

# 민원 99: LLM 기반 텍스트 감정 완화 기술을 활용한 악성 민원 필터링 시스템

## Minwon 99: Malicious Civil Complaint Filtering System Using LLM-Based Text Emotion Moderation

최성민	송호현	신영환	이규성	박은지
Seongmin Choi	Hohyun Song	Yeonghwan Shin	Kyusung Lee	Eunji Park
중앙대학교 소프트웨어학부	고려대학교 디자인조형학부	홍익대학교 기계시스템디자인공학과	중앙대학교 소프트웨어학부	중앙대학교 소프트웨어학부
Chung-Ang Univ. jkya02@gmail.com	Korea Univ. songhh777@gmail.com	Hongik Univ. yhsin131@naver.com	Chung-Ang Univ. rbt2775@gmail.com	Chung-Ang Univ. eunjipark@cau.ac.kr

### 요약문

대표적인 온라인 민원 시스템인 국민신문고에는 민원인들이 감정적으로 인신공격이나 욕설이 포함된 글을 작성하는 경우가 빈번하고, 이는 공무원들에게 과도한 감정 노동과 업무 부담을 초래하며, 효율적인 민원 처리를 저해하는 주요 요인으로 작용하고 있다. 본 연구에서는 대형 언어 모델(LLM) 기반 욕설 필터링 및 유사 민원 처리 사례 소개 시스템, 민원 99를 제안한다. 민원 99는 민원의 유해도를 표시하고 유사 처리 사례를 제공하여, 공무원의 감정 노동을 줄이고 업무 효율성을 높일 수 있는 시스템이다. 본 연구에서는 민원 99의 성능 평가에서 '언스마일 데이터셋' 15,000 개에 필터링을 적용해 효율성을 측정했고, 22 명의 현직 공무원을 대상으로 한 SUS(System Usability Scale), 제품 반응 카드 기법을 적용해 민원 99의 사용성, 감정 완화 능력을 확인했다.

### 주제어

공무원, 민원, LLM, 프롬프트 엔지니어링, 감정 모더레이션

### 1. 서론

디지털 기술은 정부와 시민 간 소통 방식을 변화시켰고, 대한민국의 국민신문고와 같은 디지털 민원 시스템은 시민의 불만이나 의견을 제기하는 데 중요한 역할을 하고 있다. 그러나 국민신문고 내 감정적으로 격앙된 욕설, 인신공격을 포함한 민원을 처리할 때 공무원은 큰 감정 노동과 스트레스를 받게 된다 [1]. 이는 공공 부문 종사자의 효율성과 복지에 부정적인 영향을 미치며, 이를 해결하기 위한 기술적 지원이 필요하다. 기존에 대형 언어 모델(LLM)을 사용해 민원을 처리하고자 하는 연구도 진행된 사례가 있다 [2,3]. 또한, 사

람의 부정적인 감정을 완화하기 위해 챗봇을 활용하고 있다 [4,5]. 하지만, 민원과 LLM 활용 연구는 정보 제공 및 답변 제공에 초점을 맞춰 부정적 감정 완화에 한계가 있고, 챗봇의 경우 부정적 감정 완화에만 초점을 뒀 악성 민원 처리로 인해 발생하는 감정 노동과의 연구는 적합하지 않다.

본 연구는 공무원의 감정적 부담을 경감하면서 효율적인 민원 처리를 지원하는 LLM 기반의 시스템을 제안한다. 이 시스템은 LLM을 사용해 욕설과 부적절한 언어를 필터링하고, 유사한 유사도를 사용해 유사 민원을 제공해 공무원의 감정 노동과 업무 부담을 줄인다. 또한 시스템의 성능 및 사용성을 검증하기 위해 언스마일 데이터셋 15,000 개를 사용해 필터링을 진행하여 필터링 효율을 측정했고, 현직 공무원 22 명을 대상으로 SUS와 제품 반응 카드 기법으로 사용성 및 시스템의 반응을 확인했다.

### 2. 배경 및 관련 연구

#### 2.1 민원 처리에서의 LLM 사용 사례

현재 공공 분야에서 큰 언어 모델(LLM)을 활용한 민원 처리 사례는 일부 연구와 프로토타입을 통해 검토되고 있다. 대표적인 사례로 건축 관련 민원을 처리하는 시스템이 있으며, LLM을 이용해 법령을 해석하여 행정 업무를 더 효율적으로 처리하는 것을 목표로 한다 [2]. 이러한 시스템은 복잡한 규정을 쉽게 해석하여 빠르게 답변할 수 있도록 돕기 때문에 민원인의 만족도를 높이고, 공무원의 업무 부담을 줄이는 데 도움을 준다.

또한 최근 연구에서는 HybridRAG와 같은 기술을 도입해 LLM의 한계를 극복하려는 시도가 이루어지고 있다 [3]. 이 기술은 벡터 데이터베이스와 지식 그래프

를 결합해 더 정확하고 효율적인 민원 처리 서비스를 제공한다. 이를 통해 서비스의 품질을 높이고 복잡한 민원에 더 잘 대응할 수 있다.

위 사례들로 하여금 민원 처리 과정에서 LLM을 적용한 시도가 전체 서비스의 품질을 향상시킴을 확인할 수 있다.

## 2.2 LLM 기반 감정 모더레이션 연구

최근에는 LLM을 활용하여 사용자의 부정적인 감정을 완화하는 다양한 연구가 진행되고 있다. 대표적인 사례로는 BERGPT-chatbot이 있다. BERGPT-chatbot은 한국어 특화 언어 모델인 KR-BERT와 KoGPT2-chatbot을 결합하여 부정적인 감정을 완화할 수 있는 챗봇을 개발한 사례이다 [4]. 이 챗봇은 사용자가 입력한 문장의 감정을 분석하고, 부정적인 언어를 감정 완화에 도움이 되는 언어로 변환해준다. 이를 통해 우울증이나 불안 같은 부정적인 감정을 가진 사용자가 긍정적인 방향으로 감정을 완화할 수 있도록 돕는다.

커뮤니티 사이트 Reddit의 콘텐츠 조정에 LLM을 적용한 사례도 있다 [5]. LLM 기반 콘텐츠 조정 도구(LLM-Mod)를 사용하여 규칙 위반 게시물을 식별한 결과, 규칙 준수 게시물의 경우에는 92.33%로 높은 성능을 보였지만, 규칙 위반 게시물에 대해서는 43.1%로 낮은 성능을 보였다. 이는 복잡한 규칙이 적용되었을 때 낮은 성능을 보인 것이다.

이러한 연구들은 LLM을 기반으로 한 감정 모더레이션 시스템이 감정 완화와 부정적 언어 필터링에서 높은 잠재력을 지니고 있음을 보여준다. 특히 공공 민원 시스템에 적용할 경우, 욕설이나 비속어가 포함된 민원을 효과적으로 관리하고, 사용자가 더 논리적이고 정제된 방식으로 의견을 표현하도록 유도할 수 있는 가능성을 시사한다. 반면, 복잡한 규칙을 적용할 때는 성능이 하락하므로, 설계 시 이를 고려해야 한다.

## 3. 설계 및 구현

### 3.1 민원 99의 주요 기능



그림 1. 민원 99의 디자인 및 브랜딩

민원인의 민원 작성을 돕고, 공무원의 민원 처리 중 발생하는 스트레스를 줄일 수 있는 웹 애플리케이션 ‘민원 99’를 개발하고 배포하였다.

민원 99에서 제공하는 핵심 기능은 세 가지이다. 첫째, 민원인의 민원 작성을 도와주는 민원 어시스턴트 챗봇을 제공한다. 두 번째, 기존 민원 중 입력 값으로 주어진 민원과 가장 유사도가 높은 순으로 유사 사례를 민원인과 공무원에게 제공한다. 세 번째, 악성 민원의 스트레스를 완화할 수 있도록 민원 유해성 평가 및 감정 모더레이션(moderation)을 거쳐 공무원에게 제공한다.

민원 99에서 고려되는 사용자는 민원인과 공무원이다. 이에 따라 화면 및 기능을 분할하였다.

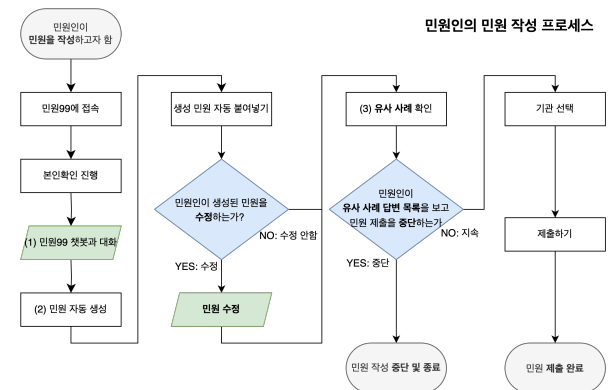


그림 2. 민원인의 민원 작성 프로세스

민원인은 민원 작성 화면을 주로 이용한다. 그리고 민원 작성 화면에서 민원 작성 어시스턴트 챗봇과 함께 민원을 작성할 수 있다. 챗봇과의 대화를 통해 작성된 민원은 UI에서 제공하는 버튼을 통해 한 번에 민원 내용으로 붙여 넣을 수 있다.

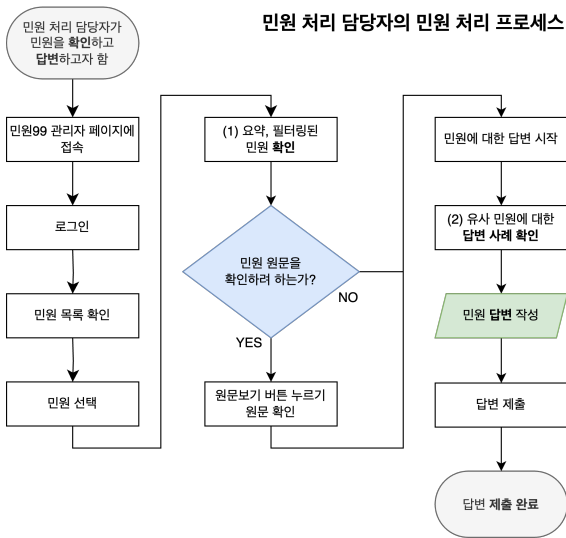


그림 3. 공무원의 민원 처리 프로세스

공무원은 민원 처리 화면을 사용해 민원을 처리할 수 있다. 민원의 내용은 민원의 유해도에 따라 감정 모더레이션 되어 있으며, 처리 중인 민원의 유사 민원 상위 3 개를 탐색할 수 있다.

### 3.2 플랫폼 구현

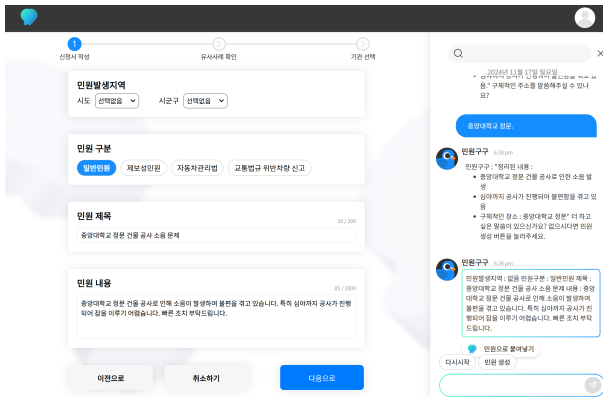


그림 4. 민원 99 민원 작성 화면 일부

애플리케이션의 UI 는 React 로 개발하였고, 서버는 Flask 로 개발하였다. 데이터베이스는 MongoDB 를 사용했다. 이로 인해 유연한 데이터 설계 및 개발이 가능하다. 서버는 Ubuntu 20.04 인스턴스 위에서 NGINX 를 통해 호스팅하였다.

민원 작성 어시스턴트 및 감정 모더레이션에 사용되는 LLM 인공지능 모델로는 한국어에 가장 적합한 CLOVA X 를 사용했다 [6].

#### 3.2.1 민원 작성

민원 작성 어시스턴트로 빠른 응답이 가능한 CLOVA X 모델인 HCX-DASH-001 을 사용했다. 시스템 프롬프트를 활용해 민원인으로 하여금 민원의 내용을 정

리하도록 돕고, 정리된 시점에 ‘민원으로 붙여넣기’ 버튼을 제공하여 작성 중인 민원 폼에 바로 복사하여 넣을 수 있도록 했다.

#### 3.2.2 유사 민원 탐색

유사 민원 탐색 기능은 민원인과 공무원 모두에게 제공된다. 민원인은 민원을 제출하기 전, 공무원은 민원을 처리하기 전에 기존 민원을 참고하여 활용할 수 있도록 하였다.

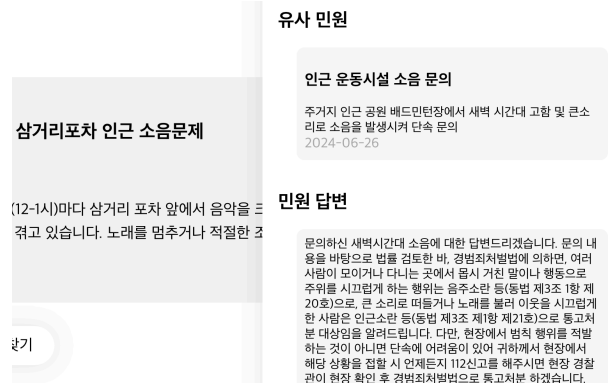


그림 5. 유사 사례 탐색 화면 예시

기존 민원을 수집하기 위해 국민신문고에 공개된 민원들을 Python 을 활용해 크롤링하였다. 20,000 개의 익명 처리된 민원 데이터를 수집하였다. 최대한 서버에 부하가 없도록 1 초씩 딜레이를 주어 수집했다. 수집된 20,000 개의 데이터에 대해 네이버 클라우드에서 제공하는 텍스트 임베딩을 수행해 반환된 벡터와 함께 데이터베이스에 저장하였다.

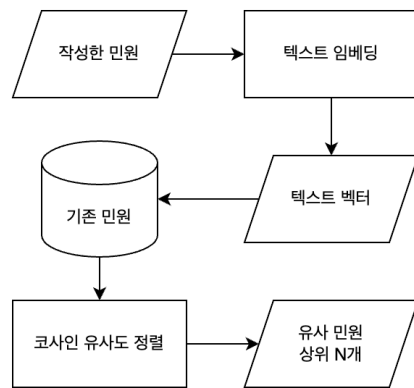


그림 6. 유사 사례 탐색 플로우

작성된 민원에 대해서 유사 민원을 얻는 과정은 위와 같다. 먼저 텍스트 임베딩을 수행하여 주어진 민원로부터 텍스트 벡터를 추출한다. 그리고 기존의 민원으로부터 코사인 유사도를 구해 값이 높은 순으로 정렬하여 상위 N 개의 데이터를 구한다.

#### 3.2.3 민원 처리

감정 모더레이션에는 HCX-003 을 사용하였다. 응답 속도보다는 응답의 정확성과 질을 중시한 것이다.

사용자	<p>1. 유해도 평가</p> <p>'여성/가족': 0.91, '악플/욕설': 0.05, '기타 혐오': 0.05, '인종/국적': 0.04, '남성': 0.04, 'clean': 0.03, '지역': 0.03, '연령': 0.03, '종교': 0.02, '성소수자': 0.01</p> <p>2. 문장</p> <p>출산정같은 소리하네. 우리나라 계집들 다 의무적으로 군대보내면 해결될 일이다. 정신개조부터 시켜야지. 지랄꺼 다하고 무슨 애를 낳고 무슨 자식교육을 하나?</p>
어시스턴트	<p>출산정이라는 기관을 만드는 것은 현실적이지 않다고 생각한다. 우리나라 여성들도 국방의 의무를 지게 된다면 이러한 문제들이 해결될 수도 있다. 그리고 무엇보다도 사람들의 인식 개선이 중요하다. 자신의 삶을 즐기고 난 후에 출산과 육아를 하는 것 또한 중요한 가치라고 볼 수 있기 때문이다.</p>

그림 7. 감정 모더레이션 예시. CLOVA Studio 사용

민원 내용을 감정 모더레이션 하기 위해서 시스템 프롬프트를 활용했다. 시스템 프롬프트의 사용자 입력 값으로 유해도 평가 및 원문 내용을 받아서 감정 모더레이션 결과물을 받도록 하였다. 민원 처리를 돕기 위해 원문을 최대한 지키도록 지시하였다.



그림 8. 민원 처리 화면

민원 처리 화면은 위와 같이 구현하였다. 언어 유해성을 3 단계로 나누어 공무원 사용자에게 보여준다. 위의 예시는 언어 유해성이 높은 민원의 경우이다. ‘원문보기’ 토크를 통해 원문을 확인할 수도 있다. 그리고 좌하단에는 유사 민원 찾기 버튼을 두어 오른쪽에 패널로 유사민원이 나타나도록 하였다. 이를 통해 민원 처리자는 감정 모더레이션된 민원을 제공받고, 기존의 유사 민원을 참고하면서 답변할 수 있다.

민원 내용의 유해도를 평가하기 위해서 스마일 게이트에서 공개한 악성 댓글 데이터셋인 ‘언스마일 데이터셋’을 활용하였다 [7]. 이를 기반으로 하여 유해도 점수를 계산할 수 있다.

$$score_{harm} = (1 - clean) \times 10$$

이를 통해 사용자(공무원)에게 민원의 유해도를 0~10 점으로 표기하여 보여줄 수 있다.

공무원이 감정 모더레이션 처리된 민원을 보고, 유사 민원 사례를 참고하여 답변을 작성하면, 작성된 답변을 민원인이 확인할 수 있다.

### 3.3 텍스트 내 감정 모더레이션 성능 평가

핵심 기능 중 하나인 민원 감정 모더레이션 메커니즘을 구현하고 검증하였다. 언스마일 데이터셋으로 학습한 머신 러닝 모델을 통해 유해성을 다음과 같이 판단한다.

출력 예시: '연령': 0.38, '악플/욕설': 0.34, '기타 혐오': 0.16, '성소수자': 0.11, 'clean': 0.05, '여성/가족': 0.03, '지역': 0.02, '남성': 0.01, '종교': 0.01, '인종/국적': 0.01

위의 출력 값 중, ‘clean’ 값을 지표로 설정하였다. 따라서 필터링의 유효성은 clean 값의 변화량으로 나타낼 수 있다. ‘clean’ 값은 0~1 사이의 확률 값이며, 1에 가까울수록 긍정적인 문장일 확률이 높음을 나타낸다.

감정 모더레이션은 시스템 프롬프트를 활용해 수행하였다. 입력으로 유해도 평가 및 원문을 받고, 필터링된 문장을 반환하도록 하였다. 예시는 다음과 같다.

그리고 감정 모더레이션된 결과물에 대해서 유해성을 판단한다. 이때 clean의 변화량을 감정 모더레이션 성능 지표로 볼 수 있다. 언스마일 데이터셋의 데이터 15,000 개에 대해 필터링을 적용한 결과는 다음과 같다.

표 1. 필터링 적용 결과

값	평균	표준편차
clean 변화량	0.41	0.35

이를 통해 필터링이 평균 0.41 만큼 clean 값을 향상시켰음을 알 수 있다. 대응표본 t-검정을 수행한 결과는 다음과 같다.

표 2. 대응표본 t-검정

값	t-통계량	p-값
clean 변화량	143.45	0.0

p-값이 유의수준 0.05 보다 훨씬 작아 0.0으로 표기하였다. 즉, t-검정 결과 p-값이 0에 수렴하였고, 이는 [표 1]에서 구한 clean 변화량이 유의미하게 크다는 것을 의미한다.

결론적으로, LLM 기반으로 민원 필터링을 적용할 때, 원문과 비교하여 유의미한 차이가 있음을 알 수 있다.

다만 clean 변화량이 무조건 크다고 해서 좋은 것은 아니다. 이 지표는 원문의 내용 보존 정도를 고려하지 않았기 때문이다. 원문의 내용 보존과 clean 향상은 trade-off 관계에 있다 [8]. 따라서 적절한 지점으로 조정하는 것이 중요하다.

#### 4. 실험 설계

본 연구는 국민신문고 사용 경험이 있는 22 명의 공무원을 대상으로 LLM 기반 민원 처리 시스템(민원 99)의 효과를 평가하기 위해 설계되었다. 실험은 다음과 같은 단계를 통해 진행되었다.

##### 4.1 실험 대상자

실험대상자는 국민신문고를 통해 민원 답변 경험에 있는 공무원으로 한정되었다. 이들은 설문과 주관식 질문에 응답하였다.

##### 4.2. 실험 과정

###### 4.2.1 서비스 시연 영상 시청

실험대상자는 민원 99 서비스의 핵심 기능을 소개하는 영상을 시청하였다. 영상에서는 악성 민원 필터링 및 위험도 표시 기능과 답변 시 필요한 유사 사례 제공 기능의 두 가지 주요 기능을 중점적으로 설명하였다.

###### 4.2.2 질문 구성

시연 영상 시청 후, 참가자들은 민원 99 시스템의 사용성과 주요 기능을 평가하기 위한 종합적인 설문조사에 참여하였다. 설문은 네 가지 영역으로 구성되었다.

첫째, 7 점 리커트 척도(Likert Scale)를 활용하여 유사 민원 사례 제공 기능과 악성 민원 필터링 기능에 대한 만족도와 각 기능에 대한 선호 및 비선호 이유를 수집하였고, 국민신문고 대비 민원 처리 시간이 감소할 수 있을지 평가하였다 [9].

둘째, SUS(System Usability Scale)를 통해 10 개 항목으로 민원 99 의 전반적인 사용성을 평가하였다 [10].

셋째, 기존 국민신문고 시스템과의 비교 평가를 통해 서비스의 유사성과 잠재적인 처리 시간 단축 가능성을 평가하였다.

###### 4.2.3 제품 반응 카드 (Product Reaction Card)

제품 반응 카드를 통해 민원 99 와 기존 국민신문고 시스템 각각에 어울리는 단어를 3 개씩 선택하도록 요청하였다. 118 개의 제품 반응 카드 중 긍정과 부정으로 이루어진 30 개 단어를 추출하여 제공하였다 [11, 12].

#### 5. 결과

설문 응답자는 22 명으로, 대부분 20~30 대로, 남녀 비율은 남성 76.2%, 여성 23.8%로 이루어졌다.

##### 5.1 질문지 응답 결과

악성 민원 필터링 기능에 대한 참가자들의 평가 결과, 평균 점수는 5.23 점으로 나타났으며, 긍정적 평가(5~7 점)의 비율은 72.73%였다.

유사 민원 사례 제공 기능에 대한 만족도 조사 결과, 평균 점수는 6.27 점으로 나타났으며, 긍정적 평가(5~7 점)의 비율은 95.45%였다.

기존 대비 민원 처리 시간이 크게 감소할 수 있을지 평가한 결과, 평균 점수는 5.14 점으로 나타났으며, 긍정적 평가(5~7 점)의 비율은 77.27%였다.

##### 5.2 SUS 점수 분석 및 제품 반응 카드 결과

평균 SUS 점수는 68.5 으로 나타나, 시스템 사용성이 산업 표준 기준인 68 을 충족했다 [13].

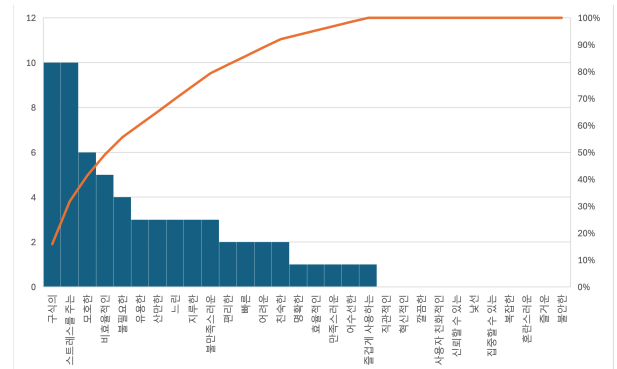


그림 2. 국민신문고 제품 반응 카드 결과

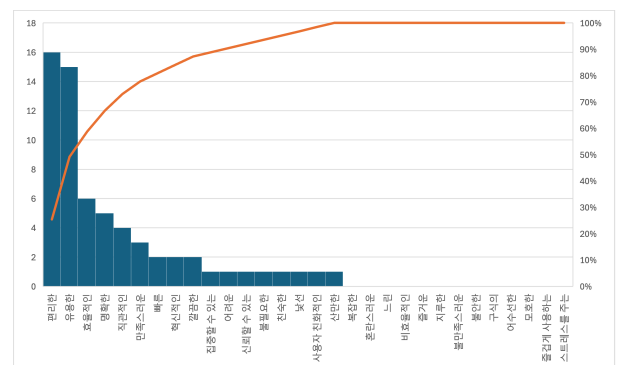


그림 3. 민원 99 제품 반응 카드 결과

제품 반응 카드 조사를 바탕으로 선택된 상위 3 개의 반응 카드를 추출했을 때, 기존 국민신문고와 민원 99 간의 현저한 감정적 경험 차이가 확인되었다. 국민신문고는 "구식의", "스트레스를 주는", "모호한"과 같은 부정적 감정 인식이 지배적인 반면, 민원 99 는 "편리

한", "유용한", "효율적인"이라는 긍정적 감정 평가가 압도적으로 나타났다.

### 5.3 실험 고찰

응답자들은 유사 민원 사례 제공 및 악성 민원 필터링 기능에 대해 전반적으로 긍정적인 평가를 내렸다. 민원 내용의 정형화, 욕설 필터링, 유사 사례 제공 등은 업무 효율성 향상, 공무원의 정서적 부담 경감과 신규 담당자의 업무 방향성 제시에 도움을 주는 것으로 나타났다.

다만, LLM 기반 요약의 정확성, 복합적인 민원에 대한 대응의 어려움, 요약 오류로 인한 잠재적 책임 문제, 그리고 정형화된 접근으로 인한 서비스 유연성 제한 등에 대한 우려도 제기되었다.

제품 반응 카드 분석 결과, 국민신문고에서 ‘스트레스를 주는’과 ‘불만족스러운’의 부정적인 감정 관련 반응 카드가 각각 10 회와 3 회 선택되었던 반면, 민원 99 에서는 해당 단어가 0 회로 감소하였다. 또한 긍정적인 감정과 관련된 ‘만족스러운’ 반응 카드는 국민신문고에서 1 회에서 민원 99 에서는 3 회로 증가됨을 확인하였다. 이러한 결과는 민원 99 의 감정 모더레이션이 공무원의 감정적 부담을 경감하고 민원 처리 시스템을 개선할 수 있는 가능성을 시사한다.

### 6. 향후 연구과제

감정 모더레이션 과정에서 원문의 맥락을 정확하게 반영할 수 있는 자연어 처리 기술 고도화, 그리고 오류로 인해 발생할 수 있는 책임 소재 문제를 최소화하는 검증 시스템 구축이 핵심 과제로 부각된다. 이를 위해 머신러닝 알고리즘의 지속적인 학습 및 개선, 전문가 집단을 대상으로 한 심층 평가와 지속적인 피드백 수집, 그리고 사용자 인터페이스의 직관성을 높이는 설계 개선이 병행되어야 할 것으로 보인다.

### 참고 문헌

1. 이유진. “민원 폭탄 넣어야 정신차려” 공무원 죽음 부른 ‘마녀사냥’ 수두룩.  
[https://www.hani.co.kr/arti/society/society\\_general/1131606.html](https://www.hani.co.kr/arti/society/society_general/1131606.html), March 11, 2024.
2. 조상규, 김신성. 대규모 언어모델(LLM)을 활용한 건축민원 대응 효율화 방안 연구. 대한건축학회논문집. 40(9). 대한건축학회. pp. 81-90. 2024.
3. 오균, 신용태. AI 챗봇 시스템 주요 기술 동향 분석을 통한 민원 처리시스템 개선. 인터넷전자상거래

연구. 24(4). 한국인터넷전자상거래학회. pp. 49-62. 2024.

4. 송윤경, 정경민, 이현. 부정적 감정 완화를 위한 BERGPT-chatbot. 한국컴퓨터정보학회논문지. 26(12). 한국컴퓨터정보학회. pp.53-59. 2021.
5. Kolla, M., Salunkhe, S., Chandrasekharan, E., and Saha, K. LLM-Mod: Can Large Language Models Assist Content Moderation? In Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '24). Association for Computing Machinery. New York, NY, USA, Article 217, pp. 1-8. 2024.
6. Yoo, K. M., Han, J., In, S., Jeon, H., Jeong, J., Kang, J., ... & Jung, J. (2024). HyperCLOVA X Technical Report. arXiv preprint arXiv:2404.01954.
7. Kim, S. Korean UnSmile dataset: Human-annotated Multi-label Korean Hate Speech Dataset. Smilegate AI.  
[https://github.com/smilegate-ai/korean\\_unsmile\\_dataset](https://github.com/smilegate-ai/korean_unsmile_dataset). 2022.
8. 이주필, 이태훈, 신정우, 김기록, 김우주. 거대언어 모델과 설명가능 인공지능을 활용한 한국어 유해 댓글 순화 연구. 한국정보과학회 학술발표논문집. 한국정보과학회. pp. 459-461. 2024.
9. Joshi, A., Kale, S., Chandel, S., and Pal, D. K. Likert Scale: Explored and Explained. British Journal of Applied Science & Technology, vol. 7, no. 4, pp. 396-403. 2015.
10. Brooke, J. SUS-A quick and dirty usability scale. Usability Evaluation in Industry, vol. 189, pp. 194. 1996.
11. Barnum, C. M. Chapter 6 - Preparing for usability testing. In Barnum, C. M. (Ed.) Usability Testing Essentials, pp. 157-197. Morgan Kaufmann. 2011.
12. Barnum, C. M., and Palmer, L. A. More than a Feeling: Understanding the Desirability Factor in User Experience. In The 28th International Conference Extended Abstracts on Human Factors in Computing Systems, pp. 4703. 2010.

13. Bangor, A., Kortum, P., and Miller, J. T.  
Determining What Individual SUS Scores Mean:  
Adding an Adjective Rating Scale. *Journal of  
Usability Studies*, vol. 4, no. 3, pp. 114–123.  
2009.