

Print ISSN: 2233-4165 / Online ISSN: 2233-5382
doi:http://dx.doi.org/10.13106/ijdb.2018.vol9.no8.51.

Irregular Bigdata Analysis and Considerations for Civil Complaint Based on Design Thinking*

비정형 빅데이터 분석 및 디자인씽킹을 활용한 민원문제 해결에 대한 고찰

Tae-Hyung Kim(김태형)** , Byung-Jae Park(박병재)*** , Eung-Kyo Suh(서응교)****

Received: July 3, 2018. Revised: July 4, 2018. Accepted: August 15, 2018.

Abstract

Purpose - Civil affairs are increasing in various forms, but civil servants who are able to handle them want to reduce the complaints and provide keywords that will help in the future due to their lack of time. While various ideas are presented and implemented as policies in solving civil affairs, there are many cases that are not policies that people can sympathize with. Therefore, it is necessary to analyze the complaints accurately and to present correct solutions to the analyzed civil complaint data.

Research design, data, and methodology - We analyzed the complaints data for the last three years and found out how to solve the problems of Yongin City and alleviate the burdens of civil servants. To do this, the Hadoop platform and Design Thinking process were reviewed, and proposed a new process to fuse it. The big data analysis stage focuses on civil complaints - Civil data extraction - Civil data analysis - Categorization of the year by keywords analyzing them and the needs of citizens were identified. In the forecast analysis for deriving insights, - The case of innovation case study - Idea derivation - Idea evaluation - Prototyping - Case analysis stage used.

Results - Through this, a creative idea of providing free transportation cards to solve the major issues of construction, apartment, installation, and vehicle problems was discovered. There is a specific problem of how to provide these services to certain areas, but there is a pressing need for a policy that can contribute as much as it can to the citizens who are suffering from various problems at this moment.

Conclusions - In the past, there were many cases in which free traffic cards were issued mainly to the elderly or disabled. In other countries, foreign residents of other area visit the areas for accommodation, and may give out free transportation cards as well. In this case, the local government will be able to set up a framework to present with a win-win scenario in various ways. It is necessary to reorganize the process in future studies so that the actual solution will be adopted, reduce civil complaints, help establish policies in the future, and be applied in other cities as well.

Keywords: Big Data Analysis, Unstructured Data, Design Thinking, Complaint Issue.

JEL Classifications: C55, D83, H79.

1. 서론

미디어의 발달로 다양한 민원제기 활동이 커지고 있으며, 이 규모도 폭발적으로 증가하고 있어, 일일이 대응하기가 어려운 것이 현실이다. 대응을 하더라도, 중복분야의 중복처리로 인하여, 많은 인적, 물적 자원이 낭비되고 있으며, 시민의 만족도 또한 처리기간이 길어지거나, 잘못된 해결방안으로 불만이 가중되고 있다. 민원의 특성상 텍스트 혹은 이미지 형식의 비정형데이터가 많은데, 이러한 데이터를 정형화하여, 시/도 차원에서 이를 활용하여 향후 정책에 반영하려는 노력도 하지

* This Paper was Modified and Developed from the Master Thesis of the Co-Author.

** First Author, Professor, Graduate School of Data Science, Dankook University, Korea. E-mail: kimtoja@dankook.ac.kr

*** Co-Author, Researcher, Software Design Convergence Center, Dankook University, Korea. E-mail: bjaepark@dankook.ac.kr

**** Corresponding Author, Professor, Graduate School of Business, Dankook University, Korea. Tel:+82-31-8005-3981, E-mail: eungkyosuh@dankook.ac.kr

않고 있는 것이 현실이다.

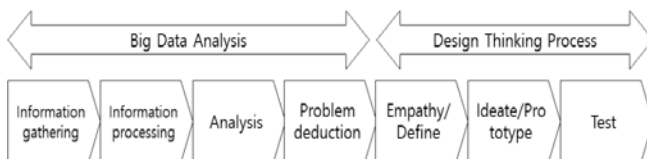
한국행정연구원 조사 자료에 의하면 최근 시민들의 의식수준 향상, 환경에 대한 관심 증가와 더불어 빠르게 진행되고 도시발전으로 인해 환경에 대한 민원이 급속도로 증가하고 있다. 그리고 통계청 자료에 의하면, 2015년 강원도의 민원서비스의 불만족 주요사유가 불친절로 나타났다. 이는 시민의 불만을 정확히 파악하지 못하여, 민원을 해결하지 못하는 공무원의 무의식과 친절도가 낮다고 표현된 것으로 예상된다. 또한 과거에 요청했으나, 처리되지 못한 것이 누적된다면, 시민의 불만은 가중될 것이다.

시청, 구청에서의 부서별 구술 민원, 전화 응대, 콜센터의 VOC 민원, 각종 온라인 민원 등으로 다양한 채널로 민원이 제기되고 있다. 이에 공무원들이 일일이 대응하기가 어렵기 때문에 한정된 인원으로 많은 민원을 대응할 방안이 필요하다.

본 연구에서는 용인시의 행정업무의 불만이나 혹은 질의, 부당한 제도 등 여러 분야의 민원데이터를 추출하기 위해 용인시 새울 전자민원창구 선택했다. 새울이란, 새롭고 올곧은 행정을 수행하는 시스템이란 의미로 2000년대부터 각종 대민 서비스와 행정 업무를 표준화/전산화된 시스템으로 안전행정부가 개발 및 보급하여 전국 자치단체에서 동일하게 사용하는 행정정보 포털 시스템이다.

검증된 이 시스템에서 최근 3년간 민원데이터를 추출하여, 빅데이터 분석을 통하여 시민의 권리 침해에 대한 시민의 목소리가 어디서 많이 발생하는지 파악하고, 해석하려고 한다. 본 연구에서는 분석한 데이터를 토대로 결론을 짓지 않고, 디자인씽킹(Design Thinking)을 통해 시민이 경험하게 될 해결책을 중심으로 분석할 것이다. 공급자 중심에서 소비자 중심으로 패러다임이 바뀌고 있는 만큼 시민의 요구와 감성을 파악하고, 과거의 흐름을 통해 향후 나아가야 하는 방향을 예측하여, 선제적으로 민원을 대응하기 위한 기틀을 마련해 나가려고 한다.

본 연구에서는 용인시의 잠재적인 문제점을 해결하기 위해 민원데이터를 활용하였다. 실제 시민이 올리는 민원이야말로, 현장의 목소리이자 염원이라 할 수 있다. 특정 분야나 키워드를 해결하기보다는 거시적인 관점에서 민원을 파악하여 향후 시정에 참고자료가 되고자 하는데 의의가 있다. 서론에서는 민원데이터를 분석하게 된 배경과 목적에 대해 설명하였다. 이론적 배경에서는 민원의 정의, 민원을 분석한 논문을 분석하였다. 그리고, 빅데이터 분석 프레임워크인 하둡을 이용해서 비정형 텍스트분석을 한 문헌을 분석하여, 이와 같은 플랫폼을 활용해서 민원을 분석하였다. 분석과 획일적인 제안에 그치지 않기 위해 디자인씽킹 프레임워크의 발전 과정을 살펴보고, 관련 문헌을 고찰하였다. 이에 창의적인 프로세스를 도출하고, 이를 통해 창의적인 아이디어를 해결방안으로 제시하였다. 아래의 <Figure 1>은 본 연구에서 지향하는 프로세스이다. 전통적인 빅데이터 분석 플랫폼에 디자인씽킹 프로세스를 통해 창의적인 해결방안을 제시하였다.



<Figure 1> Design Thinking Process through Big Data Analysis

2. 이론적 배경

2.1. 민원 연구

세상을 살아가다 보면 다양한 경험을 통해 성공과 실패를 맞게 된다. 특정 결과에 따라 이득과 손해, 희비가 갈리는 것이 현실이다. 결과를 받아들이기 어렵다면, 경우에 따라 불합리, 불평등, 불만족 반응을 보이게 된다. 시민이나 소비자는 정책에 대한 불만이나 제품에 대한 실망을 불평 행동으로 규정하고, 이를 해소시키고 의사소통하기 위해 다양한 창구를 마련하게 된다. 대표적인 것이 대한민국 국민으로써, 행정기관에 불평등이나 어떠한 문제를 신청하는 민원을 들 수 있다(Lee & Kim, 2009).

민원에 대한 연구는 다양한 분야에 연구가 수행되고 있다. Lee and Kim(2014)은 민원을 청원, 불만표현의 두가지로 분석하고, 도서관에서 민원을 분석, 요인을 도출하였다. 이에 특정 도서관 이용자를 대상으로 설문조사를 실시하였고, 요인분석, 회귀분석 등을 이용해서 통계분석을 하였다. 분석결과, 도서관 민원은 건의/청원과 불평행위로 구분하여 이해하였고, 각각의 종속요인에 영향을 미치는 독립요인으로 사서전문성, 서비스편의성, 사서와의 교류 등 6개 요인을 도출하였다.

환경분야에서는 다량의 정형 및 비정형 자료 즉, 빅데이터를 이용하여 자료의 연관성을 규명하고, 미래전략을 수립하는데 다양한 분야에서 활용되고 있다. 이에 발맞춰 국토교통부에서는 2015년 공간 빅데이터체계 구축 사업과 국토공간정보의 빅데이터 관리, 분석 및 서비스 플랫폼 기술개발사업을 통해 공간 빅데이터 체계의 데이터 제공을 위한 기반 구축 및 분석모델을 제공하고 공간 빅데이터의 활용 서비스 구축에 노력하고 있다.

Park(2010)은 천안시를 대상으로 하는 도시교통통제의 해결책으로 교통수요의 관리를 통해 교통체계 개선, 대중교통의 이용대상인 대학생이 많으므로, 대학생을 대상으로 할인율 적용하여 대중교통 이용율을 높여야 한다고 분석하였다. 그리고 대중교통 중심체계 개선을 통해 교통문제의 해결이 필요하고 도로 등 교통시설의 공급확대 가로등 망 체계 정비 및 신호운영체계의 개선을 적극적으로 추진해야 한다고 언급했다.

2.2. 빅데이터 처리와 분석

정보화 시대를 맞이하여, 하루에도 수천, 수억 건의 데이터가 다양한 채널로 쏟아지고 있고, 생산되는 데이터의 양은 실로 엄청나다. 바야흐로 빅데이터 시대인 것이다. 세계적인 소셜네트워크 회사인 페이스북을 예로 들면, 하루에 25억 개의 콘텐츠가 공유되고, 27억 건의 좋아요 클릭, 3억 장의 사진 업로드, 500테라 바이트의 새로운 데이터가 저장되어 있다. 어마어마한 양을 처리하고, 적재하려면 기존 RDB로는 한계가 있는 것이 현실이다. 이처럼 빅데이터에 대해 주목하고 있는 이유는 빅데이터의 가치에 있다. 다양한 형태(이미지, 영상, 텍스트 등)로 저장된 데이터를 분석하면 새로운 이론을 발견하고, 미래를 예측할 수 있다는 사회적 시각이 반영된 것이다. 빅데이터의 주된 속성으로 3V를 꼽는다. Volume(양), Variety(다양성), Velocity(속도)를 갖추어야 하는데, 이에 Value(가치)를 포함해서 4V라고 일컫는다. 이렇듯 빅데이터를 개발 및 분석하려는 움직임은 여러해 전부터 많은 노력이 이루어졌다. 2006년 더그 커팅에 의해 개발된 하둡은 Google의 GFS(google file system) 논문이 공개된 후, 이를 기반으로 개발되었다. 대용량 데이터를 처리하기 위해 적합한 구조로써, 분산 처리를 위해 Scale-out한다. 과거에는 서버의 처

리능력의 향상을 위해 서버의 CPU, 메모리 등을 키워 대응했다면, 현재는 개인용 컴퓨터급의 컴퓨터를 늘려서 트래픽을 분석시키는 Scale-out을 주로 사용한다.

하둠을 통한 연구도 다양한 분야에서 이루어지고 있고(Min & Bae, 2015), 민원분석에도 활용되고 있다.

Lee et al.(2014)은 부산지역 교통 관련 기사를 추출하여, 부산, '교통'이라는 키워드를 토대로 사회네트워크 분석을 통해 연관된 키워드를 추출해내는 분석을 하였다. 여기에 비정형 텍스트 데이터의 정형화를 위해 빅데이터 저장, 처리, 분석을 위한 오픈소스 프레임워크 하둠 HDFS와 맵리듀스를 리눅스 환경에서 이용하였다. 2013년도 국제신문과 부산일보의 기사를 토대로 분석하여, 당시 부산지역 화두가 부산 프리미엄 아울렛 개장과 이에 인근 지역 교통체증이 관심 키워드였고, 해운대 해수욕장, 전국 호환 교통카드 출시 등이 화제 되었음을 알 수 있었다. 이에 대용량 텍스트를 정독하지 않고도 빅데이터 분석을 통해 정보를 얻고, 분석할 수 있다고 언급하였다.

Choi(2016)는 환경/위생분야 민원분석을 통해 수요자 맞춤형 정책수립에 이를 사용해야 한다고 주장하고 있다. 이 데이터를 토대로 연도별 주요 키워드를 도출하고, 키워드들 사이의 관계를 계량적 분석을 실시하였다. 정부 3.0 정보공개요청을 통해 전주시 민원의 5개년 데이터를 수집하여 자연언어 처리 기술에 기반하여 명사를 추출하였고, 문장 안에서의 관계를 기반한 매트릭스 데이터로 변환하였다. 이를 통해 소셜네트워크 분석 툴인 넷마이너에 넣어 클러스터링 분석, 중심성지수, 그 룹밀도 등을 산출하여 시멘틱 분석을 시도하였다.

<Table 1> Big Data Analysis Process through Complaints Issues

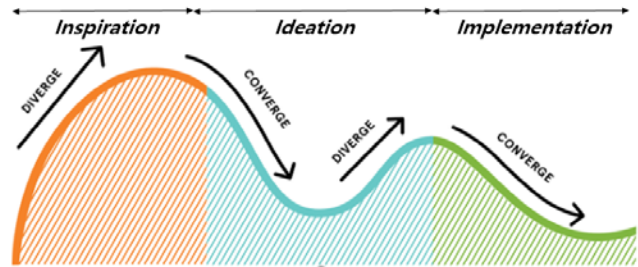
Researcher	Complaints Resolving Process
Choi (2016)	Complaints of Jeonju city → Analysis(Netminer tool) → Frequency analysis → Cluster analysis → Centrality & Density analysis → Environmental Civil Affairs Platform setup
Lee et al. (2003)	News articles about Busan city → Hadoop & Map reduce Big data analysis → Social Network analysis → Deriving keywords issues

2.3. 창의적 프로세스

2.3.1. IDEO

IDEO는 세계에서 가장 혁신적인 기업으로 꼽히는 디자인 기업으로 창의적 사고를 위해 디자인 씽킹을 혁신적인 경영방법으로 널리 소개하였고, 이 기법을 활용하여 수많은 아이디어를 통해 전 세계 기업의 제품을 디자인 하였다. Tim Brown은 그의 저서 'Change by Design'에서 디자인씽킹을 divergence와 convergence를 지속적으로 반복하는 것이라고 규정하고 있고, 이러한 가공되지 않은 방대한 양의 자료로부터, 의미 있는 행동 양식을 발견하고, 추출하는 통합작업에서 창조적인 아이디어를 발견하게 된다.(Kang et al., 2014) 아래 <Figure 2>은 IDEO의 3I 모델로 Inspiration(영감)-Ideation(발상)-Implementation(실현) 3단계로 구성된다. 1단계(Inspiration)에서는 다양한 경험, 관찰에서의 특징을 가감 없이 발언하고 조사한다. 2단계(Ideation)에서는 앞 단계에서의 내용을 토대로 분석하고 특징적인 것을 추출하는 브레인스토밍 단계이다. 3단계(Implementation)에서는 가장 좋은 아이디어를 선정해서 프로

토타입을 만들어보는 것이다. 물론 한번에 좋은 결과물을 낼 순 없으므로, 이 과정을 순환적으로 반복할 수 있다.



<Figure 2> IDEO HCD(Human Centered Design) Toolkit

2.3.2. 스탠포드 대학 d.School

세계적인 기업들이 디자인 씽킹에 주목하고 있는 현실에, 세계 최고의 창조학교로 불리고 있는 스탠포드대학의 d.School도 혁신하고 창조하는 과정을 체험을 통해 가르치고 있다. 스탠포드 대학의 d.School은 세계적인 소프트웨어 회사 SAP를 공동 창업한 Hasso Plattner가 2005년에 이 학교에 기부한 것을 모태로 만들어졌다. d.School에서 만든 디자인씽킹의 5단계는 창의적 혁신을 촉진하는 일련의 과정으로 정의하고, Empathy(공감) - Define(정의) - Ideate(아이디어) - Prototype(시제품) - Test(테스트)를 거친다고 정의하였다(Lee, Suh, & Suh, 2017).

아래의 <Figure 3>은 원거리 통학을 하는 대학원생을 페르 소나로 설정해서 그의 24시간을 들여다보고, 현재 상황을 공감(Empathy)하고, 이 학생이 처한 문제가 무엇인지 파악(Define)한 내용이다. 이 문제를 풀기 위한 방법은 무엇이 있는지 아이디어를 제시(Ideate)하고, 가장 창조적이고 획기적인 안을 가상세계에 적용해보고, 맞지 않는다면 전 단계로 돌아가서 다시 Empathy, Define 단계에서 새로운 방안을 마련해볼 수도 있을 것이다.



<Figure 3> d.School design thinking process example using Post-it

2.3.3. i-school

동경대의 i-school은 창의적인 혁신의 리더를 양성하기 위해 2009년 설립되었다. 인간 중심의 혁신, 크리에이티브 씽킹을 통한 지식체계 수립, 경험을 통한 이슈 해결을 기반으로 한다. 디자인 씽킹을 위해 Understanding-Creating-Realizing라는 3가지 단계를 제시한다. Understanding에서는 Field work를 통해 사물에 대한 직접 관찰 혹은 인터뷰를 통해 문제점에 공감하

고 이해하는 단계이다. 또한 관련 사례 분석을 통해 어떻게 극복했는지 조사한다. **Creating** 단계에서는 전 단계에서 조사한 내용을 기반으로 브레인스토밍이나, **Analogical thinking**(유추) 기법으로 창의적인 아이디어를 제시한다. 마지막으로 **Realizing** 단계에서는 채택한 아이디어를 직접 그림으로 그리거나, 다양한 도구로 만들어 봐서 실현가능성을 확인할 수 있다. i-school은 이러한 과정을 **apisnote**라는 디자인씽킹 도구를 직접 개발하여 웹기반이지만, 직접 포스트잇에 아이디어를 적어서 연결할 수 있는 기능을 제공한다.

1. Understanding	2. Creating	3. Realizing
<ul style="list-style-type: none"> Field observation User interview Fore sighting Case analyzing 	<ul style="list-style-type: none"> Opportunity area defining Analogical thinking Compulsory thinking Brainstorming 	<ul style="list-style-type: none"> Prototype creating Business model creation

<Figure 4> Process of design thinking methodology in i-school

2.2.4. 사례

Chae et al.(2016)은 analytical thinking과 사람의 intuitive thinking을 융합한 디자인씽킹 모델을 활용하여, 숙련과 창조성을 융합함으로써, 공공 피복지원체계 개선아이디어를 제시하였다. IDEO의 교육자들을 위한 디자인씽킹 Toolkit 5단계 (발견하기 - 해석하기 - 아이디어내기 - 실험하기 - 발전시키기)를 적용하였다. 1단계(발견하기)에서는 “사용자 중심의 피복지원체계를 만들어보자”라는 문제를 정의하였고, 4가지 유형의 사용자 그룹을 대상으로 FGI(focus group interview: 초점집단면접)를 시행하였다. 2단계(해석하기)에서는 FGI 결과를 검토하여 “피복을 언제 어디서나 구매할 수 있고, 집에서 배송을 받을 수 있는 시스템”구축으로 정의하였다. 3단계(아이디어내기)에서는 브레인스토밍을 통해 아이디어를 제시하고, 실현가능성을 제시하였다. 4단계(실험하기)에서는 문제의 근본을 찾아 아이디어 정련을 통해 도출된 “피복지원 인터넷/모바일 앱” 구축하기 위한 설계도로 마인드맵으로 작성하였고, 마지막으로 이를 공군에 제안하였다.

Chung(2014)은 다양한 디자인 문제해결과정을 문헌을 통해 조사하고, 자신만의 디자인 프로세스를 4단계로 정의하였다. 1 단계는 자료수집으로 문제공간 내에 존재하는 정보요소가 수집되는 단계이고, 정보요소의 정의와 분류가 이루어진다고 했다. 2단계의 문제 탐색/이해 단계는 사용자의 다양한 맥락을

이해하는 것이 핵심으로 다양한 구조적 관점에서 이해하였다. 3단계 문제정의는 정보요소들 간의 관계 부여를 통해 다양한 관점에서 통합적인 시각으로 바라보았다. 이에 시나리오 모델을 만들어서 데이터 간의 속성에 따라 새로운 형태의 관계모형이 만들어진다고 주장하였다. 4단계인 컨셉 전개에서는 시나리오에 부여된 의미 관계를 통해 문제공간에 숨어있는 핵심 의미를 파악하고, 떠오르는 아이디어를 스케치와 스케일 모형 등의 단순한 프로토타입을 이용해 실험하였다. 이를 여러 차례 반복하여 아이디어를 구체화하여 최종 선택된 디자인 컨셉으로 컨셉 시나리오와 프로토타입을 완성시켰다. 이 방법으로 적용한 사례를 연구함으로써 연구를 마무리 하였다.

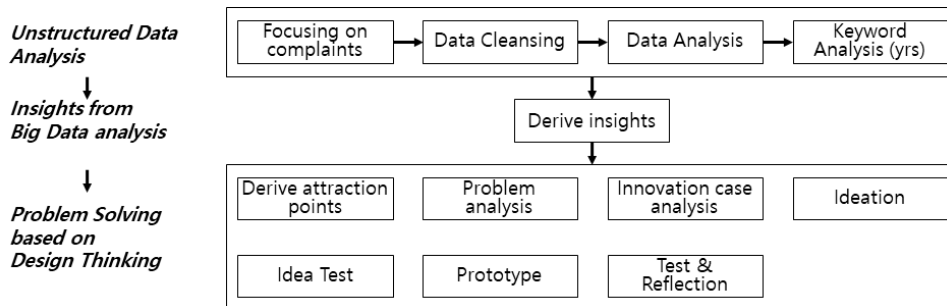
위 2가지 문헌에서 사용한 프로세스를 정의해 보았다.

<Table 2> Creative process for Problem solving

Researcher	Resolving Process
Choi et al. (2016)	Discover Interpretation → Ideation → Implementation → Develop
Jung (2014)	Information gathering Problem understanding → Problem define → Conceptualization

3. 연구방법

본 연구에서는 온라인 민원데이터를 기반으로 데이터 수집-처리-분석-디자인씽킹(Design Thinking)으로 이어지는 창의적 사고를 위한 프로세스를 제안하였다. 시민의 실질적인 목소리를 듣기 위해 민원데이터에 주목하였고, 이를 수집하기 위해 파이썬(Python) 언어를 이용하여 데이터를 수집하였고, 빅데이터 처리 플랫폼인 하둡(Hadoop)에 저장하였다. 그리고 맵리듀스(MapReduce) 프레임워크를 이용하여 문장을 단어별로 추출하여 빈도분석을 실시하였다. 또한 R의 워드클라우드를 활용해 년도 별 주요 키워드를 도식화하여, 그 해의 트렌드를 한눈에 확인할 수 있게 시각화 하였다. 그리고 디자인 씽킹을 통한 예측분석으로 창의적인 해결방안을 마련하였다. 세부 단계로는 용인시의 경우, 용인시의 매력적인 요소를 추출하고, 추출한 주요 키워드와 빅데이터 분석을 통해 도출된 문제점과의 융합을 통해 해결방안 유추할 수 있다. 그리고 지역 혹은 해외의 지역혁신사례 분석한다. 이를 통해 아이디어를 도출하고, 방법과 실현가능성을 타진하고, 평가하여 아이디어를 선정한다. 프로토타이핑을 통해 실제 아이디어의 모형을 만들거나 스케치를 통해 가능성을 타진하고, 문제가 해결되는지 확인하고, 실제 제안을 할 수 있다.



<Figure 5> Research Model

4. 연구결과

4.1. 민원데이터 분석

대한민국 대표 민원창구인 새올 민원창구의 용인시 데이터를 추출한다. 주로 고객의 개선요청, 불만제기 내용이 다수이므로, 용인시가 처해있는 문제점을 파악할 수 있는 좋은 데이터라고 판단되었다. 본 연구에서는 2014년 1월 1일부터 2016년 10월 16일까지 최근 3년간의 7,700여건의 비정형 공공 민원데이터를 파이썬의 urllib.request 모듈을 이용해서 크롤러(crawler)를 제작하였다.

분석할 데이터가 비정형 텍스트 데이터이기 때문에 명사 뿐 아니라, 조사, 대명사, 접속사 등이 섞여 있기 때문에, 정확도를 높이기 위해 명사만 추출하기로 한다. 크롤링한 텍스트 데이터를 R의 KoNLP 패키지에서 extractNoun 함수를 이용해서 조사를 제거하였고, 수동으로, 필요한 명사는 추가하고, 불필요한 감탄사, 대명사, 접속사 등은 제거하였다.

빅데이터 처리 및 분석을 위해 Amazon EC2에 하둡(Hadoop 2.7.3) 클러스터를 구성하였다. NameNode 2대, DataNode 2대를 구성하였으며, NameNode 중 한 대는 SecondaryNameNode로써, NameNode가 오작동할 때, SecondaryNameNode가 작동하기 때문에 고가용성을 확보하였다. 하지만, NameNode가 정상일 때는 DataNode 역할을 하므로 DataNode가 3대인 셈이다. 그래서 추출한 비정형 데이터를 HDFS에 올리고 맵리듀스(MapReduce)를 이용해 빈도 분석을 하였다. 가장 많이 언급된 단어가 관심도나 문제점이 될 만한 키워드라고 판단해서 이 분석을 실시하였다. 아래 순서도는 빅데이터를 처리에서 분석까지의 일련의 과정을 도식화 하였다.

최근 3년간의 7,700여건 비정형 민원데이터를 추출하는데 3 시간정도 소요되었다. 직접 case by case로 webrequest로 호

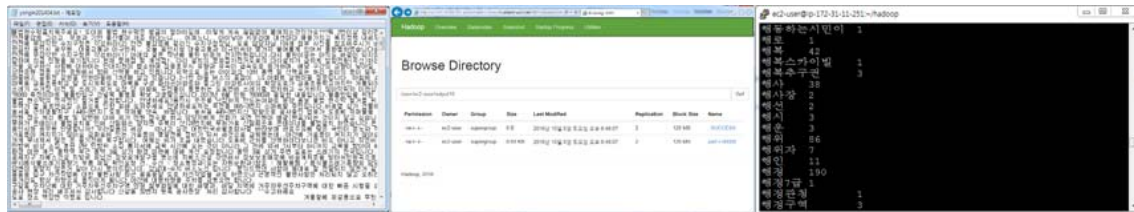
출해서 html를 가져와서 추출하는 구조라서 다소 시간이 소요되었다. 그리고 명사만 추출하여 하둡의 HDFS에 파일을 업로드 하고, 맵리듀스 프로그램을 실행한다. 정상적으로 map-reduce가 완료되었다면, 지정한 output 폴더에 결과 파일이 생성된다. 이에 하둡명령어 “.bin/hdfsdfs-cat/user/ec2-user/yong in2016-11/part-r-00000”를 치면 아래 <Figure 6>와 같은 결과를 확인할 수 있다.

4.2. 연도별 키워드 분석

아래 <Figure 7>은 민원데이터 분석 결과를 R의 WordCloud로 각 연도별로 시각화 하였다. 빈도에 따라 키워드의 크기가 정해지지만, 연관성을 고려해 insight를 도출하였고, 2014년도 민원데이터에서 추출된 키워드가 빈도에 의해 시각화되었다. 당시 광고신도시가 개발되어 입주가 진행 중이었는데, 광고가 용인과 수원을 함께 접하고 있기에, 용인에서도 다양한 이슈가 제기되었음을 알 수 있고, 버스에 대한 이슈가 많았다는 것은 버스에 의존할 수밖에 없다는 교통의 문제임을 알 수 있다.

2015년도 민원데이터를 분석한 결과 공사에 대한 불편, 소음에 시민이 고통 받았을 것으로 판단된다. 지역적으로는 여전히 광고신도시 여파로 수원에 대한 궁금증을 가지고 있어서 용인시의 거주자의 이탈을 예상해 볼 수 있다. 또한 이번에는 용인시 서천지구에 대한 궁금증이나 이슈가 있었음을 알 수 있다.

2016년도 민원데이터를 분석한 결과, 여전히 공사에 대한 시민불편이 가중되었음을 알 수 있고, 2016년도에 신분당선이 개통했지만 여전히 버스에 대한 이슈를 가지고 있음을 알 수 있다. 그리고 지역적으로는 수지구 동천지구와 기흥구 동백지구가 새롭게 이슈화되고 있음을 알 수 있다.



<Figure 6> Results of Hadoop crawling, Mapreduce process



<Figure 7> Word Cloud keyword analysis results of 2014, 2015, 2016 years

분석결과 전체적으로 버스, 공사, 설치, 아파트 순으로 언급이 많이 되고 있음을 알 수 있다. 그리고 버스, 공사의 경우 언급도 많고, 2015, 2016년도 골고루 이슈화 된 것을 알 수 있다. 그리고 설치, 차량의 경우도 2016년도에 급격히 출현 빈도가 높게 나타나고 있다. 또한 빈도는 작지만 도로, 위험, 안전, 불편에 대한 민원이 증가했음을 알 수 있다.

4.3. 빅데이터 인사이트 도출

디자인 씽킹을 통한 예측분석으로 창의적인 해결방안을 마련하였다. i-school에서 자체 개발한 온라인 창의적 아이디어 협업툴인 apinote(<https://www.apinote.com>)를 활용하여, 단계별 요소를 도출하였다.

4.3.1. 용인시 매력적인 요소

생각이나 컨셉을 정의하려고 할 때, 마인드맵이나 브레인스토밍을 많이 활용한다. 생각을 체계적으로 정리하거나 깊이 있는 사고가 필요할 때는 마인드맵, 다양한 시각의 아이디어를

정리할 경우 브레인스토밍을 주로 이용하지만, Field work(체험)을 통해 얻은 느낌을 흰색 메모에 적어놓고, 이에 파생되는 아이디어는 회색 메모에 정의한다. 다른 컨셉으로도 열거하여 연결한다. 그리고 가장 인상 깊었던 메모는 노란색 메모로 변경해 놓는다(<Figure 8>).

4.3.2. 도출된 문제점 분석 및 방안

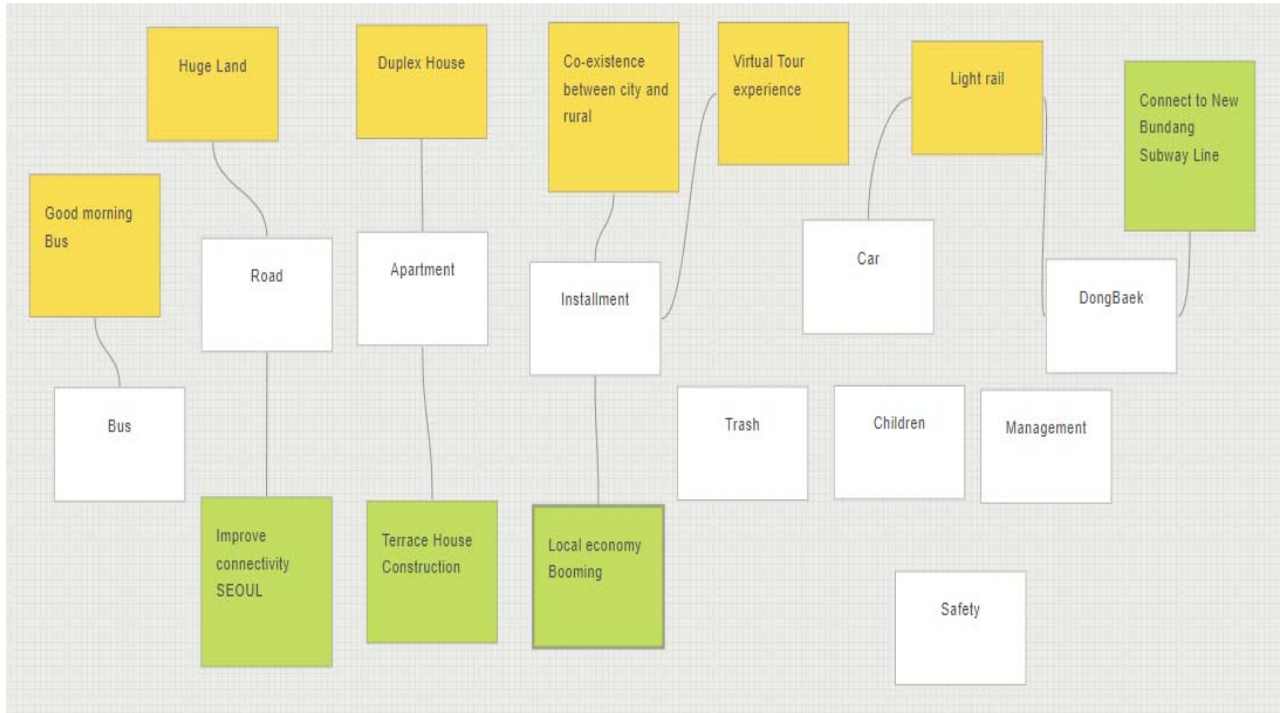
용인시의 매력적인 요소 단계에서 도출된 요소(노란색 메모)와 빅데이터 분석을 통해 도출한 키워드(흰색 메모)를 열거한다. 그리고, 키워드 별로 해결 방안이나 떠오르는 아이디어가 있다면 메모로 정리하고 녹색으로 변경한다. 그리고, 연관성이 있는 메모끼리 연결한다(<Figure 9>).

4.3.3. 혁신사례

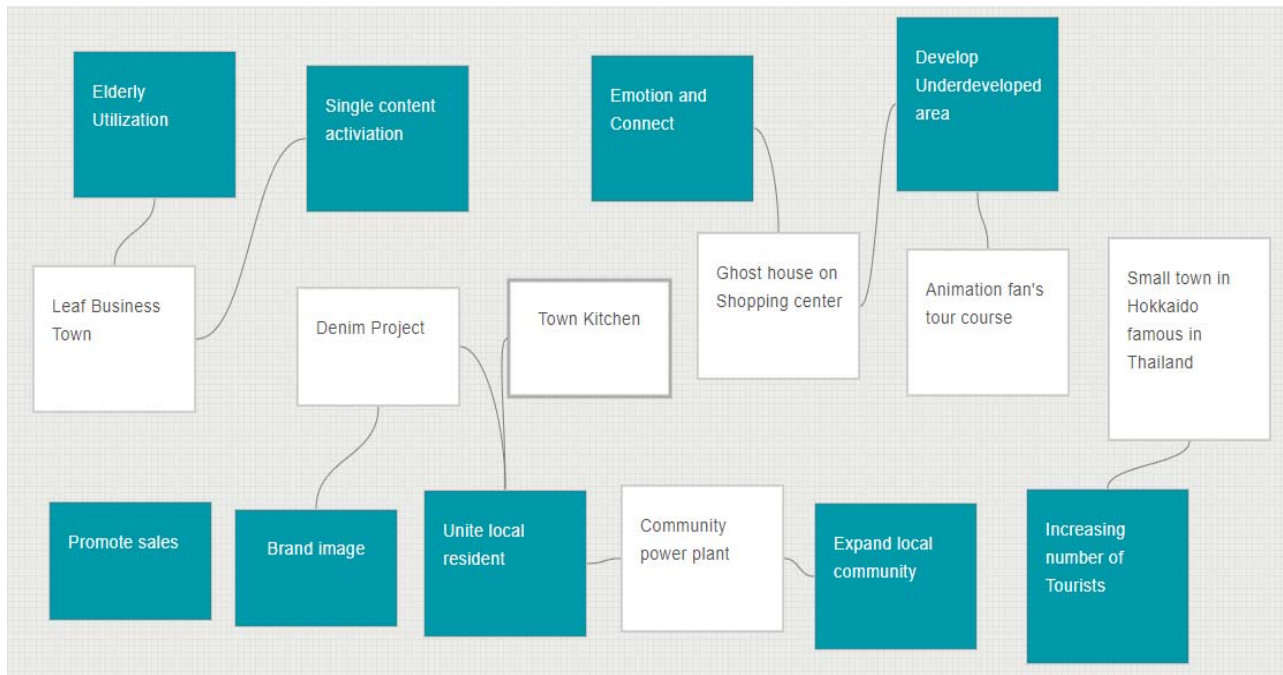
다양한 지역의 혁신사례를 기반으로 활용방안을 마련할 수 있다. 흰색메모는 지역혁신사례이고, 이를 보고 느낀점을 청록색 메모에 코멘트 했다(<Figure 10>).



<Figure 8> Eliciting elements of Yongin city attraction



<Figure 9> Convergence results of attraction point and Complaint issues

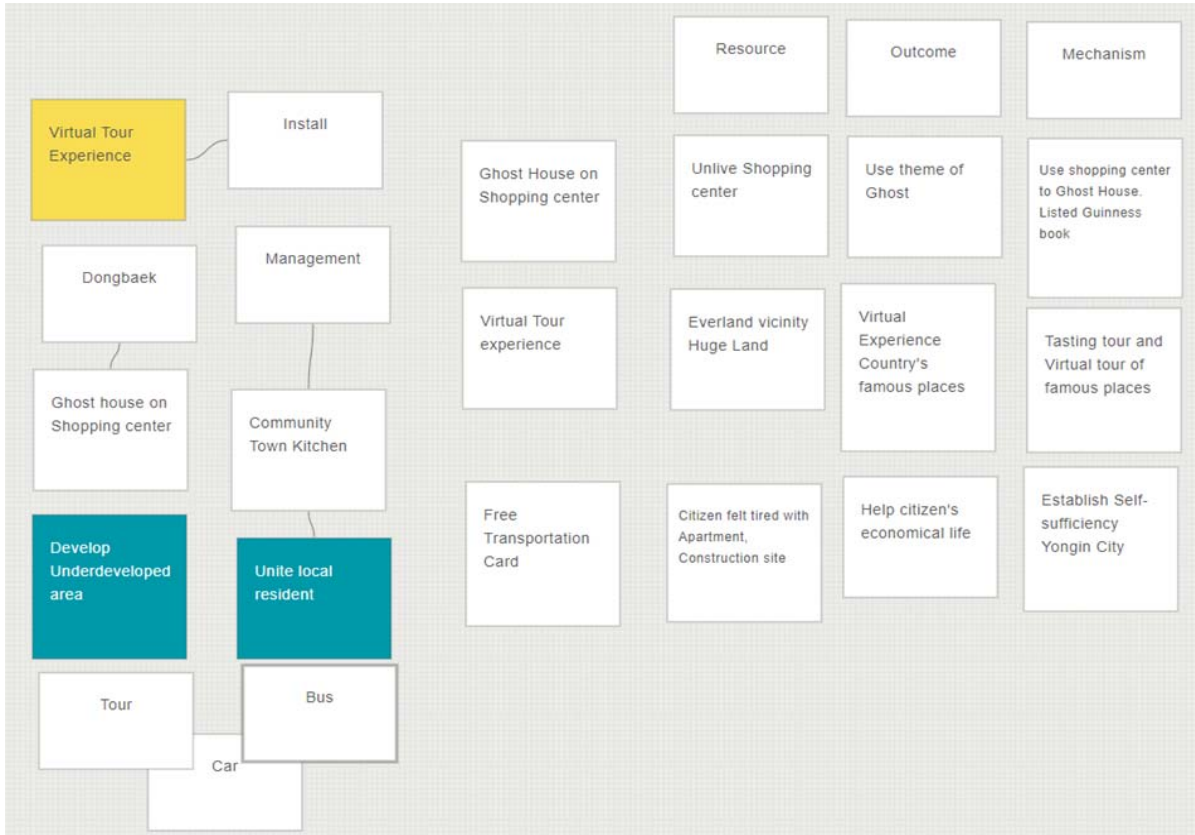


<Figure 10> Regional innovation case analysis

4.3.4. 아이디어 도출 및 방법

용인시의 매력적인 요소, 빅데이터 분석을 통한 키워드 및 해결방안, 지역의 혁신사례를 기반으로 창의적인 아이디어를 적는다. 여기서는 적용 가능한 아이디어를 구체화시키는 단계

로 어떤 자원(resource)으로 어떤 결과(outcome)를 도출할 수 있는 것인지, 또한 어떤 방법(mechanism)으로 할 것인지를 명확히 하는 단계이다(<Figure 11>).



<Figure 11> Idea generation process results

	Novelty	Impact	Feasibility	
Ghost House on Shopping Center	**	****	***	Support of Yongin city
Virtual Tour Experience	***	****	***	Connect to leisure industry
Local Community Hub	*	**	****	Participate diverse people
Free Transportation Card	***	*****	*****	Ad Profit

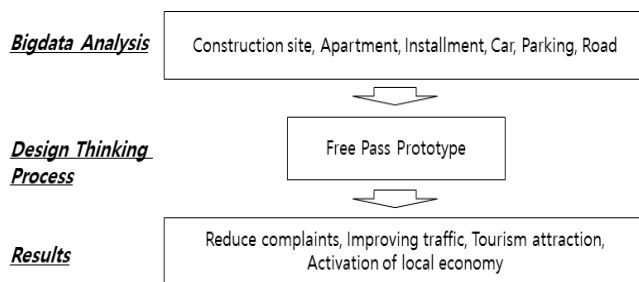
<Figure 12> Best idea selection based on Idea Evaluation

4.3.5. 아이디어 평가

구체화된 아이디어를 바탕으로 novelty(참신성), impact(효과), feasibility(실현가능성)을 평가한다. 여기서 가장 높은 점수를 얻은 아이디어를 채택한다. 여러 아이디어가 필요할 경우 다수 선택도 가능하다(<Figure 12>).

4.3.6. 프로토타이핑

채택한 아이디어를 구체화시키는 단계로, 실제 모형을 만들어 본다. 간단히 스케치해서 컨셉만 잡을 수도 있고, 모형을 만들어 볼 수도 있을 것이다. 그리고 이 아이디어를 채택하였을 경우, 해결 될 수 있는 민원의 키워드와 시민에게 미치는 효과를 정리하였다.



<Figure 13> Overall Process of Research and Results

4.3.7. 사례분석

과거 사례를 비추어 보면, 주로 노약자나 장애인 계층에 무료 교통카드를 발급하는 사례가 많았다. 그리고 해외에서는 지역경제 활성화를 위해 타지의 시민이 본 지역을 방문하여 숙박을 할 경우, 무료 교통카드를 나누어주는 경우도 있다. 이럴 경우, 지역경제 활성화에 이바지할 수 있고, 지자체에서도 광고 등 다양한 방식으로 서로 윈-윈 할 수 있는 기틀을 마련할 수 있을 것이다.

5. 결론 및 제언

4차 산업혁명의 시대에 사람들이 원하는 바는 갈수록 다양해지며 이러한 수준을 만족시키기 위하여 서비스의 수준은 갈수록 높아져가고 있다. 공공부분의 민원서비스도 마찬가지이다. 과거와는 달리 주민들이 요구하는 민원을 전체적으로 파악 및 분석하여 가장 핵심적인 문제를 해결하는 일은 모든 행정자치단체의 필수적 업무가 되어버렸다.

하지만 수많은 민원 데이터를 정독하고, 문제해결하기란 사실상 너무 방대한 양이다. 미디어의 발달로 민원을 제기하는 방법이 용이하고, 이슈가 되어, 많은 이들이 공감하고, 표현되어 중복된 민원이 각기 다른 방식의 비정형화 된 민원으로 접수되기 마련이다. 이러한 현실에 한정된 공무원이 건마다 문제를 파악하고, 해결하는 쉽지가 않다.

이 연구의 의의는 다음과 같다.

첫째, 빅데이터를 분석하여 문제를 정의하고 이를 바탕으로 디자인씽킹 프로세스에 활용하였다는 점이다. 디자인씽킹 프로세스는 연구자들마다 조금씩은 다르지만 공감단계를 통하여

문제의 대상에 대한 새로운 인사이트를 찾아내고 이를 통하여 빠지지 않은 문제를 찾아내어 정의하는데 공감, 정의 단계의 의의가 있다. 대부분의 선행연구들에 따르면 디자인씽킹 프로세스의 공감단계에서는 대상에 대한 관찰, 추적, 간접경험, 인터뷰 등의 방법을 통해 문제의 대상의 심리적 기저를 꿰뚫어 문제점(pain point)을 발견한다. 이 연구에서는 문제의 대상에 관하여 빅데이터 분석을 통해 객관적 데이터 기반의 문제의 대상에 대한 분석을 진행하였고 이를 통해 도출된 내용을 바탕으로 공감단계를 진행하는 기저의 데이터가 되었다.

둘째, 민원문제 해결을 위하여 민원데이터 빅데이터 분석을 통하여, 파악된 민원의 키워드를 디자인씽킹을 통해 해결하는 프로세스로 접근하여 해결방안을 찾아보았다는 점이다. 공공정책의 입안자들에게 있어 객관적인 데이터를 바탕으로 거시적인 관점에서 문제점을 파악하고 이를 보다 세부적인 문제들로 정의해내고, 문제의 대상에 포커스를 두어 적절한 해결방안을 제시함으로써, 정책의 수립 결과를 보다 정확하게 예측하며 입안할 수 있다는 점에서 향후 정책수립에 도움이 되고자 한다.

하지만 본 연구의 한계는 다음과 같다.

첫째, 빅데이터 분석을 통하여 민원데이터를 분석하고 이로부터 도출된 핵심적 민원을 문제로 정의하여 디자인씽킹 프로세스를 적용하고 연구를 통해 검증은 하였으나, 민원문제에 대하여 실제적인 시민들의 모습을 관찰하고, 인터뷰를 통하여 각 계층의 의견을 청취하지 못하여 보다 생생한 시민들의 문제를 분석하지 못하였다.

둘째, 디자인씽킹 프로세스 단계에서 도출된 민원문제의 해결책에 대하여 실제적 검증을 진행하지는 못한 점이다. 즉, 디자인씽킹 프로세스에서 가장 중요한 단계중 하나인 테스트 단계를 일부 진행하였다. 향후 연구에서 이러한 부분이 보완되어 실질적인 검증이 추후에 지속적으로 이루어져야 할 것이다.

아무쪼록 본 연구에서 제시한 빅데이터 분석 기반 디자인씽킹 프로세스를 통해 다양한 지자체에서 활용하여 시정의 참고자료로 사용되고 민원이 줄어드는 효과를 볼 수 있길 기대해 본다.

References

- Choi, H. O. (2016). Study on Selecting Priority Criteria Utilizing Civil Complaint Data in the Field of Environment and Sanitation. *Environmental Policy*, 24(2), 45-57.
- Chung, E. C., & Kim, E. J. (2014). Development of Creative Design Process Based on Integrated Scenario/Prototyping Model. *Design Forum*, 44, 417-430.
- Halim, R. E. (2013). The Effect of Perception and Attitude Toward Consumer Complaint Behavior. *Journal of Distribution Science*, 11(9), 17-24.
- Kang, M. J., & Lee, S. J. (2014). Abduction and Design Thinking -Theory and Practice of Creative Ideation-. *Semiotic Inquiry*, 54, 7-35.
- Lee, K. E., & Kim, K. Y. (2014). A Study on the Factors of Civil Petitions & Complaints within Public Libraries. *Journal of the Korean Society for Library and Information Science*, 48(3), 281-301.
- Lee, K. J., Noh, Y. H., Yoon, S. K., & Cho, Y. S. (2014).

- Structuring of unstructured big data and visual interpretation. *Journal of the Korean Data & Information Science Society*, 25(5), 1431-1438.
- Lee, S. Y., & Kim, Y. I. (2009). A Study on the establishment of VOC system in compliance with the shift in customer trend. *Journal of Distribution Science*, 7(2), 89-119.
- Lee, S. W., Suh, E. K., & Suh, K. S. (2017). Applying Design Thinking and Design Science to Information System Interface Research. *The Journal of Internet Electronic Commerce Research*, 17(3), 19-35.
- Lee, W. S., & Sohn, S. Y. (2013). Predicting the Pattern of Technology Convergence using Big-Data Technology on Large-Scale Triadic Patents. Proceedings of *Fall Conference on Korea Industrial Engineering Society*, 1153-1170.
- Lee, M. S., & Kim, N. K. (2016). Investigating Dynamic Mutation Process of Issues Using Unstructured Text Analysis. *Journal of Intelligent Information System*, 22(1), 1-18
- Min, J. H., & Bae, J. H. (2015). The Impact of Big Data Investment on Firm Value. *Journal of Distribution Science*, 13(9), 5-11.
- Park, H. J., Choi, B. K., Nah, Y. W., & Shin, Y. S. (2015). An Analysis of Environment Complaint Using the Spatial Big Data Mining in Incheon. *Journal of the Korean society for Geospatial Information Science*, 2015(9), 224-225.
- Park, J. K. (2010). An Analysis on the Traffic Problems of the Urban City. *The Journal of the Korea Contents Association*, 10(12), 254-266.
- Park, K. K., & Jung, J. H. (2012). A Study on the Determining Factors of the On-Line Civil Administrative Service into Civil Satisfaction. *The Korean Journal of Local Government Studies*, 16(4), 419-420.
- Song, H. J., Park, K. S., Jung, H. E., & Song, M. (2013). Trend Analysis of Korean Economy in the Economic Literature by text mining techniques. *Korea Information Management Community*, 8, 47-50.