

텍스트마이닝을 활용한 교통영향평가 교통개선대책 분석 : 경기도 도시개발사업을 대상으로

Analysis of Traffic Improvement Measures in Transportation Impact Assessment Using Text Mining : Focusing on City Development Projects in Gyeonggi Province

양 은 혜* · 강 희 찬** · 안 우 영***

* 주저자 : 국립공주대학교 도시융합시스템공학과 박사과정
** 교신저자 : 한국교통안전공단 모빌리티플랫폼처 연구위원
*** 공저자 : 국립공주대학교 도시융합시스템공학과 교수

Eun Hye Yang* · Hee Chan Kang** · Woo-Young Ahn***

* Ph.D. Student, Dept. of Urban Systems Engineering, Kongju National University
** Research Fellow, Department of Mobility Platform, Korea Transportation Safety Authority
*** Professor, Dept. of Urban Systems Engineering, Kongju National University

† Corresponding author : Hee Chan Kang, k001kr@nate.com

Vol. 22 No.2(2023)
April, 2023
pp.182~194

pISSN 1738-0774
eISSN 2384-1729
<https://doi.org/10.12815/kits.2023.22.2.182>

Received 24 February 2023
Revised 15 March 2023
Accepted 26 March 2023

© 2023. The Korea Institute of
Intelligent Transport Systems. All
rights reserved.

요 약

교통영향평가는 도시·교통 사업 시행 시에 발생할 수 있는 교통문제를 사전에 해결하는 중요한 역할을 하는 제도이지만 사업의 특성을 고려한 작성 지침과 명확한 심의 기준이 없어 심의 결과의 일관성이 떨어지는 문제점이 있다. 본 연구에서는 교통영향평가에서 심의 의결된 교통개선대책을 분석하여 심의 결과의 핵심 키워드를 찾고 교통영향평가에서 주요하게 검토해야 하는 항목을 제시하고자 한다. 이를 위해 경기도에서 심의 의결된 도시개발사업의 교통개선대책을 중심으로 TF-IDF 및 N-gram 분석을 활용한 텍스트마이닝을 하였다. 도시개발사업에 주요하게 고려되는 사항은 도로, 교차로 등 교통 기반시설의 신설이 주요한 키워드였으며 다음으로 진출입구 위치 설정, 보행 동선 연결이 주요한 사항이었다. 향후 교통영향평가 지침과 심의 기준 마련 시 본 연구에서 제시한 주요 항목을 반영함으로써, 교통영향평가의 일관성과 객관성을 높이는 심의 운영에 이용할 수 있을 것으로 기대한다.

핵심어 : 교통영향평가, 텍스트마이닝, 단어빈도-역문서빈도, 엔그램 분석

ABSTRACT

Traffic impact assessment plays a crucial role in resolving traffic issues that may arise during the implementation of urban and transportation projects. However, reported results diverge, presumably because the items reviewed differ. In this study, we analyze traffic improvement measures approved for traffic impact assessment, identify key items, and present items that should be included in assessments. Specifically, TF-IDF and N-gram analysis and text mining were performed with focus on urban development projects approved in Gyeonggi Province. The results obtained show that keywords associated with newly established transportation infrastructure, such as roads and intersections, were essential assessment items, followed by the locations of entrances and exits and pedestrian connectivity. We recommend that considerations of the items presented in this study be incorporated into future traffic impact assessment guidelines and standards to improve the consistency and objectivity of the assessment process.

Key words : Traffic Impact Assessment, Text mining, TF-IDF, N-gram

I. 서론

1. 개요

교통영향평가는 신규 개발사업 시행 시에 선제적으로 교통개선대책을 마련하여 준공 이후 발생 가능한 교통문제를 최소화하는 도시교통정비촉진법 상의 심의제도로 Cho(2019)는 교통영향평가는 사업자에게 교통 시설개선을 실제 마련하도록 하고 있어 재정적, 사회 경제적으로 긍정적인 역할을 한다고 하였다.

교통영향평가 보고서는 교통영향평가 지침에 따라 작성하는데 철도와 도로 사업을 제외한 거의 모든 사업이 같은 기준으로 보고서를 작성하고 있다. 주택 관련 개발사업과 물류 관련 개발사업에 같은 기준을 적용하고 있어 사업별 특성을 반영하지 못하고 있다. 이 밖의 교통영향평가 제도의 문제점을 알아보기 위해 교통영향평가 업무와 관련된 공무원, 심의위원(교수, 연구원), 업계 엔지니어에게 설문 조사한 결과 모든 분야에서 높게 지적된 문제점은 교통영향평가 심의 결과의 일관성 부족이었다. 교통영향평가는 명확한 심의 기준 없이 심의위원회별 의견을 중심으로 운영되고 있기에 이와 같은 결과가 나온 것으로 유추할 수 있다. 교통영향평가 심의에 대한 기준이 부재한 상황에서 일관성 있고 합리적인 기준을 제시하여 제도를 제고하기 위해서 그동안 심의해온 결과들을 체계적으로 분석함으로써 교통영향평가를 점검하고 이를 바탕으로 전문가 회의, 연구 등을 통해 합리적인 방안을 모색할 필요가 있다.

본 연구는 교통개선대책 결과의 핵심 키워드(Keyword)를 분석하여 교통영향평가에서 주요하게 개선하고자 했던 항목을 찾아 향후 교통영향평가에서 주요하게 검토해야 하는 항목을 제시하고자 한다.

우선, 교통영향평가의 대상사업 중에서 많은 사람에게 영향을 미치는 공동주택 개발을 중심에 두고 있는 도시개발사업을 연구의 대상으로 선정하고, 10년간 경기도 교통영향평가 심의위원회에서 심의 의결한 교통개선대책을 활용하여 지침에 따른 심의항목을 기준으로 적용분류 후 핵심 교통개선대책 키워드를 분석하였다.

키워드분석은 R프로그램을 이용하여 텍스트마이닝(Text Mining) 분석을 수행하였으며 특히 TF-IDF(Term Frequency-Inverse Document Frequency, 단어빈도-역문서빈도) 분석과 N-gram(엔그램, 연이어 사용된 단어쌍 분석) 분석을 활용하여 심의항목과 교통개선대책의 핵심 키워드와 키워드 간의 의미구조를 분석하였다.

II. 설문조사 및 기존 문헌 고찰

1. 교통영향평가 제도 관련 설문 조사

교통영향평가 제도개선을 위하여 경기도 내 교통영향평가 관련 참여자 102명(공무원 31, 교수·연구원 23, 엔지니어 48)을 대상으로 2022.5.24.~6.7. 간 설문 조사를 하였으며 참여자의 68%가 10년 이상의 경력자였다.

교통영향평가 제도의 문제점을 알아보기 위해 14개 보기 중 심각한 문제 우선순위 3가지를 선택하도록 하고 3순위까지 단순 합계하여 순위를 선정하였다. 1~3순위의 결과는 소속을 구분하지 않고 전체를 합친 결과이지만 공무원, 교수·연구원, 엔지니어 분야 모두에서 약간의 순위의 차이가 있을 뿐 대부분 같은 결과가 나왔으며, 이 설문 조사를 통해서 기관별 교통영향평가 심의위원회의 일관성이 결여되어 있다는 문제에 공감하고 있음을 알 수 있다.

<Table 1> Results of survey on priority issues of traffic impact assessment

1 (16%)	Subjective evaluation based on personal experience not covered in guidelines
2 (15%)	Lack of consistency due to different procedures and evaluation criteria among Gyeonggi Province and 31 cities and counties
3 (13%)	Traditional improvement measures based on convention rather than on addressing current issues

2. 교통영향평가 관련 지침과 매뉴얼

1) 교통영향평가 지침

교통영향평가 지침(국토교통부고시 제2021-317호)의 교통영향평가 내용 항목(별표 1), 교통영향평가 분석표(별지 1), 심의 결과 통보서(별지 2~5)에서 교통영향평가 심의항목을 유추할 수 있다. 사업 시행에 따른 교통의 제반 문제점 도출 및 교통개선대책 수립 내용(별표 2)을 통해서는 심의항목별로 교통개선대책 마련 시에 필요한 개괄적인 기준이 제시되어 있다.

건축물의 경우는 모든 용도가 같은 항목을 적용하며 개발사업의 도로와 철도사업을 제외한 모든 개발사업이 같은 항목을 적용하고 있다. 건축물과 도로·철도 제외한 개발사업은 항목이 유사하며 가로 및 교차로, 대중교통, 자전거, 보행, 주차, 교통안전 및 기타로 분류하여 교통영향평가 항목을 적용하고 있다.

<Table 2> Items in the traffic impact assessment guidelines¹⁾

Construction of a building	Development of a city
A. Roadway and intersections (Site access) B. Location and design of site access driveways C. Public transport, Bicycle and pedestrian accommodation D. Parking plan E. Traffic safety and others	A. External roadway system B. Internal roadway system C. Public transport, Bicycle and pedestrian accommodation D. Parking plan E. Traffic safety and others
Construction of a road	Construction of railroads
A. Access points B. Surrounding roads connected to access points C. Public transport, Bicycle and pedestrian accommodation D. Parking Plan E. Traffic safety and others F. Measures to overcome regional segregation	A. Inside the station B. Outside the station C. Parking plan D. Traffic safety and other considerations E. Measures to overcome regional division

2) 교통영향분석·개선대책 심의표준 매뉴얼 (2012)

국토해양부(현재 국토교통부)에서 2012년에 각 지자체별로 운영되고 있는 심의의 일관성을 유지할 목적으로 교통영향분석·개선대책(현재 교통영향평가) 심의표준 매뉴얼에 대한 연구 진행하여 각 지자체에 배포 하였다. 표준 매뉴얼에는 가로소통과 교차로, 진출입 동선체계, 대중교통, 보행 및 자전거도로, 교통안전 및 교통정온화, 주차로 구분하여 시설물의 위치, 용량, 운영방식 등에 대한 가이드를 제시하고 있다.

1) 교통영향평가 지침 (별표 1) 교통영향평가 내용 항목을 정리한 것임

3. 선행연구 검토

1) 교통영향평가 심의 내용 분석 관련 연구

Bae et al.(2009)는 공동주택단지의 교통영향평가의 심의 내용 중 가장 빈도가 높은 항목의 사례를 조사해 교통영향평가 심의의결서 내용 유형을 정리하였고 교통영향평가 지침의 분류(가로 및 교차로, 보행, 진출입 동선, 교통안전, 주차)에 따라 유형별 빈도분석을 하였다. 주요 심의의 반영 사례를 살펴보고 공동주택 단지의 반영 내용을 정리하여 진출입 도로까지 서비스 수준 개선, 고원식 교차로 도입, 보행 접근성 제고, 안전하고 쾌적한 진출입 동선체계 수립이 필요하다는 계획적인 측면의 결과와 교통안전 시설과 주차면에 대한 적정 설치 기준이 필요하다는 제도적인 측면에 대한 시사점을 도출하였다.

Seong et al.(1999) 교통영향평가 심의 완료된 내용을 중심으로 개선항목을 분석하였는데 개선 건수는 가로 및 교차로 시설이 가장 많은 비율을 차지하고 그다음으로 보행 시설에 대한 개선, 교통안전 시설 순으로 교통개선대책을 수립한 것으로 분석하였고, 정부와 자치단체에 의한 사업의 교통개선 건수가 민간사업에 따른 교통개선 건수보다 2배 많은 것으로 분석하였다.

2) 텍스트마이닝을 활용한 정책과 제도에 대한 연구

Ann(2022)는 지방자치단체의 경관계획 경향을 파악하기 위해 지방자치단체의 경관계획 보고서를 수집하고 빈도분석, N-gram 분석을 활용한 텍스트마이닝 분석을 통해 키워드를 찾고 경관계획의 경향을 파악하였다. 더욱 효율적인 경관계획 수립을 위해 경관법 개정과 가이드라인 필요성 등 개선점을 도출하였다.

Cho(2020)는 환경영향평가 제도의 변화를 살펴보기 위해 31년간 택지개발 사업의 환경영향평가 협의 내용을 수집하고, 빈도분석, 키워드분석 중심의 텍스트마이닝을 활용하여 키워드를 추출하고 항목별로 구분하였다. 시기별로 평가항목의 변화 양상을 확인하고 언론 보도를 통한 사회적 인식의 변화가 제도에 미치는 영향의 정도, 사업 시행 주체 별로 차별성이 있다는 것을 확인하였다.

So et al.(2019)는 전 세계 스마트시티와 스마트시티 교통부문의 개념을 정립하기 위해 국내·외 보고서 등에서 정의 내용을 수집하고 텍스트마이닝을 활용하여 단어별 빈도분석 수행하였다. 빈도분석을 통해 추출된 단어를 범주로 구분하고 핵심개념들의 조합하여 스마트시티를 8가지 핵심 기능으로 정의하고 스마트시티 교통부문은 5가지 핵심 기능으로 정의하였다.

Jo et al.(2023)는 개인의 여가 통행데이터를 구축하기 위해 소셜 미디어 플랫폼의 정보를 수집하고 텍스트마이닝을 활용하여 토픽을 추출하고 공간분석을 위해 핫스팟 분석을 시행했다. 지역별로 주요 여가활동이 다르고 지역적 특성이 교통수단 선택행위에 있어 유의미한 영향을 미친다는 점을 확인하였다.

Ghomi and Hussein(2022)는 보행자 위반 행동에 대한 원인을 분석하고 미래 연구 방향 등을 제시하기 위해 21년간 보행자 위반 행동과 관련된 연구들에서 텍스트마이닝을 활용하여 토픽을 추출하였다. 보행자 위반 가능성을 높이는 주요요인으로 횡단 대기시간, 불법 주정차, 교통량, 버스정류장의 유무 등이 있고 기반시설 마련, 교육 프로그램개발 등 다양한 방법으로 보행자 위반 행동을 줄이는 방법을 제시하였다.

4. 연구의 차별성

교통영향평가 지침이나 심의표준 매뉴얼에서는 일반적인 구분하고 있는 가로, 교차로, 대중교통, 보행, 주차, 안전 등의 사항을 대상 사업마다 구분하지 않고, 일괄적으로 적용하고 있어 대상 사업별로 특성을 고려한 기준이라고 하기에는 미흡하며, 심의표준 매뉴얼은 현재는 활용하고 있지 않아 그 기능은 유명무실해졌다.

우리나라의 교통영향평가 심의 내용에 대한 논문은 거의 찾아볼 수 없으며 Bac et al.(2009)의 연구는 도시 계획과 건축계획과 관련하여 공동주택 단지계획 설계 시 반영해야 하는 교통영향평가 반영항목을 도출하기 위해 심의의견에 대해 빈도분석을 하였으며 실제 심의 결과를 활용한 것은 실제 적용성을 높일 수 있어 그 결과가 의미가 있다고 볼 수 있다. 다만 빈도분석만으로 분석되어 일부 중요하지 않은 내용이 반영되고 핵심 내용이 제외되어 주요 단어별 연계성 검토를 통해 실제 적용 가능한 결론이 도출이 필요하다.

기존 연구검토를 통해 최근 빅데이터 기법의 하나인 텍스트마이닝 기법이 다양한 분야에서 오랫동안 운영해 온 제도나 연구되어온 분야에서 토픽을 찾고 새롭게 도입되는 기술에서 개념을 정립하는 등 비정형 데이터의 텍스트마이닝을 통해서 새로운 정보를 만들어 정책이나 제도를 뒷받침하는데 활용하는 것을 살펴볼 수 있었다.

본 연구는 텍스트마이닝을 활용하여 그 동안 실제 운영해온 교통영향평가 심의 의결된 자료의 비정형 데이터의 단순 빈도분석이 아닌 핵심 키워드분석, 키워드 간 관계분석을 하였다. 이를 통해 현재 운영되고 있는 교통영향평가의 현황을 살펴보고 분야별로 주요하게 검토가 필요한 항목을 제시하는 점과 이를 활용한 제도의 발전 방향을 제시하여 기존 연구와 차별성이 있다.

Ⅲ. 연구방법론

1. 분석 대상

최근 10년간 경기도 교통영향평가 심의위원회에서 심의의결 완료된 도시개발사업 중 도시기반시설(도로, 공원, 학교 등)을 제외한 부지 중 80% 이상이 주거용도(공동주택, 단독주택, 준주거) 토지이용계획을 갖는 도시개발 지역 13곳(김포시 풍무양도·유현·결포·신곡, 화성시 반월, 양평군 다문, 수원시 망포·장안, 광주시 역세권·소화, 평택시 궁평·일광, 하남시 현안)을 선정하고 교통영향평가 심의의결서(개선대책 683건)의 교통개선대책의 내용을 수집하여 분석하였다.

대부분 보고서의 교통개선대책은 지침에서 정하고 있는 사업지구 외부(가로 및 교차로), 사업지구 내부(가로 및 교차로), 대중교통·자전거·보행, 주차, 교통안전 및 기타로 구분되어 있었다. 주차 분야에서 동일하게 ‘노외주차장 확보’사항을 제시하고 있고 교통안전 및 기타 분야에서도 표지판, 노면표시등 안전 시설물의 설치에 관한 내용이 담겨 있어 주차와 교통안전 및 기타 분야는 제외한 3분야만 분석 대상으로 선정하였다.

2. 분석 방법

본 연구에서 분석 대상으로 정한 교통영향평가 교통개선대책 데이터는 텍스트로 작성된 비정형 데이터이다. Williams and Gong(2014)는 일반적으로 자연어로 구성된 비정형 데이터에서 패턴 또는 관계를 추출하여 의미 있는 정보를 찾아내기 위해서 텍스트마이닝을 분석을 주로 활용한다고 하였다.

특히 주요 키워드를 추출하고 의미 관계를 분석하기 위해 텍스트마이닝 기법의 중 TF-IDF 분석과 N-gram 분석을 활용하고자 한다. TF-IDF 분석은 문서 내에서 어떤 단어가 중요한지를 나타내는 지표로서 키워드 추출을 할 수 있으며, N-gram 분석은 문서 내에서 단어나 언어 패턴을 분석하여 문서를 분류하는 데 사용된다. TF-IDF 분석으로 추출한 키워드를 N-gram으로 분석하여 해당 키워드와 연관성 있는 다른 단어나 언어 패턴을 찾아낼 수 있으며, N-gram으로 분석한 결과를 TF-IDF에 반영하여 단어의 중요도를 보다 정확하게 측정할

수도 있다.

TF-IDF 분석은 빈도 가중치 모델로 문서 내에서 각 단어가 얼마나 자주 등장하는지를 측정하여 그 단어와 다른 단어들과 비교하여 해당 단어가 상대적으로 얼마나 중요한지를 가중치로 나타낸다. 문서 내에서 자주 등장하는 단어일수록 TF 값이 크고, 전체 문서에서 드물게 등장하는 단어일수록 IDF 값이 크며, 이 두 값의 곱으로 계산된 TF-IDF 값이 커지는 것으로 TF-IDF 모델을 활용하면 특정 단어나 키워드가 문서 내에서 얼마나 중요한지를 파악할 수 있다.

$$TF-IDF = TF \times \log \frac{N}{DF}$$

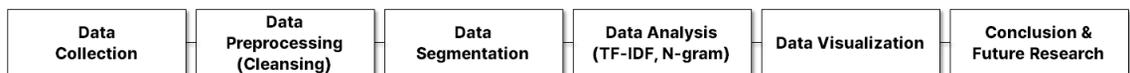
TF : Term Frequency, N : Total Number of Documents, DF : Document Frequency

N-gram(Park and Suh, 2015) 분석은 빈도분석을 통해 나온 키워드들이 원래의 자기 자리에 있던 순서를 완전히 무시하고 처리되므로 키워드 자체가 가지는 의미를 제대로 표현하지 못할 수 있어 이를 보완하기 위해 문서에서 인접한 키워드들을 쌍으로 묶어 표현하는 적용방법이다. TF-IDF 분석은 한 개의 단어만을 분석 대상으로 삼았지만, N-gram은 한 단어 이상의 단어 배열(sequence)을 분석 대상으로 삼아 단어를 N-gram으로 나눈 후 해당 단어들을 하나의 사건으로 보고 연속으로 나올 확률을 곱하여 각 단어 간의 연관성을 고려하여 한 문장의 확률을 계산하는 모델이다.

$$P(W) = \prod_{i=1}^T P(w_i | w_{k-1}, w_{k-2}, \dots, w_{k-n+1})$$

N-gram 분석에서 추출된 키워드들은 문서에서 중요한 단어들의 언어 패턴을 찾아 내용을 함축적으로 요약해주기 때문에 TF-IDF 분석과 연계하여 정확한 핵심 키워드를 선정하는 데 도움이 된다.

본 연구에서는 TF-IDF를 활용한 단순 빈도분석, TF-IDF 키워드분석을 하여 교통개선대책 분야별로 키워드의 중요도를 알아보고 N-gram 네트워크 그래프를 통하여 키워드들 사이에 어떤 관계를 형성하고 구조화되어 있는지 살펴보았다. 데이터 텍스트마이닝 분석을 위해 오픈 소스인 통계분석용 소프트웨어 R프로그램(버전 4.2.0)을 이용하여 비정형 텍스트 문서를 구조화하고, 분석 및 시각화하는 작업을 시행하였다.



<Fig. 1> Flowchart

IV. 분석 결과

전체 교통개선대책 683건을 사업지구 외부, 사업지구 내부, 대중교통·자전거·보행로 구분하여 빈도분석(TF)과 TF-IDF 가중치를 활용한 주요 키워드분석, N-gram 분석의 결과를 도출하고 워드 클라우드 그래프 및 네트워크 그래프로 시각화하였다.

도출된 키워드는 교통개선대책의 분류에 따라 분석하여 데이터 클리닝 단계를 거친 후 키워드의 배제 없이 분석하였다. 단, 데이터 클리닝 단계에서 각각의 개발사업에만 해당하는 키워드인 도로명칭, 교차로명칭,

<Table 5> Frequency analysis of keywords - Public transport, Bicycle, and Pedestrian

Keyword	TF
Installation	0.273101
Bicycle	0.075975
Sidewalk	0.069815
Bus	0.053388
Pedestrian	0.045175
Bus stop	0.039014
Crosswalk	0.032854



<Fig. 4> Visualization of frequency analysis of keywords - Public Transport, Bicycle, and Pedestrian

데이터 클리닝을 통하여 단순 빈도분석을 한 결과 ‘설치’, ‘교차로’, ‘차선’ 등의 키워드가 높은 빈도를 갖는 결과로 도출되었지만, 교통개선대책에 관한 내용을 파악하기 어렵고 분야별로 주요한 내용을 찾는 것은 더욱 어렵다. 교통영향평가를 다년간 수행해 본 연구자로서 교차로의 구조적, 운영적 개선의 필요성과 회전차로, 가감속 차로의 설치에 대한 교통개선대책에 관한 내용이라는 것을 유추할 수 있으나 이는 정확성이 떨어지는 것으로 TF-IDF 분석을 통해 분야별로 관련된 핵심 키워드를 추출할 필요성이 있다.

2. TF-IDF 분석

TF-IDF 분석은 반복적으로 나타나는 키워드들을 포함한 단순 빈도분석이 아닌 전체 문서에서 중요하게 생각되는 키워드를 추출하는 방법으로 사업지구 외부, 사업지구 내부, 대중교통·자전거·보행으로 구분하여 분야마다 중요 키워드를 분석하였다.

빈도분석에서 세 분야 모두에서 가장 높게 나온 키워드는 ‘설치’ 이다. 이는 도시개발사업의 특성상 도로 기반시설을 함께 시공하는 사업이기 때문에 나온 결과로 볼 수 있다. 대중교통·자전거·보행 분야를 제외한 사업지구 외부와 내부 모두에서 ‘교차로’가 높게 나왔다.

TF-IDF 분석에서 사업지구 외부에는 ‘차선’, ‘좌회전’, ‘확장’, ‘개선’, ‘폭원’이 중요 키워드로 분석되었고, 사업지구 내부에서는 ‘차선’, ‘설정’, ‘진출입구’, ‘불허구간’, ‘감속’이 중요 키워드로 분석되었다. 이는 사업지구 외부에서는 좌회전 차선 등을 설치하기 위한 도로의 확장에 관한 내용이 중요하게 다뤄지고, 사업지구 내부에서는 진출입구 주변의 불허구간 설정이나 진출입을 위한 감속차선 설치 등이 중요하게 다뤄진다는 것을 알 수 있다.

대중교통·자전거·보행 분야는 ‘버스’, ‘구축’, ‘정류장’, ‘횡단로’, ‘버스쉘터’ 에 대한 키워드가 중요 키워드로 버스정류장 및 횡단에 관한 내용이 중요하게 다뤄지는 것을 알 수 있다.

<Table 6> TF-IDF keywords

Rank	Outside the Development District		Inside the Development District		Public Transport, Bicycle, and Pedestrian	
	Keyword	TF-IDF	Keyword	TF-IDF	Keyword	TF-IDF
1	Lane	0.044128	Lane	0.031899	Bus	0.021647
2	Left turn	0.017907	Designation	0.022909	Construction	0.018047
3	Expansion	0.015349	Entrance/exit	0.019985	Bus stop	0.015819
4	Improvement	0.014709	No entry zone	0.019985	Crosswalk	0.013535
5	Widening	0.01343	Deceleration	0.011914	Bus shelter	0.011279

TF-IDF 키워드분석을 통해 ‘설치’ 등 단순 반복되는 단어들을 제외되었고, 분야별로 주요 교통개선대책의 키워드를 추출되어 키워드 간의 관계성이 높아진 것을 유추할 수 있었다. 키워드의 연계성을 자세히 살펴보기 위해서 N-gram 네트워크 분석을 추가로 시행하였다.

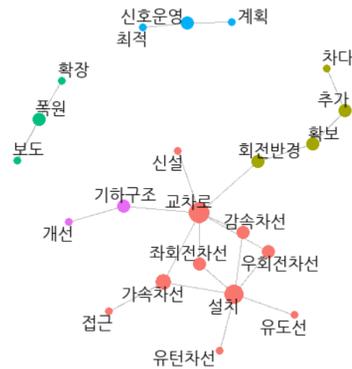
3. N-gram

N-gram 네트워크 그래프는 언어어 사용될 때 의미를 지니는 키워드들이 전반적으로 어떤 관계를 형성하는지 표현하고 있으며 키워드와 다른 키워드들이 얼마나 밀접하게 연결되어 있는지를 나타내는 연결 중심성을 보기 위해 키워드의 노드의 크기를 조정하였으며, 키워드 간의 그룹 성을 보기 위해 커뮤니티 변수를 추가하여 N-gram 네트워크 그래프를 표현하여 분석하였다.

사업지구 외부 분야의 N-gram 네트워크 그래프는 <Fig. 5>에 제시하였다. 빈도분석에서 빈도가 높게 나타난 키워드인 ‘교차로’, ‘설치’를 중심성을 가지며, TF-IDF 키워드분석에서 중요 키워드로 분석된 ‘차선’, ‘좌회전’, ‘확장’ 등이 연결성을 가지며 ‘기하구조’, ‘신설’, ‘좌회전차선’, ‘감속차선’, ‘설치’ 등이 연쇄적으로 나타나는 특징을 발견할 수 있었다.

<Table 7> N-gram analysis - Outside of the development district

Keyword	Keyword 2	Frequency
Intersection	Left turn lane	26
Left turn lane	Installation	24
Road widening	Expansion	21
Intersection	Road structure	19
Road structure	Improvement	19
Deceleration lane	Installation	13
Intersection	Deceleration lane	13
Acceleration lane	Installation	12



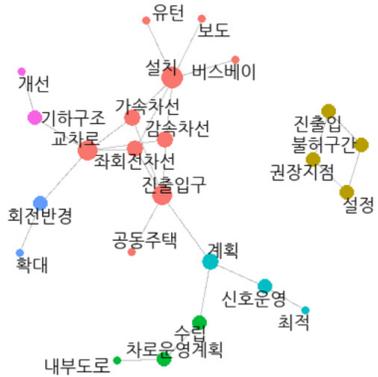
<Fig. 5> N-gram Graph - Outside the development district

사업지구 내부 분야의 N-gram 네트워크 그래프는 <Fig. 6>에 나타나듯이 빈도분석에서 빈도가 높게 나타난 키워드인 ‘설치’, ‘교차로’, ‘진출입구’가 중심성을 갖는다. TF-IDF 키워드분석에서 도출된 중요 키워드인 ‘진출입’, ‘불허구간’, ‘권장지점’, ‘설정’이 별도의 커뮤니티를 구성하고 있는 것을 알 수 있다.

사업지구 외부 분야와 내부 분야 모두 가로 및 교차로에 대한 교통개선대책을 수립하는 것으로 그 내용이 유사하다고 생각될 수 있었으나 N-gram 관계분석을 통해서 분야별 교통개선대책과 차별성이 있는 것을 알 수 있었다.

<Table 8> N-gram analysis - Inside of the development district

Keyword	Keyword 2	Frequency
Deceleration lane	Installation	29
Acceleration lane	Installation	24
Left turn lane	Installation	24
Intersection	Left turn lane	17
Intersection	Turning radius	16
No entry zone	Installation Instructions	16
Entrance/exit	No entry zone	16
Entrance	Deceleration lane	

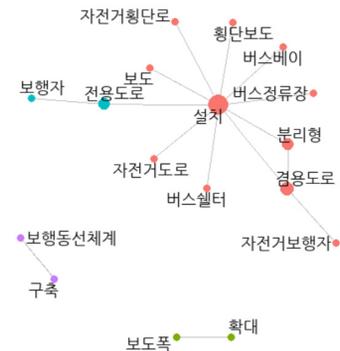


<Fig. 6> N-gram Graph - Inside the development district

대중교통·자전거·보행 분야의 N-gram 분석은 <Fig. 7>에서 보이듯이 빈도분석에서 빈도가 높게 나타난 키워드인 ‘설치’를 중심으로 ‘횡단보도’, ‘버스베이’, ‘버스정류장’, ‘자전거도로’ 등이 있으며, ‘보행 동선체계’, ‘구축’과 ‘보도폭’, ‘확대’는 별도로 구분된 것을 볼 수 있다.

<Table 9> N-gram analysis - Public Transport, Bicycle, and Pedestrian

Keyword	Keyword 2	Frequency
Sidewalk	Installation	17
Bus stop	Installation	13
Bicycle and Pedestrian	Shared path	12
Sidewalk width	Expansion	11
Crosswalk	Installation	11
Bus bay	Installation	8
Pedestrian	Exclusive lane	8
Exclusive lane	Installation	8



<Fig. 7> N-gram Graph - Public Transport, Bicycle, and Pedestrian

N-gram 네트워크 분석을 통해서 빈도분석에서 가장 높은 빈도로 나온 키워드인 “설치”, TF-IDF 키워드분석에서 도출된 분야별 주요 키워드 간의 관계성을 알아보고 네트워크가 몇 개의 주요 하위구조로 세분되었는지 알 수 있었다. 특히 빈도분석이나, TF-IDF 키워드분석은 단어 빈도수와 중요도를 고려하여 분석하므로 단어들 사이의 관련성을 고려하기 어려운 반면 N-gram 네트워크 분석은 단어 간의 연결성을 고려하여 분석하므로 단어들 사이의 의미적 관련성 및 복잡한 상호작용을 고려할 수 있어 더욱 정교한 분야별 주요 교통개선대책의 결과를 얻을 수 있었다.

V. 결 론

1987년 도시교통정비촉진법의 제정으로 시작된 교통영향평가 제도는 심의위원의 검토의견을 중심으로 운

영되고 있는 심의제도이며, 개별 사업의 승인관청 소속의 심의위원회에서 심의를 진행하고 있어 심의위원의 개별 역량마다 심의 결과가 다르게 나타나는 현상을 보이고 있다. 일부 지방 기초지자체에서는 심의위원을 구성하기 어려워 전문성이 다소 결여된 심의위원들이 심의하고 있는 경우가 발생한다. 이러한 문제는 경기도에서 진행한 설문 조사 결과에서도 문제점으로 지적되었다. 심의위원회를 운영하는 공무원, 심의를 받는 엔지니어(대행사), 심의하는 심의위원들이 유사한 문제점을 지적하고 있으며 이는 일관성 부족이라는 것으로 그 공통점을 찾을 수 있다. 1997년 처음 마련된 교통영향평가 지침과 현재 지침은 크게 다른 점이 없이 분석 내용이 추가되었을 뿐이고 개별 사업별로 다르게 적용되어야 하는 교통개선대책 마련 기준 등이 부재하여 모든 사업이 같은 기준으로 보고서를 작성하고 있다. 사업마다 보고서 작성 기준은 동일한데 심의위원회마다 심의 기준과 결과는 다른 아이러니한 현상이 발생하고 있는 것이다.

본 연구에서는 교통영향평가 제도 고도화를 위한 기초단계로 대상사업마다 주요하게 개선하고자 하는 항목을 찾아 교통영향평가 심의 시 주요하게 검토해야 하는 항목을 찾고자 하였다. 분석 대상은 경기도 교통영향평가 심의위원회에서 심의 의결된 주거용도 중심의 도시개발사업 자료를 분석하였다. 텍스트마이닝에서 핵심 키워드와 관계성을 분석하기 위한 TF-IDF와 N-gram 분석을 사용하여 분야별로 주요 키워드와 관계성을 알아보았다.

가로 및 교차로에 대한 부분은 사업지구 외부와 내부로 구분되는데 사업지구 외부의 교통개선대책은 주로 ‘교차로’와 ‘설치’가 연결 중심성을 가지며 좌회전차선, 감속차선, 가속차선, 우회전 차선, 유턴차선 등 차선의 설치 대한 내용이 주요한 항목이었으며 교차로의 신설, 교차로의 회전반경 확보하는 대책도 주요 내용으로 포함되어 있었다. 보도 폭원 확장 및 최적 신호 운영계획도 주요한 결과로 나타났다. 사업지구 외부는 기존 현황도로와의 접하는 곳이 많아서 추가로 용량 확보를 위한 교통대책이 주요한 것을 알 수 있으며, 사업지구 외의 주변 지역까지의 교통 영향을 고려하여 최적 신호 운영방안을 제시하는 것이 주요하게 분석되었다.

사업지구 내부의 교통개선대책의 일부 교차로 개선 부분에 대해서는 사업지구 외부의 분석 결과와 비슷한 것으로 보이나 N-gram 분석에서 볼 수 있듯이 ‘진출입구’와 ‘감속차로’와 ‘가속차로’의 설치가 연결성 있게 분석된 것은 그 설치 위치가 사업지구 외부와는 다르다는 것을 의미한다. 사업지구 내부의 감속 및 가속차로는 진출입구를 중심으로 진출입을 위한 완화차로의 역할을 하는 것이며 사업지구 외부의 감속 및 가속차로는 교차로를 중심으로 속도를 조절하여 접근하기 위한 완화차로의 역할을 하는 것이다. 또한 ‘진출입’, ‘불허구간’, ‘권장구간’이 주요 키워드로 분석되었으며 사업지구의 블록마다 차량 진출입구의 적정위치를 중요시하고 있고, 이를 관리하여 원활한 교통 흐름을 만들고자 하는 것을 알 수 있다.

대중교통·자전거·보행 분야에는 ‘보행동선체계 구축’ 등 보행자를 중심으로 한 교통개선대책이 눈에 띄며 ‘설치’를 중심으로 ‘횡단보도’, ‘버스베이’, ‘버스정류장’, ‘자전거도로’를 설치하는 것을 주요하게 고려하고 있다.

교통영향평가의 텍스트마이닝 분석을 통해 도시개발 사업의 특성을 검토한 결과 기존 지역에 건축물을 짓는 것과는 다른 새로운 도시를 개발하는 것을 고려하여 교통에 대한 기반시설을 새롭게 ‘신설’하는 내용을 주로 담은 결과가 도출된 것을 볼 수 있었다.

교통개선대책 수립 시 도로, 회전차로, 보도, 자전거도로, 회전교차로 등 다양한 도로 구성요소 마다의 설계 지침, 기준 등을 적용하여 설계되고 있다. 하지만 각 설계 지침에서 교통영향평가에 활용되는 항목은 일부분이다. 이에 본 연구에서 도출된 주요 항목에 대해 교통개선대책 설계의 법적 적정성 검토를 위한 체크리스트, 교통영향평가 심의 가이드라인을 만들 수 있을 것이다. 이를 통해서 도로 구조물이 적정한 규모로 설치할 수 있도록 하고 객관적이고 일관성 있는 심의 결과를 도출하여 사업자의 부담은 줄고 그 효과를 최대

로 발휘할 수 있을 것이다.

본 연구는 주거 용도를 중심으로 한 도시개발사업에 관한 심의 내용 중에서 주차와 교통 안전분야를 제외하고 분석을 진행하여 개발사업의 규모별 차이점과 도로와 교차로, 보행 등과 연결된 교통안전 시설물 설치에 대해 추가로 검토할 필요성이 있다. 또한, 향후 다른 사업과 건축물 개발에 따른 교통개선대책도 검토하여 교통영향평가 대상사업 별로 비교하여 검토한다면 교통영향평가에서 추구하는 공통된 교통개선대책과 대상 사업별 특징이 있는 교통개선대책이 구분되고 이에 따른 현황조사와 수요분석까지 고려할 수 있을 것이다.

35여 년이라는 오랫동안 실질적으로 우리 생활에서 느껴지는 교통개선대책을 마련하고 있는 교통영향평가의 중요성을 높이고 신뢰성을 회복하기 위해서 본 연구는 그 기초자료가 될 것이다. 향후 연구로 남겨 놓은 연구를 비롯하여 현재 심의되고 있는 교통개선대책의 현황을 살펴보고 지침 개정이나 심의 운영 시에 적용한다면 효과적일 것이다.

REFERENCES

- Ann, Y. B.(2022), *Analysis of the Characteristics of Local Government Landscape Plans: Focusing on Word Frequency Analysis Using Text-Mining Method*, Master's Thesis, SEOUL Women's University.
- Bae, W. K., Oh, S. U. and Cho, P. K.(2009), "Analysis and Direction for the Resolution of Traffic Impact Assessment about Multi-Family Housing", *Journal of Urban Design Institute of Korea*, vol. 10, no. 3, pp.149-164.
- Cho, D. S.(2019), "Urban Transportation & Traffic Impact Assessment", *Transportation Technology and Policy*, vol. 16, no. 2, pp.4-6.
- Cho, N. W.(2020), *An empirical study on the institutional change of environmental impact assessment: Focused on agreed terms*, Doctoral Dissertation, The Graduate School Yonsei University.
- Ghomi, H. and Hussein, M.(2022), "An integrated text mining, literature review, and meta-analysis approach to investigate pedestrian violation behaviours", *Accident Analysis & Prevention*, vol. 173, 106712, pp.1-18.
- Go, W. Y. and Hong, J. Y.(2022), "Framework for Key Performance Index of Use Cases in Smart Transportation Using Text Mining: Focusing on Daegu Metropolitan City", *Journal of Korean Society of Transportation*, vol. 40, no. 4, pp.449-461.
- Jo, J. S., Kwak, J. H., Ku, D. U., Yoo, Y. T. and Lee, S. J.(2023), "Estimation of Leisure Behavior Using Unstructured Data: Focusing on Seoul City Wall", *Journal of Korean Society of Transportation*, vol. 41, no. 1, pp.135-147.
- Ministry of Land, Infrastructure and Transport(2021.04.13), *Traffic Impact Assessment Guidelines(교통영향평가 지침)*, Ministry of Land, Infrastructure and Transport Notice No. 2021-317.
- Park, J. Y. and Suh, C. W.(2015), "Analysis of Changes in the Housing Market Using TF-IDF Weight Model", *Korea Real Estate Academy*, vol. 63, pp.46-58.
- Seong, N. J., Woo, M. J. and Park, B. H.(1999), "Effectiveness Questions of Transportation Impact

- Assessment in the Case of Chungbuk”, *Journal of Accident Analysis & Prevention*, vol. 13, no. 1, pp.169-176.
- So, J. H., Kim, T. H., Kim, M. J., Kang, J. W., Lee, H. and Choi, J. M.(2019), “A Study on the Concept of Smart City and Smart City Transport”, *Journal of Korean Society of Transportation*, vol. 37, no. 2, pp.79-91.
- Williams, T. P. and Gong, J.(2014), “Predicting Construction Cost Overruns Using Text Mining, Numerical Data and Ensemble Classifiers”, *Automation in Construction*, vol. 43, pp.23-29.