

정보통신용어표준화 추진 현황과 인공지능 활용에 대한 고찰

조은주, 김학훈, 차순일

한국정보통신기술협회(TTA)

jej@tta.or.kr, hakhoonk@tta.or.kr, sicha@tta.or.kr

Status of ICT Terminology Standardization and Implications for Applying AI

Jo Eunju, Kim Hakhoon, Cha Soonil

Telecommunication Technology Association

요약

한국정보통신기술협회(TTA)는 1993년부터 정보통신기술 발전에 따라 생성되는 정보통신용어를 해설하고 표준화하여, 전문가뿐만 아니라 비전문가들도 올바르게 사용할 수 있도록 정보통신용어사전(terms.tta.or.kr)을 서비스하고 있다. 본 논문에서는 정보통신용어표준화 추진 절차, 표준용어 선정과 작성 현황을 알아보고, 일반인도 알기 쉬운 표준용어를 인공지능을 활용하여 어떻게 잘 선정하고 만들 수 있을지 현 시점에서의 이점, 한계 등을 살펴보았다.

I. 서론

한국정보통신기술협회(TTA)는 1993년부터 정보통신기술 발전에 따라 생성되는 정보통신용어를 해설하고 표준화하여, 전문가뿐만 아니라 비전문가들도 올바르게 사용할 수 있도록 정보통신용어사전(terms.tta.or.kr)을 서비스하고 있다.

용어사전은 사용 범위에 따라 크게 표준용어, 시사용어, TTA 표준용어로 구분하여 서비스한다. 표준용어는 정보통신용어표준화(이하 '용어표준화') 절차에 따라 채택된 정보통신용어를 말한다. 전문가뿐만 아니라 일반인이 일상생활에서 ICT 관련 제품이나 서비스와 관련하여 사용하는 용어다. 시사용어는 사회 전반에 걸쳐 대다수가 사용하는 ICT 용어로 주로 신문사와 같은 언론에서 사용한다. 실생활과 관련되는 서비스, 법/제도, 트렌디한 유행어도 포함되어 있다. 마지막으로 TTA 표준은 정보통신단체표준에서 정의하여 사용하는 전문 기술 용어다. 전문가가 기술 전달 및 구현 등을 위해 사용한다.

본 논문에서는 표준용어가 어떻게 만들어지는지 용어표준화 추진 절차, 작성 원칙 등을 알아보고 인공지능 시대에 표준용어가 어떤 변화를 가질 수 있을지 향후 과제 등을 고찰해 보기로 한다.

II. 본론

1) 용어표준화 추진 절차

용어표준화의 대상이 되는 용어는 ICT 융합, 방송, 이동통신, AI 등 ICT 전 분야에서 사용하는 용어다. 새로운 기술에 대한 적합한 용어 제시, 이미 사용되고 있는 용어 정리, 혼용 용어 표준화, 외래 표현의 우리말 표기 제시 등을 목적으로 한다. 협회는 '23년 12월 기준 총 25,363건의 표준용어를 보유·서비스하고 있으며 매년 약 120건을 발굴, 제정 및 개정 중이다.



[그림 1] 용어표준화 추진 절차

용어표준화는 발굴 → 집필 → 심의·채택 → 서비스의 과정으로 추진한다. 용어 제안은 국민 누구나 할 수 있으며 제안된 용어는 필요성, 중요성 등을 검토하여 수록 여부를 결정한다. 표준용어 초안은 ① 집필 전문가 초안 작성, ② 검토 전문가 초안 검증 ③ 국어전문가 국어 감수의 과정으로 집필하고, 약 20명의 ICT 기술 분야별 전문가, 국어전문가가 활동하는 정보통신용어표준화위원회에서 최종 심의 후 서비스한다. 정보통신용어사전 웹/앱 그리고 네이버, 다음과 같은 포털뿐 아니라 기관, 기업에도 제공한다. '23년 국가지식정보'로 지정되어 '디지털집현전'(k-knowled ge. kr)에서도 서비스중이며, 메타버스와 같은 기술분야에서 사용하는 용어들을 모아 제작한 기술용어집을 일반인 등에 배포하고 있다. 또한 모든 용어는

1) 『국가지식정보 연계 및 활용 촉진에 관한 법률』 제2조에 따라 국가적 이용가치가 있어 국가지식정보위원회의 지정을 받은 정보

서비스 채널 내에서 오류, 보완 등에 대한 사용자 의견을 상시 접수하여 유지보수한다.



[그림 2] 정보통신용어사전 웹/앱

2) 표준용어 선정과 작성

표준용어는 내외부 및 언론 기사에 많이 사용되거나 활용이 기대되는 기술 등으로 정부 보고서, 보도자료, 기술 동향 및 언론에 많이 노출된 용어 중 중요도, 활용도 등을 고려하여 선정한다. 단, 제품이나 상표, 회사와 같은 고유명사, 특정인이나 조직 소유의 자료에서만 사용된 용어 등은 지양한다.

표준용어는 공공언어 성격이 있다. 공공언어는 좁은 의미에서 공공기관이 일반 국민을 대상으로 공공의 목적으로 사용하는 언어이지만 넓은 의미에서 일반 국민을 대상으로 사용하는 모든 언어를 말한다. 따라서 표준용어를 작성할 때는 정확한 개념을 표현한 용어, 이해하기 쉽고 오해할 가능성이 적은 용어, 어문규범에 맞는 용어를 사용할 필요가 있다.

이를 위하여 TTA는 알기 쉽고 올바른 국어 표현으로 표준용어를 작성하는 것을 원칙으로 용어표준화 추진 절차에 국어 운문·감수와 국어순화를 포함하여 운영한다. 특히 비전문가 대상의 용어, 즉 국민 생활과 관련성이 높거나 공공성이 강하고 국민 일반이 보도자료, 홈페이지 등에서 자주 접할 수 있는 용어는 좀 더 쉽고 자연스럽게 작성될 수 있도록 한다. 다만 해당 업계 종사자 또는 유사 계열 전공자(대학생 이상)를 대상으로 하거나 용어 자체의 전문성과 고유성이 인정되는 경우, 널리 사용하는 표현이 있을 경우는 원어 등을 사용할 수 있도록 예외로 둔다. 무리한 국어순화는 활용을 떨어뜨리고 용어에 대한 혼란을 초래할 수 있기 때문이다. 이때의 비전문가는 정규 고등학교 교육과정을 수료한 일반인을 말한다. 전문가는 대학 전공자 이상을 말한다.

일반인의 전문가 시대를 맞이하면서 TTA의 전문용어 순화 노력에 더하여 더 좋은 양질의 결과물을 잘 만들어 낼 수 있을지 인공지능을 활용해 보았다.

3) 인공지능의 활용과 한계

Open AI사가 개발한 자연어 생성 모델 GPT-3.5, GPT-4가 출시되면서 전 세계적으로 사회 모든 분야에서 인공지능 붐이 일기 시작했다. 주어진 프롬프트(prompt)를 기반으로 자연스럽게 창의적인 텍스트를 만들어 주고, 질문에 답변을 제공하고, 다른 언어를 번역해 주고, 대화해 준다. 사용자가 궁금한 용어가 있을 때 사전을 이용하지 않고 인공지능에게 물어 해결할 것 같은 수준이다.

인공지능을 표준용어 대상 선정과 작성에 활용할 수 있을지 ChatGPT Plus로 부분 시도해 보았다. 사이버 보안 등과 같은 기술에서 주로 사용하

는 주요 용어, 최근 이슈가 있는 기술 용어 등을 뽑아주고 선정 기준, 개수에 맞게도 추려준다. 사전에 적합한 대상일지, 중요도 등을 고려해 전문가가 검토할 수 있는 초기 데이터로 활용할 수 있었다.

용어사전 집필 양식을 주면 해설(안)을 그럴듯하게 만들어 준다. 국외 연구보고서 등 참고할 자료 등을 적절하게 번역·설명해 주고 최종 결과물도 빠르고 효율적으로 만들어 작업 능력을 높인다. 반면 몇 가지 문제점이 보였다.

- 용어별 중점적으로 설명 필요한 요소가 다를 수 있고, 필요한 답을 끌어내기 위해 다양한 관점을 사용하거나 답변해 주는 내용을 분석해 다시 질문하는 다중, 심층 프롬프트가 필요할 수 있다.
- 응답 내용을 이해할 수는 있으나 구체적이지 않고 국어적으로 활용하기에 부자연스러움이 있다.
- 전문가의 검증이 필요하다. 즉, 참고로만 활용하는 한계가 있다.
- 해결되지 않은 저작권 이슈가 있다.

그 외, 인공지능이 제안해 준 보다 쉬운 표현을 사용할 수 있을지를 검토해 보았을 때 일부 설명은 말이 되는 것 같으나 실제로 검토했을 때 내용의 정확성이 떨어지기도 했다.

III. 결론

국어기본법 제17조(전문용어의 표준화 등)에 따르면 ‘국가는 국민이 각 분야의 전문용어를 쉽고 편리하게 사용할 수 있도록 표준화하고 체계화하여 보급하여야 한다’라고 규정되어 있다. 이에 따라 TTA는 용어표준화를 통해 바르고 이해가 쉬운 표현을 사용한 표준용어를 만들고 전문가를 포함한 일반에 보급·확산하기 위해 노력 중이다.

본 논문에서는 그 일환으로 TTA 정보통신용어사전의 용어표준화 추진 절차 등 일반 현황을 알아보고, 전문 분야 용어사전에 수록할 만한 표준용어 대상 선정과 다양한 사용자의 이해를 돕기 위한 정확하고 알기 쉬운 표현 사용에 인공지능을 활용해 본 이점과 한계를 살펴보았다. 향후 인공지능은 발전을 거듭할 것이다. 더 자연스럽게 정확한 설명을 만들어 내겠지만 용어사전과 같이 하나의 전문 플랫폼에서 여러 분야의 전문가들이 검토하고 심의한 신뢰도 있는 콘텐츠를 만들어 내기에는 어쩔 수 없는 한계가 있어 보인다. 인공지능은 사전을 대체할 새로운 도구가 아닌 콘텐츠 제작 시간을 단축하고 품질과 효율성을 높이는 등 우수한 조력자로서의 활용 측면에서 좀 더 검토될 필요가 있다.

ACKNOWLEDGMENT

본 논문은 2024년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No. 2022-0-00007, ICT 국내 표준화 연구)

참 고 문 헌

- [1] 국립국어원, 한눈에 알아보는 공공언어 바로 쓰기, 2022.12.
- [2] 한국정보통신기술협회(TTA) 정보통신용어사전, <http://terms.tta.or.kr/>