

<https://doi.org/10.7236/JIIBC.2023.23.1.203>
JIIBC 2023-1-29

실시간 스트리밍 서비스에서 신조어 정보 제공 시스템

System for Neologism Information Support in Real-Time Streaming Service

이승용*, 김능희**

Seungyong Lee*, Neunghoe Kim**

요약 최근 실시간 스트리밍 서비스가 MZ 세대의 각광을 받고 있으며 시장 규모 또한 지속적으로 커지고 있다. 사전에 촬영된 영상, 편집된 동영상은 단방향적 커뮤니케이션의 모습만 보여주지만 이와 달리 양방향 통신인 실시간 스트리밍 서비스는 사용자의 질문이나 요구에 즉각적으로 대응할 수 있다는 데에 강점이 있다. 또한 COVID-19 팬데믹으로 인해 대면 문화에서 비대면 문화로 전환됨에 따라 수업, 회의, 기타 여가 활동 등을 실시간 영상으로 소통하는 이용자가 폭발적으로 증가하였다. 하지만 실시간 스트리밍 서비스가 활성화되며 다양한 세대가 참여함에 따라 신조어를 포함하여 서로 간의 언어 차이에 의한 갈등 상황이 생기는 문제점이 발생하였다. 이를 해결할 방안으로 본 논문에서는 위키피디아 API를 활용하여 신조어의 뜻을 수집 후 출력하여 서로에게 확인시켜주는 방법을 통해 서로에게 의도를 파악하는 방법을 고안하였다.

Abstract Recently, real-time streaming services are gaining popularity among the MZ generation and the market size is continuously growing. Pre-recorded and edited videos only show one-way communication, but real-time streaming services have the advantage of responding immediately to questions and requests from users, as they enable two-way communication. With the transition from face-to-face culture to non-face-to-face culture due to the COVID-19 pandemic, the number of users communicating in real-time video for activities such as classes, meetings, and leisure has dramatically increased. However, as real-time streaming services become more active and diverse generations participate, there is a problem of conflicts arising from language differences, including the use of neologisms. To address this issue, this paper proposes a method of collecting the meaning of neologisms through the Wikipedia API and presenting them to each other, so that they can understand each other's intentions.

Key Words : real-time streaming service, neologism, wikipedia

*학생회원 군산대학교 소프트웨어학과

**정회원 군산대학교 소프트웨어학과(교신저자)

접수일자 2022년 11월 15일, 수정완료 2023년 1월 15일

게재확정일자 2023년 2월 3일

Received: 15 November, 2022 / Revised: 15 January, 2023 /

Accepted: 3 February, 2023

**Corresponding Author: nunghoi@kunsan.ac.kr

Department of Software Science & Engineering, Kunsan National University, Korea

I. 서 론

최근 실시간 스트리밍 서비스가 다양한 세대의 각광을 받으며 시장의 규모가 커지고 있다. 앱애니가 2021년 9월 발간한 소셜 미디어 앱의 진화 보고서에 따르면 2021년 상반기 전 세계적으로 실시간 스트리밍 분야의 상위 5개 앱은 지난 3년간 연평균 성장률 25%를 기록하며 사진 및 비디오 앱의 연평균 성장률 15%를 뛰어넘은 것으로 나타났다. 이와 더불어 COVID-19 팬데믹으로 인해 대면 문화에서 비대면 문화로 전환됨에 따라 수업, 회의, 기타 여가 활동 등에서 실시간 영상으로 소통하는 이용자가 폭발적으로 늘어나게 되었고 국내외 플랫폼들의 스트리밍 서비스 시장을 점유하기 위한 움직임이 활발하다^[1].

실시간 스트리밍 서비스는 실시간 채팅과 동영상 기능을 통해 다양한 세대와 서로 소통할 수 있다는 것이 강점이다. 사전에 촬영, 편집된 동영상은 단방향적 커뮤니케이션으로 정보와 모습만 보여준다. 이와 다르게 양방향 소통인 실시간 스트리밍 서비스는 사용자의 질문이나 요구에 즉각적으로 대응하여 관계를 형성하고 소속감을 쌓을 수 있다. 그리고, 실시간 스트리밍 진행자와 시청자 뿐만 아니라 함께 스트리밍 서비스를 시청하는 소비자들끼리 대화하며 서로의 의견을 공감하여 유대를 쌓을 수 있다는 점도 있다.

하지만 실시간 스트리밍 서비스가 활성화되며 다양한 세대가 참여함에 따라 신조어를 포함하여 서로 간의 언어 차이에 의한 갈등 상황이 생기는 문제점이 발생하였다.

이를 해결할 방안으로 위키피디아에서 제공하는 API(Application Programming Interface)를 활용하여 신조어의 뜻을 찾고 출력해서 서로에게 확인시켜주는

방법과 위키피디아에 없는 신조어의 경우 사전에 구축해 둔 데이터베이스에서 신조어 내용 탐색 후 서로에게 의도를 파악하게 하는 방법을 고안하였다.

II. 배경

1. 위키피디아

웹 기반의 다언어 백과사전으로서 중립적이고 검증이 가능한 자유형 콘텐츠 백과사전을 제공하는 목적을 가진 프로젝트로 전 세계 누구나 참여 가능하며 문서를 수정하고 발전시킬 수 있다. 위키피디아는 익명의 인터넷 사용자가 대가 없이 공동으로 문서를 작성하고 문서를 훼손하지 못하게 편집을 제한할 경우를 제외한 인터넷을 이용하는 누구나 위키피디아 문서를 편집할 수 있다. 종이로 출간된 백과사전과 달리 실시간 편집이 가능하며 끊임없이 새로 작성되고 갱신되어 일반화되고 중립적인 시각으로 변한다^[2]. 최근에는 지식을 활용하는 많은 인공지능 기술들이 위키피디아를 사용하고 있으며 웹에 존재하는 자원들을 연결 및 의미 부여를 통해 지식 표현과 공유하는 용도로 사용되고 있다^[3].

III. 실시간 스트리밍 서비스에서 신조어 정보

제공 시스템 설계 및 구현

1. 신조어 정보 제공 시스템 설계

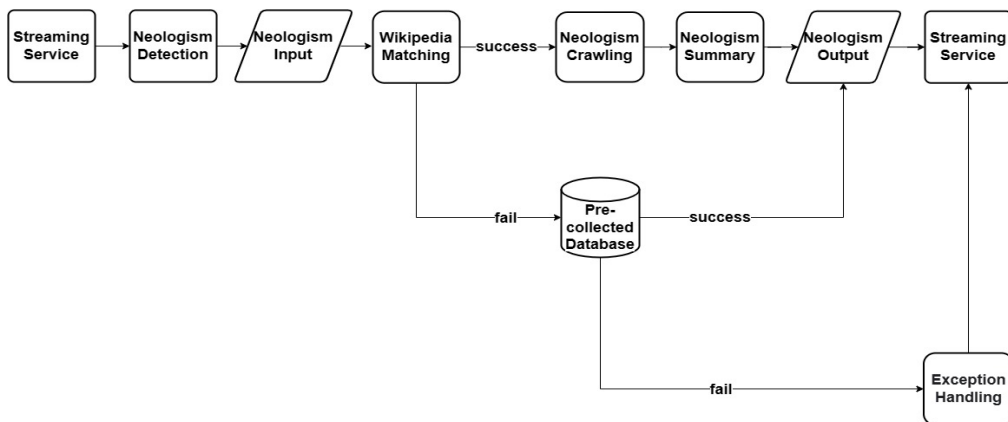


그림 1. 전체 프로세스
Fig. 1. Overall process

해당 연구는 wikipediaAPI의 기능 중 wiki.page를 활용하여 신조어의 정보를 사용자에게 보여주는 방법과 데이터베이스를 활용하여 신조어의 정보를 사용자에게 보여주는 방법을 토대로 연구를 진행하였다.

전체적인 연구 프로세스는 그림 1의 전체 전개도를 기반으로 한다. 실시간 스트리밍 서비스에서 이용자가 신조어를 확인 후 신조어에 관해 검색하였을 경우 wikipediaAPI를 통해 해당 신조어의 문서를 확인한다. 신조어 내용이 존재 시 해당 내용을 수집 후 wikipediaAPI에서 제공하는 summary 기능을 통해 신조어의 요약물 실행한다. 요약된 신조어 내용을 스트리밍 서비스에 출력하는 절차를 거친다. wikipediaAPI를 통해 검색된 신조어 내용이 존재하지 않을 시 사전에 신조어 내용을 수집한 데이터베이스를 통해 해당 내용을 찾고 데이터베이스 내에 존재 시 해당 신조어 내용을 스트리밍 서비스에 출력한다. 만약 위의 두 경우에 모두 충족되지 않을 시 스트리밍 서비스에 예외 처리를 출력한다.

2. 신조어 정보 제공 시스템 구현

신조어 정보 제공 시스템 설계에 따라 시스템을 구현하였으며 시스템 구현의 환경 정보와 실행화면은 다음과 같이 구성하였다.

표 1. 컴퓨터 환경

Table 1. Computer environment

Computer	Spedification
OS	Windows 10 Education ver.20H2
CPU	Intel(R) Core(TM) i3-10105 CPU @ 3.70GHz
RAM	16.0GB
Discord	API v10
Wiki API	Semantic Versioning 2.0.0

표 2. 실험 환경

Table 2. Experimental environment

실험 환경	사전 데이터	사용 라이브러리
Discord	.csv	discord.py wikipedia.py beautifulsoup

컴퓨터 환경은 표 1과 같은 환경에서 진행하였고 실험 환경은 표 2와 같은 환경에서 진행하였다. 실험 환경에 대해 세부적으로 설명하자면 실험에 사용된 플랫폼은 Discord이다. 사전 신조어 수집 데이터는 .csv 파일이고 사용 라이브러리는 Discord 환경에 제안 시스템을 적용

하기 위해 Bot 사용을 위한 discord.py를 이용하였다. wikipedia.py는 위키피디아 문서에 접근 및 내용 요약 을 위해 사용하였다. beautifulsoup는 사전 데이터를 수집하는 용도로 이용하였다.

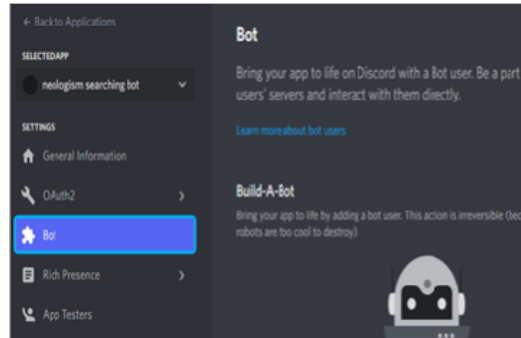


그림 2. DiscordBot 생성 화면

Fig. 2. Screen for creating a DiscordBot

그림 2는 DiscordBot으로 실시간 방송플랫폼 트위치나 유튜브 등에서 사용되는 NightBot 기능과 최대한 유사한 환경을 조성하기 위해 Discord에서 제공하는 discord.py를 활용하여 DiscordBot을 이용하였다.

표 3. 데이터베이스 예시

Table 3. Database example

단어	뜻
가즈아	도박이나 투자에서 긍정적인 기대를 표현하는 감탄사로, 소망을 말할 때 사용되는 단어로 의미가 확장되었다.
금수저	부모의 연 소득과 가정 환경 등에 부족함이 없는 가정에서 태어나 호의호식하는 사람을 뜻한다.
플렉스	속어로 '과시하다', '사치를 부리다'라는 뜻인 'flex'의 차용어이다.
어라 랍스타	어라? 혹은 어라라? 추임새 뒤에 랍스타를 붙여서 어라랍스타??로 사용된다.

표 3은 신조어 데이터베이스로 표 2의 사전 데이터를 데이터베이스 시스템으로 구축한 예시를 보여주고 있다.

구현 완료 후 wikipediaAPI를 실행하였고 시스템 설계에 따라 신조어를 검색하였을 경우 위키피디아에 존재하여 보이는 실행화면과 위키피디아에 존재하지 않아 데이터베이스로부터 가져와 보이는 실행화면은 다음과 같다.

그림 3의 실행화면 부분은 실시간 방송 이용자가 '뉴트로'라는 신조어 입력 시 봇이 이를 감지하고 해당 신조어 내용을 출력해 주는 실행화면을 보여주고 있다. 신조어 입력 시 '뉴트로'라는 신조어가 위키피디아 문서로 존

재할 경우를 나타내고 있다. 그림 3의 terminal 화면은 내부에서 실행 중인 모습을 나타낸 것으로 입력받은 신조어가 위키피디아 문서에 존재 여부를 확인하는 부분이다. Page - Exists는 위키피디아 문서 내의 존재 여부를 보이는 기능으로 위키피디아 문서가 존재하여 True로 나타내었다. 그림 3의 terminal 화면의 내부 모습은 '뉴트로'라는 단어 입력을 받고 해당 신조어 문서의 존재 여부를 wikipediaAPI를 통해 확인받은 뒤 문서 존재 시 summary로 요약된 내용을 출력한다.

실행 화면	<p>실시간 방송 이용자: 뉴트로</p> <p>디스코드 신조어 봇 : 뉴트로(newtro) 또는 신복고(新復古)는 2010년대 후반부터 복고풍이 새롭게 유행하는 현상을 뜻하는 대한민국의 신조어이다.</p> <p>#일번에 메시지 보내기</p>
terminal 화면	<p>Page - Exists: True</p> <p>Page - Summary: 뉴트로(newtro) 또는 신복고(新復古)는 2010년대 후반부터 복고풍이 새롭게 유행하는 현상을 뜻하는 대한민국의 신조어이다.</p>

그림 3. 신조어가 위키피디아에 있는 경우
Fig. 3. When the neologism is on Wikipedia

그림 4의 실행화면 부분은 실시간 방송 이용자가 '오하운'이라는 신조어 입력 시 봇이 이를 감지하고 해당 신조어 내용을 출력해 주는 실행화면을 보여주고 있다. 신조어 입력 시 '오하운'이라는 신조어가 위키피디아 문서 내에 존재하지 않는 모습을 나타내고 있다. 그림 4의 terminal 화면 또한 그림 3의 terminal 화면과 같이 내부에서 실행 중인 모습을 나타낸 것으로 wikipediaAPI와 비교한 뒤 위키피디아 문서 내에 '오하운'이라는 신조어 문서가 존재하지 않을 경우로 Page - Exists가 False로 된다. 입력받은 신조어가 위키피디아 문서 내에 없는 것을 확인 후 데이터베이스로 넘어가 사전에 수집한 내용을 출력한다.

실행 화면	<p>실시간 방송 이용자: 오하운</p> <p>디스코드 신조어 봇 : 오하운: 오늘 하루 운동</p> <p>#일번에 메시지 보내기</p>
terminal 화면	<p>Page - Exists: False</p> <p>Page - Summary: False</p> <p>오하운: 오늘 하루 운동</p>

그림 4. 신조어가 위키피디아에 없는 경우
Fig. 4. When the neologism is not on Wikipedia

IV. 실험 및 결과

인터넷 커뮤니티에서 100명을 대상으로 본 논문에서 제안한 방식의 유용성에 대해 설문조사를 진행하였다.

설문조사는 신조어로 인한 문제점 2개와 본 논문에서 제안하는 방식을 사용했을 시에 대한 결과 2개로 총 4개의 질문을 사용하여 수행하였다. 설문조사 결과는 그림 5와 같다.

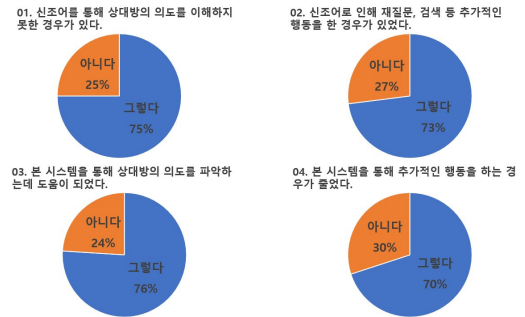


그림 5. 설문조사 결과
Fig. 5. Survey results

신조어로 인해 상대방의 의도를 이해하지 못한 경우에 대해 전체 응답자 중 75%가 '그렇다'라고 응답하였다. 신조어로 인해 의미 전달이 원활하지 않음을 확인해 볼 수가 있다.

신조어로 인해 재질문, 검색 등 추가적인 행동을 한 경우에 대해서는 전체 응답자 중 73%가 '그렇다'라고 응답하였다. 신조어로 인해 재질문, 검색 등 추가적인 수고가 발생하고 있다는 것을 알 수가 있다.

본 시스템을 통해 상대방의 의도를 파악하는데 도움이 되었다는 응답자가 76%였다. 본 논문에서 제안하는 방식이 의미 전달을 원활하게 하는데 도움이 된다는 것을 확인해 볼 수가 있다.

본 시스템을 통해 추가적인 행동을 하는 경우가 줄었다는 응답자가 70%였다. 본 논문에서 제안하는 방식이 재질문, 검색 등 추가적인 수고를 감소시킨다는 것을 알 수가 있다.

V. 결론

본 논문에서는 실시간 스트리밍 서비스 이용자를 위해 위키피디아를 이용하여 신조어 정보 제공 시스템을 실시

간 스트리밍 서비스 환경에 구현하였다. 본 제안 방법은 이용자에게 실시간 스트리밍 서비스 내에서 빠르게 신조어의 내용에 접근이 가능해지고 실시간 스트리밍 서비스 이용 중 신조어에 대한 내용을 제공한다. 이를 통해 의미 전달을 원활하게 하는데 도움이 되고 재질문, 검색 등 추가적인 수고를 감소시킨다는 것을 확인하였다.

향후 연구로는 신조어 탐색 시에 사용하는 위키피디아 API의 의존도를 줄이고 자체적인 데이터베이스를 기반으로 강화된 신조어 수집 및 요약을 수행할 예정이다.

References

- [1] Chang-Sup Nam, A Study on the Quality Monitoring and Prediction of OTT Traffic in ISP. Journal of Korea Institute of Information, Electronics, and Communication Technology, Vol. 14(2), pp. 115-121, Apr 2021. DOI: <http://dx.doi.org/10.17661/jkiect.2021.14.2.115>
- [2] Wikipedia, URL : <https://ko.wikipedia.org/>
- [3] JaeHun Kim, Myungjin Lee, Knowledge Extraction Methodology and Framework from Wikipedia Articles for Construction of Knowledge-Base. Journal of Intelligence and Information Systems, Vol. 25(1), pp. 43-61, Mar 2019. DOI: <http://dx.doi.org/10.13088/jiis.2019.25.1.043>
- [4] Yunji Moon, The Effect of Social Affordances in Social Live Streaming Service. Journal of Information Technology Applications & Management, Vol. 27(6), pp. 31-51, Dec 2020.
- [5] Sang-Min Park, Chul-Won Na, Min-Seong Choi, Da-Hee Lee, Byung-Won On, KNU Korean Sentiment Lexicon - Bi-LSTM-based Method fo Building a Korean Sentiment Lexicon -. Journal of Intelligence and Information Systems, Vol. 24(4), pp. 219-240, Dec 2018. DOI: <http://dx.doi.org/10.13088/jiis.2018.24.4.219>
- [6] Chang, Jae-Young, Topical Clustering Techniques of Twitter Documents Using Korean Wikipedia. The Journal of the Institute of Internet, Broadcasting and Communication, Vol. 14(5), pp. 189-196, Oct 2014. DOI: <http://dx.doi.org/10.7236/JIIBC.2014.14.5.189>
- [7] Kang, Hyeonah, Song, Hyunju, Constructing Sentiment Lexicon for Subject-Specific Sentiment Analysis. Korean Linguistics, Vol. 93, pp. 83-110, Nov 2021. DOI: <http://dx.doi.org/10.20405/kl.2021.11.93.83>
- [8] Yoon Sung-Hee, A Study on Word Semantic Categories for Natural language Question Type Classification and Answer Extraction. Journal of the Korea Academia-Industrial cooperation Society, Vol. 5(6), pp. 539-545, Dec 2004.
- [9] Cho, Byung-Ho, A Social Learning as Study Platform

using Social Media. Journal of Korea Institute of Information, Electronics, and Communication Technology, Vol. 5(4), pp. 180-185, Dec 2012.

저 자 소 개

이 승 용(학생회원)



- 2017년 3월 ~ 현재 : 군산대학교 소프트웨어학과 학사과정
- 주관심분야 : 자연어처리, 소프트웨어공학

김 능 회(정회원)



- 2008년 8월 : 고려대학교 컴퓨터학과 (이학석사)
- 2017년 8월 : 고려대학교 컴퓨터·전파통신공학과 (공학박사)
- 2021년 3월 ~ 현재 : 군산대학교 소프트웨어학과 조교수
- 주관심분야 : 요구공학, 소프트웨어공학