

유해화학물질 취급시설 관련 국민신문고 빅데이터 분석

박유현*, 김연진**, 조승범***¹, 류태인***², 김경배****

Big Data Analysis of National Complaints Portal on Hazardous Chemical Handling Facilities

Yoohyun Park*, Yeonjin Kim**, Seungbum Jo***¹, Tae In Ryu***², and Gyoungbae Kim****

본 연구는 환경부 화학물질안전원 주관으로 서원대학교 산학협력단에서 수행한 “화학안전 기술 선진화 방안 마련 연구”의 지원을 받아 수행된 연구임

요약

국민권익위원회가 운영하는 「국민신문고」 시스템은 정부에 대한 모든 민원·제안·신고 및 정책토론 등을 인터넷으로 간편하게 신청하고 처리하는 범정부 국민 소통 포털이다. 국민신문고는 국민 개인의 의견을 청취하고 민원을 해결하는 데 중요한 역할을 하고 있으며, 국민신문고에 접수된 대량의 민원 데이터를 빅데이터 분석을 통해 국민들의 요구사항을 파악하고, 정부와 관련 기관들이 이에 대한 대응 방안을 마련하는 데 도움이 될 수 있다. 국민신문고와 같은 공공기관을 대상으로 빅데이터 분석을 진행한 사례들이 다수 있었지만, 대부분은 전체 민원 또는 특정 기관 대상으로 키워드 분석을 하였고 특정 목적의 세부 분석은 다소 부족하였다. 본 논문에서는 국민신문고의 민원을 접수하는 기관 중의 하나인 A 기관의 사례를 통해, 민원에 관련된 규정별로 키워드 분석하여 민원이 집중되고 있는 규정, 민원에 관련된 연관 규정, 해당 규정에서의 주요 키워드들을 추출하여 향후 규정 개정 등에 데이터 기반의 개선을 할 수 있도록 하였다.

Abstract

The ePeople site, run by the Anti-Corruption & Civil Rights Commission, is a portal for cross-governmental communication that allows people to easily file and manage complaints, suggestions, reports, and policy discussions related to the government through the internet. This platform is essential in hearing individual opinions and resolving complaints, and the wealth of data received can be analyzed to identify public demands, helping government agencies develop effective response strategies. Several cases of big data analysis targeting public institutions such as ePeople site have been conducted. However, most of these cases primarily focused on keyword analysis of overall complaints or specific institutions, with somewhat limited in-depth analysis for specific purposes. In this paper, we present a case study of Institution A, one of the agencies responsible for receiving complaints through the ePeople site. Through keyword analysis categorized by regulations related to complaints, we identified concentrated regulations, associated regulations, and key keywords within those regulations. This analysis aims to provide a data-driven basis for future regulatory improvements, including revisions, in order to address complaints more effectively.

Keywords

big data analysis, people's complaints, hazardous chemical, text mining, natural language processing

* 동의대학교 컴퓨터소프트웨어공학 교수
- ORCID: <https://orcid.org/0000-0002-1579-6625>
** 서원대학교 산학협력단 연구원
- ORCID: <https://orcid.org/0009-0003-4070-538X>
*** 화학물질안전원 사고예방심사2과
- ORCID¹: <https://orcid.org/0000-0003-3141-9464>
- ORCID²: <https://orcid.org/0000-0002-6542-6796>

**** 서원대학교 소프트웨어응용 교수(교신저자)
- ORCID: <https://orcid.org/0000-0001-6131-8846>
· Received: Jul. 27, 2023, Revised: Oct. 10, 2023, Accepted: Oct. 13, 2023
· Corresponding Author: Gyoungbae Kim
Dept. of Software, Seowon University, 28674, Musimseo-ro 377-3,
Cheongju, Chungcheongbuk-do, Korea
Tel.: +82-43-299-8952, Email: gbkim@seowon.ac.kr

I. 서 론

국민권익위원회가 운영하는 「국민신문고」 시스템 [1]은 정부에 대한 모든 민원·제안·신고 및 정책토론 등을 인터넷으로 간편하게 신청하고 처리하는 범정부 국민 소통 포털이다. 「국민신문고」 시스템은 모든 행정기관, 사법부 및 주요 공공기관과 연결되어 원-스톱 소통 서비스 제공하고 있으며 각급 행정기관은 국민신문고 웹프레임 연계를 통해 자체 홈페이지에서 민원 등 신청·접수·처리결과 통지 서비스를 제공하고 있다. 「국민신문고」 시스템과 연동하여 민원분석 시스템을 운영하고 있는데, 민원분석 시스템은 주, 월간 단위 민원동향 분석 및 전 행정기관(일부 유관기관 포함) 배포, 민원 확산 및 피해 예방을 지원하기 위한 민원 예보 발령, 특정 주제·분야에 관한 이슈, 기획 분석결과 보고, 부처 요청, 범정부 주요 회의 등 개최 시에 맞춤형 민원 분석을 지원하는 기능을 수행한다. 또한, 2020년 개통된 「차세대 국민신문고 민원 시스템」은 인공지능을 활용한 빅데이터 분석, 인공지능 기반 자연어 처리기술을 활용하여 다양한 추천기능 및 병합 처리기능을 장착하여 운영되고 있다[2].

국민신문고는 국민 개인의 의견을 청취하고 민원을 해결하는 데 중요한 역할을 하고 있으며, 국민신문고에 제출된 대량의 민원 데이터를 빅데이터 분석을 통해 분석하면 국민들의 요구사항을 파악하고, 정부와 관련 기관들이 이에 대한 대응 방안을 마련하는 데 도움이 될 수 있다. 국민신문고에서 수집된 민원 데이터를 분석하면 민원 제기 대상, 제기 경로, 처리 속도 등을 파악할 수 있다. 이를 통해 필요한 경우 추가적인 법안 제정을 통해 민원을 근원적으로 해결할 수 있는 방안을 제시할 수도 있으며, 기존 법령의 적용이 국민의 일상생활에 얼마나 큰 영향을 미치는지 계량적으로 파악할 수 있다.

국민신문고와 같은 공공기관을 대상으로 빅데이터 분석을 진행한 사례들이 다수 있었지만, 대부분은 키워드 중심의 분석에 그쳤다. 본 논문에서는 국민신문고의 민원을 접수하는 기관 중의 하나인 A기관의 사례를 통해, 민원에 관련된 규정별로 키워드 분석을 하고자 한다. 이를 통해 A 기관에 접수되는 민원에 관련이 많은 규정 등을 파악할 수 있

고 이를 통해 규정 개정 등 후속조치의 방향을 설정할 수 있다.

본 논문의 구성은 다음과 같다. II장에서는 국민신문고 빅데이터 분석에 관한 기존의 관련연구를 살펴보고, III장에서는 본 논문의 분석 대상이 되는 국민신문고 중 A기관에 처리한 민원 데이터의 특징 및 분석 방법, 분석결과에 대해서 설명한다. 마지막으로 IV장에서는 결론을 맺는다.

II. 관련 연구

국민신문고 데이터를 대상으로 교통경찰에 대한 국민의 치안 수요를 탐색한 연구[3]에서는 교통경찰 관련 국민신문고 민원데이터를 대상으로, 주요 토픽을 추출하고 높은 비중을 차지한 위반 신고에 대해 분석하였다. 분석 결과 교통경찰 관련 민원은 시설 개선, 신호에 따른 교차로 통행방법, 번호판 영치, 개인형 이동장치 등의 토픽으로 분류하였고, 이를 통해 교차로 내 위반과 이륜자동차의 위반에 대한 단속을 강화하고 무인 교통 단속장비, 횡단보도, 신호등의 설치 및 운영에 대한 선제적인 조치, 최근 개정된 법령과 시행된 정책, 경찰 교통민원 사이트, 단속 사후 절차에 대한 더욱 활발한 홍보가 필요하다고 제안하였다.

청와대 홈페이지의 국민청원 자료를 분석하여 유아교사의 요구 및 관심사를 알아보고자 수행한 연구[4]에서는 유아교사들은 인건비 및 복지 개선, 교사 대 유아 비율 감축 등 지속적인 요구를 제기하고 있었으며, 사립유치원의 비리와 코로나19 등 각종 사회적 이슈에 따른 즉각적인 요구 역시 제기하고 있음을 확인하였다.

안전신문고를 대상으로 안전신문고의 신고내용이 향후 일어날 수 있는 재난과 연관성을 조사한 연구[5]에서는 안전신문고에 신고된 자료와 안전신문고 분석 기간 동안의 신문기사를 분석하여 안전신문고와 신문 기사 내용 간의 상관관계를 분석하였다. 그 결과 응답 및 확인 관련 보고서의 수가 증가함에 따라 몇 달 내 사고가 발생하였으며, 사회의 불안에 대해 사전에 보고된 안전신문고의 내용을 분석하면 미래 재난 예측에 활용될 수 있을 것이라 예상하였다.

민원 데이터를 대상으로 분석한 연구로는 전주시의 민원 데이터를 활용하여 환경·위생 분야의 정책 시행에서 데이터 기반 의사결정을 제안한 연구[6]와 진주시의 전자민원 데이터를 분석하여 연중 민원 집중 시기, 연간 민원량 예측, 지역에 따른 주요 민원 사유 등을 조사한 연구[7][8]도 있었다. 또한, 서울시에서 운영하는 전자 민원 사이트인 응답소에 게재된 교통 민원 데이터를 대상으로 토픽 모델링과 소셜 네트워크 분석 기법을 활용하여 연도별 주요 민원 주제와 각 주제에 해당하는 민원의 의미를 탐색한 연구[9]와 천안시 민원 데이터에 소셜 네트워크 분석을 적용해 계절별, 그리고 민선 기수별 주요 민원 주제와 키워드의 차이를 분석[10]한 연구가 있었으며, 소셜네트워크 분석 기법으로 부산시에 접수된 버스 관련 민원의 패턴을 분석한 연구[11]도 있었다. 또한, 부산시 민원 게시판의 민원 데이터를 활용하여 TF-IDF와 공기어(CoOccurrences) 구조를 분석해 민원 데이터로부터 부산 시민들의 요구사항을 체계적으로 분석[12][13]한 연구와 부산시 민원 데이터에 토픽 모델링을 적용하여 분기별 민원 동향을 파악[14]한 연구가 있었다. 그리고, 딥러닝을 이용한 텍스트 분석 기법으로 부산시 민원을 자동으로 분류하는 모델을 보고한 연구[15]와 인천광역시의 민원 데이터를 활용하여 인천 시민의 공적 관심사를 분석, 시민들이 인천시에 요구하는 시정 운영 사항들을 분석하고자 국민신문고에서 민원 데이터를 수집하여 워드 임베딩과 K-평균 군집화를 사용해 민원 텍스트를 분석한 연구[16]도 있었다.

국민신문고 자체적으로도 다양한 수준의 빅데이터 분석을 시행하고 있으며 「한눈에 보는 민원 빅데이터」 사이트에서 시각화하여 제공하고 있다[17]. 해당 사이트에서는 지역별, 기관별, 키워드별, 분야별 통계자료를 제공하고 있지만, A 기관의 독자적인 통계 분석 자료는 제공되지 않고 있다.

민원 데이터와 다른 종류이지만 온라인 리뷰 분석에 관한 연구들도 지속적으로 진행되어 왔다. 특히 리뷰 내용에서 가짜 리뷰(Fake review)를 LSTM 등 다양한 인공지능 기법을 활용하여 탐지하는 방법[18]도 제시되었다.

본 논문에서는 국민신문고의 A 기관에 관한 민원을 대상으로 다양한 기준으로 분석을 하고자 한

다. 이를 통해 A 기관에 접수되는 민원에 관련이 많은 규정 등을 파악할 수 있고 데이터 기반의 규정 개정 등 후속조치의 방향을 설정할 수 있다.

III. 국민신문고 A기관 민원 데이터 분석

국민신문고 데이터를 기반으로 「한눈에 보는 민원 빅데이터」 사이트[17]에서는 기관별, 지역별 등 민원 빅데이터 자료를 시각화 하여 제공하고 있다. 하지만, 단순히 시계열 민원 건수, 10여개의 키워드 클라우드, 민원인의 지역, 성별, 연령, 분야 등의 정보만을 제공하고 있어 각 기관별로 특성에 맞는 데이터 분석이 추가적으로 필요하다.

A기관은 화학사고를 예방, 대응하는 환경부 소속 전문 기관으로 환경부와 유역·지방환경청, 전국 7개 합동방재센터를 지원하고 화학사고·테러 시 전문 인력과 장비, 위험범위 예측 평가, 과학적 대응 기술과 정보를 제공하기 위하여 설립되어 운영되고 있다. 국민신문고에 접수된 민원 중 A기관에 관련된 민원들이 A기관으로 전달되고, 해당 부서에서 민원에 응대하고 있다. A기관에서 처리하는 민원에서 특정 사항에 대해 법령의 해석이나 화학물질의 보관, 취급 등에 대한 세부적인 질의가 많은 편이다. 이러한 이유로 본 연구에서는 민원을 관련 법령으로 분류하고 해당 법령에 관한 민원 대상으로 키워드 분석을 하고자 한다.

3.1 분석 대상 데이터

분석대상 국민신문고 민원은 2014년 2월부터 2022년 5월까지의 총 5,534건이며 자세한 내용은 표 1과 같다.

A 기관이 접수한 민원은 총 53개의 필드로 구성되어 있으며, 이 중에서 분석에 유의미한 필드는 민원 제목, 날짜, 내용, 요약, 답변, 담당부서 등 8개이다.

A 기관에 접수된 민원들은 표 2의 부서에서 처리되고 있으며, 가장 많은 민원이 접수되고 있는 부서는 「사고예방심사2과」로, 전체 민원의 77%에 해당한다. 「사고예방심사2과」는 내부적으로 국민신문고 민원처리를 가장 많이 하고 있는 화학안전 규제개선TF팀, 화학사고예방관리계획서 적합 이후

14 유해화학물질 취급시설 관련 국민신문고 빅데이터 분석

사업장에서 이행해야 하는 부분을 총괄 관리하는 이행관리팀, 화관법 관련, 안전원 고시법 개정 관리, 화관서 통계 관리, 사업장 정보 관리, 민원 관리 제도 관리 중 나오는 문제점 및 해결법 관리하는 제도운영팀이 있다.

본 논문에서는 전체 민원 중 비율이 가장 높은 「사고예방심사2과」의 담당 민원을 대상으로 분석하고자 한다.

표 1. A기관의 국민신문고 민원 수

Table 1. Number of complaints about public institution A

no.	year	# of complaints	ratio(%)
1	2014	5	0.1
2	2015	98	1.8
3	2016	120	2.2
4	2017	734	13.3
5	2018	1,121	20.3
6	2019	1,353	24.4
7	2020	937	16.9
8	2021	916	16.6
9	2022	250	4.5
total		5,534	100

표 2. 담당부서별 민원 수

Table 2. Number of complaints by department

no.	department	#of complaints	ratio(%)
1	Education and training innovation team	159	2.87
2	Planning and operations division	6	0.11
3	Incident response coordination division	633	11.44
4	Incident prevention assessment unit 1	451	8.15
5	Incident prevention assessment unit 2	4,285	77.43
total		5,534	100

3.2 A 기관 민원 관련 규정

「사고예방심사2과」에서 처리하는 민원 중에서 관련 규정에 관련된 것들은 대부분 표 3의 내용을 포함한다. 2021년 4월 27일 환경부고시 제2020-204, 205호인 「위해관리계획서 작성 등에 관한 규정」과 「장외영향평가서 작성 등에 관한 규정」은 폐지되고, 안전원고시 제2021-07호 「화학사고예방관리계획서작성 등에 관한 규정」으로 운영하고 있다.

3.3 민원 데이터 분석 과정

A 기관이 접수한 민원 구성 중 민원 제목, 날짜, 내용, 요약, 답변, 담당부서 등 8개 필드를 대상으로 Python으로 주요 항목을 추출하였다. Python 프로그램에서 사용한 라이브러리들은 표 3과 같다. 데이터 분석을 위해서 Pandas와 Numpy, Matplotlib을 사용하였고, 상관도 분석을 위한 다중 라벨 이진화 (MultiLabel Binarizer)하는 전처리를 위해서 Scikit-learn을 사용하였다. 또한, 한국어 형태소 분석을 위해 KoNLPy를 사용하였다.

표 3. 데이터 분석 활용 Python 라이브러리

Table 3. Python libraries used in data analysis

Library name	Description
Pandas	Python library tailored for handling large datasets, enabling efficient data manipulation, analysis, and exploration
Numpy	Python library for numerical computing, providing efficient array operations and mathematical functions essential
Matplotlib seaborn	Python library used for creating high-quality visualizations and plots
sklearn	Python library offering a comprehensive set of tools for machine learning tasks
KoNLPy	Python library specifically designed for natural language processing (NLP) tasks in Korean
hanspell	Korean spelling and grammar checking library in Pythonbig data applications

민원 데이터를 분석하기 위해 먼저, 민원의 제목, 내용, 요약, 답변 데이터를 대상으로 키워드별 빈도를 구하고 취급시설 및 화학물질명 사전을 생성하였다.

민원 데이터에서 민원 제목, 민원 내용 등 국민청원자가 작성한 글은 특정한 규칙이 없이 자유롭게 작성되는 반면, 담당자가 작성한 답변은 대부분 구조적으로 구성되어 있다. 특히, 답변의 내용 중에서는 답변자의 이름, 소속뿐만 아니라 관련 규정 등에 대해서도 명시하는 경우가 많다.

따라서, 해당 민원이 어떤 규정에 관련된 민원인지를 분석하기 위해 민원의 답변에서 표 4의 규정의 주요 키워드들을 사용하여 분류하였다.

표 4. 「사고예방심사2과」 업무 관련 주요 규정

Table 4. Main regulations related to 「Accident prevention and assessment division 2」

class		name of regulation	
취급시설	시설 유형 기준 (9종)	안전원고시 제2020-7호	유해화학물질 실내 보관시설 설치 및 관리에 관한 고시
		안전원고시 제2020-6호	유해화학물질 실내 저장시설 설치 및 관리에 관한 고시
		안전원고시 제2020-9호	유해화학물질 실외 보관시설 설치 및 관리에 관한 고시
		안전원고시 제2020-8호	유해화학물질 실외 저장시설 설치 및 관리에 관한 고시
		안전원고시 제2020-5호	유해화학물질 제조·사용시설 설치 및 관리에 관한 고시
		안전원고시 제2020-10호	유해화학물질 지하 저장시설 설치 및 관리에 관한 고시
		안전원고시 제2019-11호	유해화학물질 차량 운반시설 설치 및 관리에 관한 고시
		안전원고시 제2020-11호	유해화학물질 차량 운송시설 설치 및 관리에 관한 고시
		안전원고시 제2020-12호	유해화학물질 사외배관 이송시설 설치 및 관리에 관한 고시
	업종별 기준	안전원고시 제2021-2호	표면처리업종 유해화학물질 취급시설설치 및 관리에 관한 고시
		안전원고시 제2021-3호	염색업종 유해화학물질 취급시설설치 및 관리에 관한 고시
		안전원고시 제2022-2호	항만구역 내 유해화학물질 보관시설설치 및 관리에 관한 고시
	소량 기준	안전원고시 제2021-5호	유해화학물질 소량 취급시설에 관한 고시
		안전원지침 제2021-1호	유해화학물질 소량 취급시설의 설치·정기·수시검사의 방법 등에 관한 세부지침
물질별		안전원 고시 제2022-4호	유해화학물질별 구체적인 취급기준에 관한 규정
화학사고 예방관리 계획서	작성 (3종)	안전원고시 제2021-07호	화학사고예방관리계획서작성 등에 관한 규정
		안전원지침 제2021-2호	사고 영향범위 산정에 관한 기술지침
		안전원지침 제2021-3호	사고시나리오 선정 및 위험도 분석에 관한 기술 지침
	검토	안전원고시 제2021-08호	화학사고예방관리계획서 검토 등에 관한 규정
		안전원고시 제2021-09호	화학사고예방관리계획서 이행 등에 관한 규정
		환경부고시 제2020-205호	장외영향평가서 작성 등에 관한 규정
		환경부고시 제2020-204호	위해관리계획서 작성 등에 관한 규정

예를 들어 특정 민원이 가장 많은 민원이 집중된 「안전원지침 제2021-1호」와 관련된 민원인지를 판단하기 위해서 먼저 지침명인 “안전원지침 제2021-1호”를 포함하거나, “소량 취급시설”과 “설치”, “정기검사”, “수시검사”라는 용어가 함께 나타나는지를 조사하였다. 이러한 조건을 만족하는 민원인 경우는 「안전원지침 제2021-1호」와 관련된 민원으로 분류하였고, 이러한 방식으로 분류한 결과는 표 5와 같다.

지침 2021-1호에 관련된 민원이 총 988건으로 23%에 해당되며, 고시 제2020-5호에 관련된 민원이 850건으로 20%, 고시 제2021-5호에 관련된 민원은 379건, 고시 제2020-8호에 관련된 민원은 341건으로 이들 4개 규정에 관한 민원이 전체의 60% 정도 차지하고 있는 것으로 분석되었다. 안전원 고시 제2021-1호에 관련된 민원이 많은 이유로는 다른 고시들은 시설의 설치 및 관리에 관한 내용인 반면, 해당 고시는 설치 및 검사 방법에 관한 세부지침으로 배관, 안전밸브, 검지경보설비, 피해저감시설, 스크러버 등 검사항목이 다양하기에 검사 세부 방법에 대

한 질문 사항이 많기 때문으로 보인다. 화학물질중합정보시스템의 2020년 통계조사 자료에 따르면[19], 유해화학물질을 취급하는 30,961개의 사업장 중 제조업이 18,108개로 이와 연관된 유해화학물질 제조·사용시설 설치 및 관리에 관한 고시인 제2020-5호에 관한 민원이 많이 발생한 것으로 보인다.

하나의 민원에 대해서 두 개 이상의 규정으로 분류되는 경우도 있기 때문에 표 5의 민원 수 합은 표 1, 2에 나타난 총 민원 수의 합보다 크다. 하나의 민원에 대해서 두 개 이상의 규정으로 분류된 결과는 표 6와 같다.

또한, 규정 간 상관도를 측정한 결과는 표 7과 같다. 결과와 같이 「안전원고시 제2021-7호」와 「안전원고시 제2021-8호」의 상관도(PCC, Pearson Correlation Coefficient)가 매우 높은 것을 알 수 있다. 피어슨 상관 계수(PCC)[20]란 두 변수 X 와 Y 간의 선형 상관 관계를 계량화한 수치로, +1과 -1 사이의 값을 가진다. +1은 완벽한 양의 선형 상관 관계, 0은 선형 상관 관계 없음, -1은 완벽한 음의 선형 상관 관계를 의미한다.

표 5. 규정별 민원 수

Table 5. Number of complaints by regulation

name of regulation	#of complaints	ratio(%)
안전원지침 제2021-1호	988	23.1
안전원고시 제2020-5호	850	19.8
안전원고시 제2021-5호	379	8.8
안전원고시 제2020-8호	341	8
안전원고시 제2020-6호	258	6
환경부고시 제2020-205호	209	4.9
안전원고시 제2020-7호	157	3.7
환경부고시 제2020-204호	132	3.1
안전원고시 제2021-08호	101	2.4
안전원고시 제2021-07호	99	2.3
안전원고시 제2020-11호	83	1.9
안전원지침 제2021-3호	38	0.9
안전원고시 제2020-9호	34	0.8
안전원고시 제2020-10호	34	0.8
안전원고시 제2020-12호	31	0.7
안전원고시 제2019-11호	20	0.5
안전원고시 제2021-2호	13	0.3
안전원고시 제2021-09호	9	0.2
안전원지침 제2021-2호	5	0.1
안전원고시 제2021-3호	2	0
안전원고시 제2022-2호	0	0
안전원고시 제2022-4호	0	0
미분류	1974	46.1

「안전원고시 제2021-7호」, 「안전원고시 제2021-8호」, 「안전원고시 제2021-9호」의 경우 화학사고예방관리계획서의 작성, 이행, 검토에 관한 규정으로 화학사고예방관리계획서 작성 시 3가지 고시가 모두 함께 검토되는 경우가 많기 때문에 상관도가 높게 나온 것으로 보인다.

한편, 표 5에서 접수된 민원 중에서는 단순 문의나 질의 취하 등 규정과 연관이 없거나 규정명에 들어 있는 키워드로 분류가 되지 않는 사항들이 46% 차지하고 있다.

3.4 주요 규정별 세부 분석

3.4에서는 국민신문고 접수가 가장 많은 「안전원지침 제2021-1호」와 「안전원지침 제2020-5호」의 세부 분석 내용을 기술한다.

표 6. 복수 연관 규정 민원 수

Table 6. Number of complaints by multiple regulations

names of regulations	#
안전원고시 제2021-5호, 안전원지침 제2021-1호	152
안전원고시 제2020-5호, 안전원고시 제2021-5호, 안전원지침 제2021-1호	92
안전원고시 제2020-5호, 안전원지침 제2021-1호	74
안전원고시 제2020-6호, 안전원고시 제2020-5호	37
안전원고시 제2020-8호, 안전원고시 제2020-5호	37
안전원고시 제2020-5호, 안전원고시 제2021-5호, 안전원지침 제2021-1호, 환경부고시 제2020-205호	29
안전원고시 제2020-6호, 안전원고시 제2020-8호	28
안전원고시 제2020-8호, 안전원지침 제2021-1호	27
안전원고시 제2021-7호, 안전원고시 제2021-8호	26
안전원고시 제2020-6호, 안전원지침 제2021-1호	22
안전원고시 제2020-7호, 안전원고시 제2020-5호	22
환경부고시 제2020-205호, 환경부고시 제2020-204호	20
안전원지침 제2021-1호, 환경부고시 제2020-205호	19
안전원고시 제2021-5호, 안전원지침 제2021-1호, 환경부고시 제2020-205호	16
안전원고시 제2020-11호, 안전원지침 제2021-1호	13
안전원고시 제2020-7호, 안전원고시 제2021-5호, 안전원지침 제2021-1호	13
안전원고시 제2021-7호, 안전원고시 제2021-8호, 환경부고시 제2020-205호	13

표 7. 규정 간 상관도 분석

Table 7. Correlation analysis between regulations

regulation #1	regulation #2	PCC
안전원고시 제2021-7호	안전원고시 제2021-8호	0.76
안전원지침 제2021-1호	안전원고시 제2021-5호	0.46
안전원고시 제2021-7호	안전원고시 제2021-9호	0.42
안전원고시 제2019-11호	안전원고시 제2020-11호	0.38
안전원고시 제2021-9호	환경부고시 제2020-204호	0.33
안전원고시 제2021-3호	안전원지침 제2021-3호	0.31
안전원고시 제2021-8호	안전원고시 제2021-9호	0.28
안전원고시 제2020-7호	안전원고시 제2020-9호	0.27
안전원고시 제2021-7호	환경부고시 제2020-204호	0.26
환경부고시 제2020-204호	환경부고시 제2020-205호	0.25
안전원고시 제2021-8호	환경부고시 제2020-204호	0.23
안전원고시 제2020-6호	안전원고시 제2020-8호	0.17
안전원지침 제2021-3호	안전원고시 제2021-5호	0.15
안전원고시 제2020-5호	안전원고시 제2021-5호	0.13
안전원고시 제2020-11호	안전원고시 제2020-8호	0.12
안전원고시 제2020-5호	안전원고시 제2021-3호	0.12
안전원고시 제2020-6호	안전원고시 제2020-9호	0.12
안전원고시 제2021-3호	안전원고시 제2021-7호	0.10

세부 분석을 위한 방법은 3절에서 분류한 규정별 민원을 대상으로 모든 키워드를 Python 프로그램을 활용하여 추출하고, 그 중 화학물질과 취급시설에 관한 내용들을 분류하였다. 화학물질 및 취급시설을 기술하는 과정에서 표준화되지 않은 표현들이 발견되어 이를 위한 추가적인 연구가 필요한 것으로 파악되었다.

주요 분석 내용은 민원 접수 날짜별 분석, 화학물질 빈도수, 취급시설 빈도수이다. 이 분석을 통해 해당 규정에 관련성이 높은 화학물질과 취급하고 있는 장소에 대한 내용을 파악할 수 있다.

3.4.1 안전원고시 제2021-1호

「안전원고시 제2021-1호」는 「유해화학물질 소량 취급시설의 설치·정기·수시검사의 방법 등에 관한 세부지침」으로 분류된 민원의 수는 988건이며, 월별 민원 수는 그림 1과 같다. 민원은 2018년 7월에 36건으로 가장 많았고, 점차 감소하는 추세를 보이고 있다. 2018년 7월부터 2019년 3월까지 민원은 22건에서 36건으로 다른 시기에 비해 민원이 많은 편인데, 이는 해당 규정의 제정일(2018.9.10.), 개정일(2019.2.11.)과 관계가 있는 것으로 보이며, 후속 개정일(2021.4.1.) 근처에서 접수 민원 수가 유의미하게 증가함을 알 수 있다.

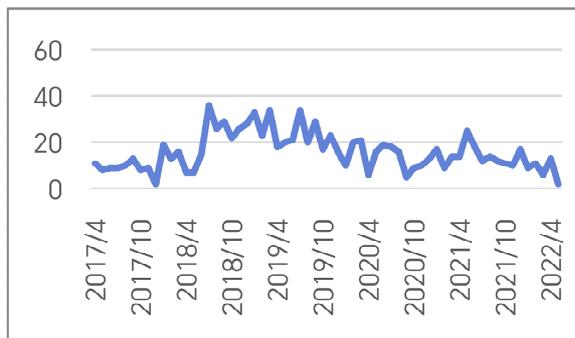


그림 1. 「안전원고시 제2021-1호」 관련 민원 접수일 분포

Fig. 1. Distribution of dates for complaints related to the 「Notices No. 2021-1」

표 8은 「안전원지침 제2021-1호」 관련 민원 내 화학물질 명 빈도수를 나타낸다. 나트륨, 수산화, 산화나트륨, 수산화나트륨은 띄어쓰기 등의 문제로 구분되어 있지만 하나의 화학물질로 보이며, 다른 화

학물질에 비해 월등히 높은 빈도를 보여주고 있다. 이외에 황산, 폐수, 독물, 염산 등의 순서로 빈도수가 많은 것을 알 수 있다. 수산화나트륨은 염기성 물질로 금속을 세척하는 도금업종이나 폐수처리 하는 곳에서 주로 사용하는 물질이기 때문에 화학물질을 취급하는 사업장에서 전반적으로 사용되고 있어 이에 대한 민원인 많이 발생한 것으로 보인다.

표 8. 「안전원고시 제2021-1호」 관련 민원 내 화학물질 명 빈도수

Table 8. Chemical substance frequency in complaints related to the 「Notices No. 2021-1」

chemical substance	freq.	chemical substance	freq.
sodium	132	nitric acid	38
hydroxide	129	hydrogen peroxide	38
sodium peroxide	123	oxidation state	38
sodium hydroxide	123	ammonia	36
sulfuric acid	95	ethyl	36
wastewater	80	methyl	29
poison	77	chlorine	21
hydrochloric acid	50	plastic	19
toxic substance	49	chloride	18
flammable substance	46	petroleum	18

표 9는 「안전원지침 제2021-1호」 관련 민원 내 취급시설의 빈도수를 나타낸다. 관련 민원 내에서 탱크, 배관, 저장탱크 등의 취급시설 순으로 빈도수가 큰 것을 알 수 있다. 이는 화학물질을 취급하는 사업장에서는 주로 저장탱크에 보관 후 사용하거나, 제조한 물질을 저장탱크에 보관하기에 저장탱크 관련 민원 비율이 높으며, 이와 연결된 배관 기준에 관한 민원도 같이 높게 나온 것을 알 수 있다.

표 9. 「안전원고시 제2021-1호」 관련 민원 내 취급시설 빈도수

Table 9. Facility frequency in complaints related to the 「Notices No. 2021-1」

facility	freq.	facility	freq.
tank	335	warehouse	81
pipeline	284	speed bump	77
storage tank	223	building	75
underground	178	concrete	62
container	157	laboratory	57
barrier	151	vehicle	55
detector	147	fireworks	53
road	134	tanker truck	45
valve	128	sewage system	42
trench	103	wall structure	36

3.4.2 안전원고시 제2020-5호

「안전원고시 제2020-5호」는 「유해화학물질 제조·사용시설 설치 및 관리에 관한 고시」로 분류된 민원의 수는 850건이며, 월별 민원 수는 아래 그림2와 같다.

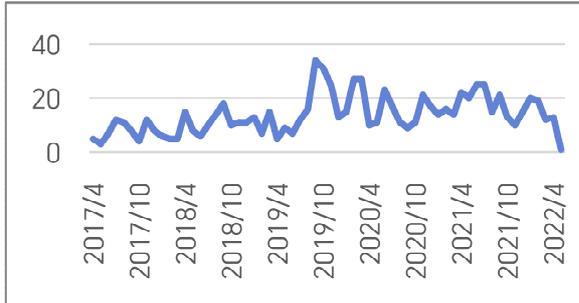


그림 2. 「안전원고시 제2020-5호」 관련 민원 접수일 분포

Fig. 2. Distribution of dates for complaints related to the 「Notices No. 2020-5」

민원 날짜 중 해당 규정의 제정일(2019.9.2.)이 있는 달의 민원 건수는 34건으로 가장 많은 것으로 조사되었다.

표 10은 「안전원고시 제2020-5호」 관련 민원 내 화학물질명 빈도수를 나타낸다. 「안전원지침 제2021-1호」와 마찬가지로 나트륨, 수산화, 산화나트륨, 수산화나트륨은 띄어쓰기 등의 문제로 구분되어 있지만 하나의 화학물질로 보이며, 다른 화학물질에 비해 월등히 높은 빈도를 보여주고 있다. 이외에 독물, 폐수, 인화성물질, 염산 등의 순서로 빈도수가 많은 것을 알 수 있다.

표 10. 「안전원고시 제2020-5호」 관련 민원 내 화학물질명 빈도수

Table 10. Chemical substance frequency in complaints related to the 「Notices No. 2020-5」

chemical substance	freq.	chemical substance	freq.
sodium	88	nitric acid	37
hydroxide	86	hydrogen peroxide	35
sodium peroxide	84	oxidation state	35
sodium hydroxide	84	hydrochloric acid	30
poison	59	ammonia	26
wastewater	59	methyl	24
flammable substance	54	chloride	19
sulfuric acid	50	ethyl	19
toxic substance	46	drainage	19
volcanic ash	38	toluene	18

표 11은 「안전원고시 제2020-5호」 관련 민원 내 취급시설의 빈도수를 나타낸다. 관련 민원 내에서 배관, 지하, 탱크, 감지기 등의 취급시설 순으로 빈도수가 큰 것을 알 수 있다.

표 11. 「안전원고시 제2020-5호」 관련 민원 내 취급시설 빈도수

Table 11. Facility frequency in complaints related to the 「Notices No. 2020-5」

facility	freq.	facility	freq.
pipeline	278	building	87
underground	241	speed bump	74
tank	231	fireworks	72
detector	174	waste water collection facilities	67
storage tank	159	warehouse	51
container	146	wall structure	50
road	119	entrance	41
valve	113	scrubber	37
barrier	98	waterworks	34
trench	97	concrete	34

IV. 결 론

국민권익위원회가 운영하는 「국민신문고」 시스템은 정부에 대한 모든 민원·제안·신고 및 정책토론 등을 인터넷으로 간편하게 신청하고 처리하는 범정부 국민 소통 포털이다.

국민신문고는 국민 개인의 의견을 청취하고 민원을 해결하는 데 중요한 역할을 하고 있으며, 국민신문고에 제출된 대량의 민원 데이터를 빅데이터 분석을 통해 분석하면 국민들의 요구사항을 파악하고, 정부와 관련 기관들이 이에 대한 대응 방안을 마련하는 데 도움이 될 수 있다. 국민신문고에서 수집된 민원 데이터를 분석하면 민원 제기 대상, 제기 경로, 처리 속도 등을 파악할 수 있다. 이를 통해 필요한 경우 추가적인 법안 제정을 통해 민원을 근원적으로 해결할 수 있는 방안을 제시할 수도 있으며, 기존 법령의 적용이 국민의 일상생활에 얼마나 큰 영향을 미치는지 계량적으로 파악할 수 있다.

국민신문고와 같은 공공기관을 대상으로 빅데이터 분석을 진행한 사례들이 다수 있었지만, 대부분은 키워드 중심의 분석에 그쳤다. 본 논문에서는 국민신문고의 민원을 접수하는 기관 중의 하나인 A기관의 사례를 통해, 민원에 관련된 규정별로 빅데이터 분석을 하였다.

A 기관에서 답변한 2014년부터 2022년까지 국민신문고 데이터 총 5,534건을 대상으로 분석한 결과, 대부분(4,285건, 77.43%)은 사고예방심사2과 담당의 질의이며, 질의가 많은 고시는 유해화학물질 소량 취급시설의 설치·정기·수시검사의 방법 등에 관한 세부지침(안전원고시 제2021-1호)으로 988건(23.1%)이며, 질의 중 연관 고시 중 화학사고예방관리계획서작성 등에 관한 규정(안전원고시 제2021-7호)와 화학사고예방관리계획서 검토 등에 관한 규정(안전원고시 제2021-8호)의 상관도가 가장 높았다. 또한, 주요 규정에 대해 주요 화학물질 및 취급시설에 대한 키워드 분석을 수행하였다.

본 논문은 국민신문고를 대상으로 관련 규정 중심으로 분석하여 각 규정에 대한 국민 활용도를 파악하는데 기본적인 데이터를 제공하는 데 기여점이 있다. 하지만, 자연어로 작성된 질의 및 답변을 기반으로 관련 규정을 정확히 분류하지 못한 한계점을 가진다. 이를 해결하기 위해서는 질의 내용을 정확히 파악하여 답변하는 답변자가 관련 규정을 명시할 필요가 있다. 또한 화학물질명에 대한 데이터베이스 구축이 추가적으로 필요하다. 분석 대상의 화학물질로 수산화나트륨, 황산, 염산 등의 화학물질을 한정할 수도 있고, 폐수, 독물, 유독물질 등의 용어도 포함시킬 수도 있을 것이다. 마찬가지로, 취급 시설명에 대한 데이터베이스 구축도 필요할 것으로 예상된다.

References

- [1] e-people, <https://www.epeople.go.kr/index.jsp> [accessed: Apr. 09, 2023]
- [2] S.-J. Eom, "Democracy and Public Administration in the Age of Artificial Intelligence: Focusing on the case of 「e-People」 System Advancement", *The Korean Journal of Public Administration*, Vol. 30, No. 2, pp. 35-64, Jun. 2021. <http://dx.doi.org/10.22897/kipajn.2021.30.2.002>.
- [3] S. Lee, "An Analysis of Civil Complaints about Traffic Policing Using the LDA Model", *The Journal of The Korea Institute of Intelligent Transport Systems*, Vol. 20, No. 4, pp. 57-70, Aug. 2021. <http://dx.doi.org/10.12815/kits.2021.20.4.57>.
- [4] H. R. Yu and J. Lee, "Content Analysis of Early Childhood Teachers' Interests and Needs: Focused on Their Petitions Posted on Korean National Petition Bulletin Board", *Childhood Education Research & Review*, Vol. 25, No. 6, pp. 169-193, Dec. 2021. <http://dx.doi.org/10.32349/ECERR.2021.12.25.6.169>.
- [5] J. Lee, J. Shin, S. Cho, and S. Lee, "Research Suggestion for Disaster Prediction using Safety Report of Korea Government", *Journal of Korean Society of Disaster and Security*, Vol. 12, No. 4, pp. 15-26, Dec. 2019.
- [6] H. O. Choi, "Study on Selecting Priority Criteria Utilizing Civil Complaint Data in the Field of Environment and Sanitation", *Journal of Environmental Policy and Administration*, Vol. 24, No. 2, pp. 45-57, Jun. 2016. <http://dx.doi.org/10.15301/jepa.2016.24.2.45>.
- [7] T. H. Won and H. H. Yoo, "Pattern Analysis for Civil Complaints of Local Governments Using a Text Mining", *Journal of the Korean Society of Surveying, Geodesy, Photogrammetry and Cartography*, Vol. 34, No. 3, pp. 319-327, Jun. 2016. <http://dx.doi.org/10.7848/ksgpc.2016.34.3.319>.
- [8] T. Won, H. Yoo, and M. Seo, "Spatial Pattern and Trend Analysis of Parking-related Electronic Civil Complaints in Jinju-Si", *Journal of Cadastre & Land Information*, Vol. 47, No. 1, pp. 5-14, Jun. 2017. <http://dx.doi.org/10.22640/lxsiri.2017.47.1.5>.
- [9] J. Y. Kim and Chang Justin S, "Analysing Civil Traffic Complaints using Latent Dirichlet Allocation", *Proc. of the KOR-KST Conference*, pp. 106-111, Sep. 2018.
- [10] I. Jeon, O. Jun, M. Choi, H. Kim, and J. Chung, "Characteristics of Civil Complaints to a Local Government based on Social Network Analysis : Focused on Cheonan City E-Bulletin Board (Allso 365)", *Journal of Regional Studies*, Vol. 25, No. 2, pp. 117-141, Jun. 2017.

[11] S. E. Ryu, S. Hong, T. Lee, and N. Kim, "A Pattern Analysis of Bus Civil Complaint in Busan City Using the Text Network Analysis", Korean Computers and Accounting Review, Vol. 16, No. 2, pp. 19-43, Aug. 2018. <http://dx.doi.org/10.32956/kaoca.2018.16.2.19>.

[12] H. Jeong, T. Lee, and S. Hong, "A Copus Analysis of Electronic Petitions For Improving the Responsiveness of Public Services : Focusing on Busan Petition", The Korean Journal of Local Government Studies, Vol. 21, No. 1, pp. 423-436, May 2017. <http://dx.doi.org/10.20484/klog.21.1.17>.

[13] H. Kim, T. H. Lee, S. E. Ryu, and N. R. Kim, "A Study on Text Mining Methods to Analyze Civil Complaints - Structured Association Analysis", Journal of the Korea Industrial Information Systems Research, Vol. 23, No. 3, pp. 13-24, Jun. 2018. <http://dx.doi.org/10.9723/jksiis.2018.23.3.013>.

[14] J.-S. Park and S.-M. Lee, "Big Data Analysis of Busan Civil Affairs Using the LDA Topic Modeling Technique", Informatization Policy, Vol. 27, No. 2, pp. 66-83, Jun. 2020. <https://doi.org/10.22693/NIAIP.2020.27.2.066>.

[15] N. Kim, and M. R. Adyan, "Automatic Classification of Civil Complaint Data Using CNN and Bidirectional LSTM : The Case of Busan, South Korea", Korean Computers and Accounting Review, Vol. 17, No. 2, pp. 81-98, Dec. 2019. <http://dx.doi.org/10.32956/kaoca.2019.17.2.81>.

[16] J. Lee and C. Kim, "What Incheon Wants: Analysis of Citizens' Opinions Using Civil Petitions Data from e-People", Korean Journal of Urban Studies, Vol. 20, pp. 183-223, Dec. 2021. <http://dx.doi.org/10.34165/urbanr.2021..20.183>.

[17] big data visualization of e-people, <https://bigdata.epeople.go.kr/bigdata/bigMainPage.npaid> [accessed: Apr. 09, 2023]

[18] E. F. Cardoso, R. M. Silva, T. A. Almeida, "Towards automatic filtering of fake reviews", Neurocomputing, Vol. 309, pp. 106-116, Oct.

2018. <https://doi.org/10.1016/j.neucom.2018.04.074>.

[19] Integrated Chemical Information System, <https://icis.me.go.kr> [accessed: May 18, 2023]

[20] W. Kirch, "Pearson's Correlation Coefficient", Encyclopedia of Public Health, pp. 1090-1091, 2008. https://doi.org/10.1007/978-1-4020-5614-7_2569.

저자소개

박 유 현 (Yoohyun Park)



1996년 2월 : 부산대학교

전자계산학과(학사)

1996년 2월 : 부산대학교

전자계산학과(석사)

2008년 8월 : 부산대학교

전자계산학과(박사)

2000년 1월 ~ 2000년 12월 :

한국국방연구원 연구원

2001년 1월 ~ 2009년 2월 : 한국전자통신연구원

선임연구원

2009년 3월 ~ 현재 : 동의대학교 컴퓨터소프트웨어공학과
교수

관심분야 : 클라우드, 빅데이터, IT융합서비스

김 연 진 (Yeonjin Kim)



2021년 2월 : 경북대학교

나노소재공학부 신소재공학전공
(학사)

2023년 8월 : 서원대학교

정보통신공학(석사)

2023년 4월 ~ 현재 : 서원대학교

산학협력단 연구원

관심분야 : 빅데이터, 인공지능, 재난재해

조 승 범 (Seungbum Jo)



2004년 2월 : 아주대학교

기계공학과(학사)

2015년 12월 : 캘리포니아대학

샌디에고 기계공학과(박사)

2019년 4월 ~ 현재 :

화학물질안전원 공업연구관

관심분야 : 빅데이터, 전산유체,

안전기준, 방재기술

류 태 인 (Tae In Ryu)



2008년 2월 : 부산대학교 화학과
(학사)

2016년 8월 : 성균관대학교
화학공학과(박사)

2019년 1월 ~ 현재 : 환경부
화학물질안전원 공업연구사

관심분야 : 화학물질 분석,

화학물질 특성 예측, 빅데이터, 머신러닝

김 경 배 (Gyoungbae Kim)



1992년 2월 : 인하대학교
전자계산공학과(학사)

2000년 2월 : 인하대학교
컴퓨터공학(박사)

2000년 3월 ~ 2004년 2월 :
한국전자통신연구원 선임연구원

2004년 3월 ~ 현재 : 서원대학교

소프트웨어학부 교수

관심분야 : 빅데이터, 인공지능, 재난재해, 클라우드
컴퓨팅