의미 있는 정보를

추출한다면 많은 시간과 비용을 절약할 수 있을 것이다.

## 5. 결론 및 향후연구

본 연구에서는 언론과 SNS의 데이터를 수집하여 경기도에서 추진 중인 '경기지역화폐'에 대한 시민 의견 모니터링이 가능함을 확인하고자 하였다. 정책평가는 주로 전문가 또는 전문가관에서 평가지표를 바탕으로 근거하여 이루어졌으나 본 연구에서는 텍스트 마이닝을 통해 수요자들의 직접적인 의견을 수렴하여 보다 자동적으로 효율적인 중간평가를 하였다는 차별성이 있다.

연구의 공헌도는 다음과 같다. 학술적으로는 본 연구에서 텍스트 마이닝을 적용한 정책과 제도에 대한 중간평가를 제시하였다는 점에서 향후 후속 연구를 유발할 것으로 예상된다. 또한 실무적으로는 정부 정책의 새로운 평가 방법으로서 국민들이 SNS 상에 남긴 텍스트를 분석하여 정부 사업이 올바른 방향으로 진행되고 있는지를 판별할 수 있는 기초자료로 활용될 수 있을 것이다. 기술의 발달에 따라 갈수록 다양해지고 있는 언어사용 환경에서 사용자들의 실제적인 의견을 수집하여 효과적이고 적절한 방법을 통해 분석한다면 앞으로 시행될 정책과 제도의 안착에 성공적인 효과를 거둘 수 있다고 생각된다.

## REFERENCES

- [1] M. M. Yoo. (2019). A Study On Types, Characteristics and Establishment of Regional Currency Movement. *The Journal of Labor Studies*, 39(2019.12), 131-157.
- [2] J. Choi, D. Jeon & S. Yoon. (2016). *A Study on the Utilization plan of the Community Currency in Gyeonggi-Do*. Suwon : Gyeonggi Research Institute.
- [3] H. J. Woo & Y. H. Kim. (2017). Spatial Distribution Patterns of Twitter Data with Topic Modeling. *Journal of the Korean Association of Regional Geographers*, 23(2), 376- 387.  
DOI : 10.26863/jkarg.2017.05.23.2.376
- [4] S. Son, D. Kim, S. Lee, M. Gil, & Y. Moon. (2017). Storm-Based Dynamic Tag Cloud for Real-Time SNS Data. *KIPS Transactions on Software and Data Engineering*, 6(6), 309-314.  
DOI : 10.3745/PKIPS.y2016m10a.47
- [5] Y. Kim & H. Kang. (2016). An Analysis of Relationship Between Word Frequency in Social Network Service Data and Crime Occurrences. *KIPS Transactions on Computer and Communication Systems*, 5(9). 229-236.

DOI : 10.3745/KTCCS.2016.5.9.229

- [6] S. Hong, S. Ryu & S. Ahn. (2019). Sentimental analysis on urban regeneration policy: Focused on reviews of Gamcheon culture village. *Journal of the Korean Data & Information Science Society*, 30(6), 1233-1244.  
DOI : 10.7465/jkdi.2019.30.6.1233
- [7] S. Ahn, S. Ryu & S. Hong (2020). A Sentiment Analysis Model for Small-Scale Unstructured Policy Data Using Transfer Learning. *Journal of the Korean Data & Information Science Society*, 31(2), 405-414.  
DOI : 10.7465/jkdi.2020.31.2.405
- [8] I. H. Kim. (2019). Newspaper Big Data and Text Mining for Digital Humanities. *The Journal of Language & Literature*, 78, 41-62.  
DOI : 10.15565/jll.2019.06.78.41
- [9] S. Lee & H. Kim. (2009). Keyword Extraction from News Corpus using Modified TF-IDF. *The journal of society for e-Business Studies*, 14(4), 59-73.
- [10] Y. Noh, J. Lim, K. bok & J. Yoo. (2020). Hot Topic Prediction Scheme Using Modified TF-IDF in Social Network Environments. *KIISE Transactions on Computing Practices*, 23(4), 217-225.  
DOI : 10.5626/KTCP.2017.23.4.217
- [11] J. H. Ryu & Y. Y. Yoo. (2018). The Fourth Industrial Revolution Core Technology Association Analysis Using Text Mining. *Journal of Digital Convergence*, 16(9), 129-136.  
DOI : 10.14400/JDC.2018.16.8.129
- [12] S. H. Cha. (2007). Comprehensive survey on distance/similarity measures between probability density functions. *International Journal of Mathematical Models and Methods in Applied Science*, 1(4), 300-307.
- [13] S. Lee & H. Kim. (2009). Comparison of responses to issues in SNS and Traditional Media using Text Mining -Focusing on the Termination of Korea-Japan General Security of Military Information Agreement(GSOMIA)-. *Journal of Digital Convergence*, 18(2), 277-284.  
DOI : 10.14400/JDC.2020.18.2.277
- [14] S. Kim & J. Lee. (2016). Naver's influence on the public opinion, ahead of KBS·Chosun, Journalists Association of Korea,  
<http://www.journalist.or.kr/news/article.html?no=38363>
- [15] M. Kobayashi and K. Takeda. (2000). Information Retrieval on the Web. *ACM Computing Surveys*, 32(2), 144-173  
DOI : 10.1145/358923.358934

안 순 재(Soon-Jae Ahn)

장학원



- 2018년 8월 : 연세대학교 박사(공학)
- 2019년 5월 ~ 현재 : 동아대학교 스마트거버넌스연구센터 전임연구원
- 관심분야 : 빅데이터, 인공지능, 텍스트마이닝
- E-Mail : asj8652@gmail.com

이 새 미(Sae-Mii Lee)

장학원



- 2017년 2월 : 세종대학교 박사(경영학)
- 2010년 1월 : Northeastern University 석사(리더십)
- 2018년 11월 ~ 현재 : 동아대학교 스마트거버넌스연구센터 전임연구원
- 관심분야 : 서비스 마케팅, 리더십, ICT, 텍스트마이닝

· E-Mail : emailme6@naver.com

유 승 의(Seung-Eui Ryu)

장학원



- 2014년 5월 : University of Texas at El Paso 박사(Information Systems)
- 2018년 9월 ~ 현재 : 동아대학교 스마트거버넌스연구센터 전임연구원
- 관심분야 : 국제경영학, 빅데이터, 인공지능, 텍스트마이닝
- E-Mail : juim0928@dau.ac.kr