

# ChatGPT API 를 이용한 개인화된 뉴스 요약 시스템의 설계 및 구현

홍성호, 김의연, 정석현, 윤민혁, 최용훈, 송민경\*, 류시훈\*  
광운대학교, \*한국혁신센터 워싱턴

{gyogook608, dmldussla93, jayze3736, gural9368}@gmail.com, [yhchoi@kw.ac.kr](mailto:yhchoi@kw.ac.kr),  
\*{minsong, sryu}@kicdc.org

## Design and Implementation of a Personalized News Summarization System Using ChatGPT API

Seongho Hong, Euiyeon Kim, Sukhyun Jeong, Min-Hyeok Yun, Yong-Hoon Choi,  
Minkyong Song\*, Sihoon Ryu\*  
Kwangwoon Univ, \*Korea Innovation Center Washington DC

### 요약

본 논문에서는 사용자가 등록한 키워드를 기반으로 가장 연관성이 높은 뉴스 기사를 선별하여 제공하는 시스템을 제안한다. 이 시스템의 주요 기능은 다음과 같다. 첫째, 신뢰할 수 있는 뉴스 플랫폼에서 크롤링된 뉴스를 사용자가 등록한 키워드를 통해 필터링하여 메일로 제공함으로써, 사용자가 불필요한 정보를 걸러내고 중요한 뉴스만 효율적으로 소비할 수 있게 한다. 둘째, ChatGPT API 를 이용한 뉴스 요약 기능을 통해 사용자는 기사 전문을 읽지 않고도 주요 내용을 파악할 수 있어, 짧은 시간 안에 관심 있는 뉴스를 효과적으로 소비할 수 있다. 셋째, 시스템은 관리자 페이지를 통해 메일 발송 주기와 크롤링 주기를 설정할 수 있으며, 키워드를 자유롭게 추가, 수정, 제거할 수 있도록 설계되었다. 결론적으로, 본 시스템은 정보의 효율적인 소비를 가능하게 하여 사용자의 시간을 절약하고, 관련성 높은 콘텐츠에 쉽게 접근할 수 있도록 지원한다.

### I. 서론

현재 모바일과 인터넷 기술의 급속한 발전과 보급으로 인해 사람들이 뉴스 기사를 접하는 방식이 크게 변화하고 있다. 특히 최근에는 개인의 관심사를 반영한 뉴스 서비스를 이용하려는 수요가 증가하면서, 개인의 특성을 고려한 소셜 네트워크 기술을 활용한 맞춤형 뉴스 제공 서비스에 대한 연구가 활발히 진행되고 있다<sup>[1]</sup>. 전통적인 뉴스 추천 방식은 다수의 사용자 데이터를 기반으로 한 뉴스 검색 특징 데이터를 사용했으나, 최근에는 구독자의 선호도를 반영하여 유사한 성향을 가진 사람들이 검색한 뉴스 기사의 특징을 활용한 추천 시스템에 대한 관심이 높아지고 있다.

한편, 현재 다양한 뉴스레터 시스템이 존재하지만, 추천 시스템이 주로 개발된 웹사이트 내에서의 상호작용에 의존하거나<sup>[2][3]</sup> 사용자가 뉴스를 어떤 우선순위와 기준에 따라 제공받는지 명확히 알기 어렵다. 더불어, 포털 사이트들은 뉴스 기사를 추천할 때 요약되지 않은 원문 기사를 제공함으로써 사용자가 기사의 내용을 한눈에 파악하기 어렵다는 문제도 있다<sup>[4]</sup>. 이러한 문제점들은 사용자의 만족도를 저하시키고, 뉴스 소비를 불편하게 만드는 요인이 된다.

따라서 본 논문은 신뢰할 수 있는 뉴스 플랫폼에서 뉴스를 가져와 미리 정의된 키워드를 통해 연관도 높은 뉴스를 선별하고, 이를 ChatGPT API<sup>[5]</sup>를 이용해 요약하여 맞춤형 뉴스레터를 생성하여 이를 사용자들에게 전송하는 새로운 맞춤형 뉴스레터 시스템을 제안한다.

이러한 접근 방식은 정보의 효율적인 소비를 가능하게 하며, 사용자의 시간을 절약하고 더욱 관련성 높은 콘텐

츠에 접근할 수 있도록 한다. 본 논문에서는 이러한 시스템의 전체적인 흐름과 구현 방법을 자세히 설명하고, 간단하지만 효율적으로 연관된 뉴스를 선별하는 방법을 제시한다.

### II. 본론

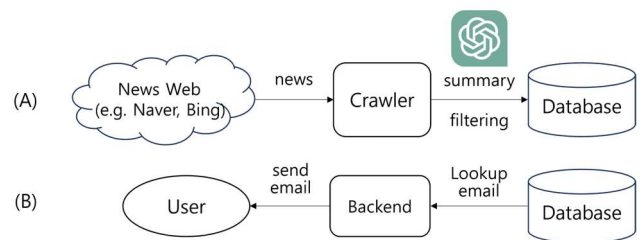


그림 1. 제안된 시스템의 전체 흐름도

본 시스템의 구조는 두 가지 주요 프로세스로 구성되어 있으며, 그림 1에 전체적인 구조가 나타나 있다.

그림 1(A)는 뉴스 수집 및 요약 저장 과정을 보여준다. 이 과정에서 시스템은 우선 Naver, Bing 과 같은 뉴스 플랫폼으로부터 실시간으로 뉴스를 크롤링한다.

이후, 필터링 과정이 두 단계로 진행된다. 첫 번째 단계에서는 TF-IDF 를 사용하여 기존 뉴스 기사와의 유사도를 측정하며, 유사도가 0.9 이상인 기사는 제외한다. 이는 재업로드되거나 소폭 수정된 뉴스를 걸러내기 위함이다.

두 번째 단계에서는 기사들을 (1)과 (2)에 따라 연관도 점수를 매기고, 가장 연관성이 높은 5 개의 뉴스 기사를 선별한다.

$$R_{ij} = f_{ij} \times w_j \quad (1)$$

(1)에서  $i$  는 뉴스 기사의 인덱스를 나타내며 ( $i = 1, 2, \dots, N$ ),  $j$  는 키워드 인덱스를 나타낸다 ( $j = 1, 2, \dots, K$ ). 여기서  $N$  은 뉴스의 총 개수,  $K$  는 키워드의 총 개수를 의미한다.

$f_{ij}$  는  $i$  번째 뉴스기사에  $j$  번째 키워드가 등장한 횟수를 나타내며,  $w_j$  는 각 키워드에 대해 사용자에게 의해 정의된 가중치를 의미한다. 따라서  $R_{ij}$  는  $i$  번째 뉴스 기사에 대한  $j$  번째 키워드의 연관도 점수를 나타낸다.

$$T_i = \frac{1}{K} \sum_{j=1}^K R_{ij} \quad (2)$$

(1)에서 구한  $R_{ij}$  를 모든 키워드에 대해 평균을 취한  $T_i$  는  $i$  번째 뉴스기사의 연관도 점수가 되며, 이는 (2)로 나타낼 수 있다.

예를 들어,  $i$  번째 뉴스 기사에  $j$  번째 키워드가 여러 번 등장하고, 그 키워드의 가중치가 높을수록  $R_{ij}$  의 값은 커지게 된다. 이를 통해 뉴스 기사의 연관성을 정량적으로 평가할 수 있으며, 최종적으로 연관도 점수가 가장 높은 상위 5 개의 뉴스를 선정하게 된다.

선정된 5 개의 뉴스는 ChatGPT API 를 사용하여 요약되며, 최종적으로 뉴스의 제목, 링크, 작성일 그리고 요약 내용이 데이터베이스에 저장된다.

#### Google Pixel Tips for Making the Most of Your Phone

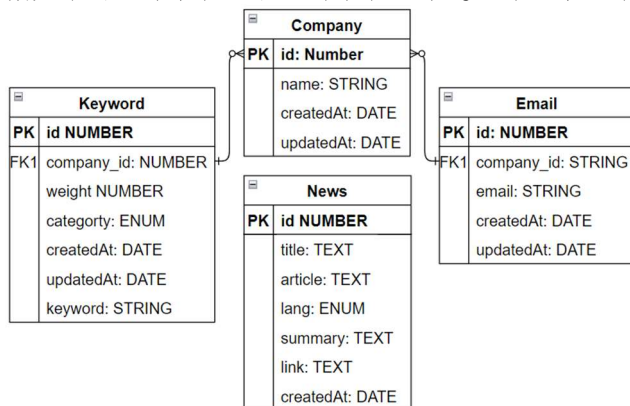
Google's Pixel 8 and 8 Pro introduce AI-powered features like changing photo expressions and object manipulation. Key features include Circle to Search, a temperature sensor, and macro photography. Other Pixels offer tools like Live Caption, Photo Unblur, and Astrophotography. Features like call screening and Quick Tap shortcuts enhance user experience across various Pixel models.

#### Your Google Photos library can now appear in OnePlus and Xiaomi's gallery apps

Google Photos is now integrated into the stock gallery apps of OnePlus, OPPO, Realme, and Xiaomi devices. This integration allows users to back up, view, and edit their Google Photos library directly from their device's gallery app. However, some limitations include the lack of album sync and potential duplicates. Despite these caveats, this integration simplifies photo management for many users.

#### 그림 2. 제안된 시스템에서 생성된 슛폼 뉴스레터 일부

그림 1(B)는 사용자에게 이메일을 보내는 과정을 보여준다. 데이터베이스에 저장된 뉴스 요약 정보는 Node.js 의 NodeMailer 패키지를 사용하여 설정된 주기마다 등록된 이메일로 발송된다. 그림 2 에서 볼 수 있듯이 뉴스레터에는 뉴스 제목과 요약 정보가 포함된다.



#### 그림 3. 제안된 시스템의 데이터베이스 개체-관계 다이어그램

이 시스템은 사용자가 직접 키워드, 키워드별 가중치, 메일 발송 주기 등을 설정할 수 있게 하여, 자동으로 사용자에게 가장 관련성 높은 뉴스를 정해진 주기마다 제공하는 자동화된 뉴스 크롤링 및 요약 전송 시스템을 구현한다. 해당 시스템에서 사용되는 데이터 요청/응답에 대해서는 Node.js 의 라이브러리인 express 를 사용하여 구현되었다. 시스템은 키워드, 이메일 발송 주기, 사용자 이메일 등을 설정할 수 있는 관리자 페이지를 포함하고 있다. 이 관리자 페이지는 Node.js 의 React 라이브러리를 사용하여 개발되었으며 제안하는 시스템에서 사용되는 모든 데이터베이스는 Mysql 을 기반으로 하며, 그림 3 에서 개체-관계 다이어그램을 보여준다.

#### III. 결론

본 논문은 웹 크롤링과 ChatGPT API 를 활용하여 사용자 맞춤형 뉴스 콘텐츠를 제공하는 시스템을 제안하고 구현했다. 이 시스템은 사용자가 선호하는 키워드를 기반으로 최신 뉴스 기사를 자동으로 수집하고, ChatGPT 를 사용하여 요약함으로써 정보의 효율적인 소비를 가능하게 한다.

이러한 접근 방식은 빅 데이터의 잠재력을 최대한 활용하여, 특히 뉴스 크롤링과 요약 분야의 기술적 발전을 실용적인 형태로 구현하였다는 점에 의의가 있다. 그러나 시스템의 현재 단점 중 하나는 크롤링 과정에서 많은 시간이 소요된다는 점이다. 이를 해결하기 위해 향후 연구에서는 크롤링 프로세스를 병렬적으로 진행하여 크롤링 시간을 대폭 줄임으로써 시스템의 전체적인 효율성을 향상할 예정이다.

#### ACKNOWLEDGMENT

이 논문은 2024 년도 정부(산업통상자원부)의 재원으로 한국산업기술진흥원의 지원을 받아 수행된 연구임(RS-2024-00406796, 2024 년 산업혁신훈인재성장지원사업)

#### 참 고 문 헌

- [1] Z. LU, Content-based collaborative filtering for news topic recommendation, In AAAI. pp. 217-223, 2015
- [2] Young-In Kim, Il Kyu Eom, & Ho-Sang Park (2016). Design and Implementation of a Responsive Website for Personalized News Article Recommendation and Newspaper Production System. 한국지식정보기술학회 논문지, 11(2), 205-214.
- [3] Lu, Z., Dou, Z., Lian, J., Xie, X., & Yang, Q. (2015, February). Content-based collaborative filtering for news topic recommendation. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 29, No. 1).
- [4] Park Seonae, Yoon Ju Hyun, Kim Ga young, Park Soyoung, Kim Yeon Su, & Jinho Ahn (2021-11-25). Customized Health News App Using Big Data Crawling. Proceedings of KIIT Conference, 제주.
- [5] OpenAI. (2023). ChatGPT(August 3 version) [Large language model]. (<https://chat.openai.com/chat>)