

기계번역 품질에 대한 비판적 고찰

A critical look at machine translation quality

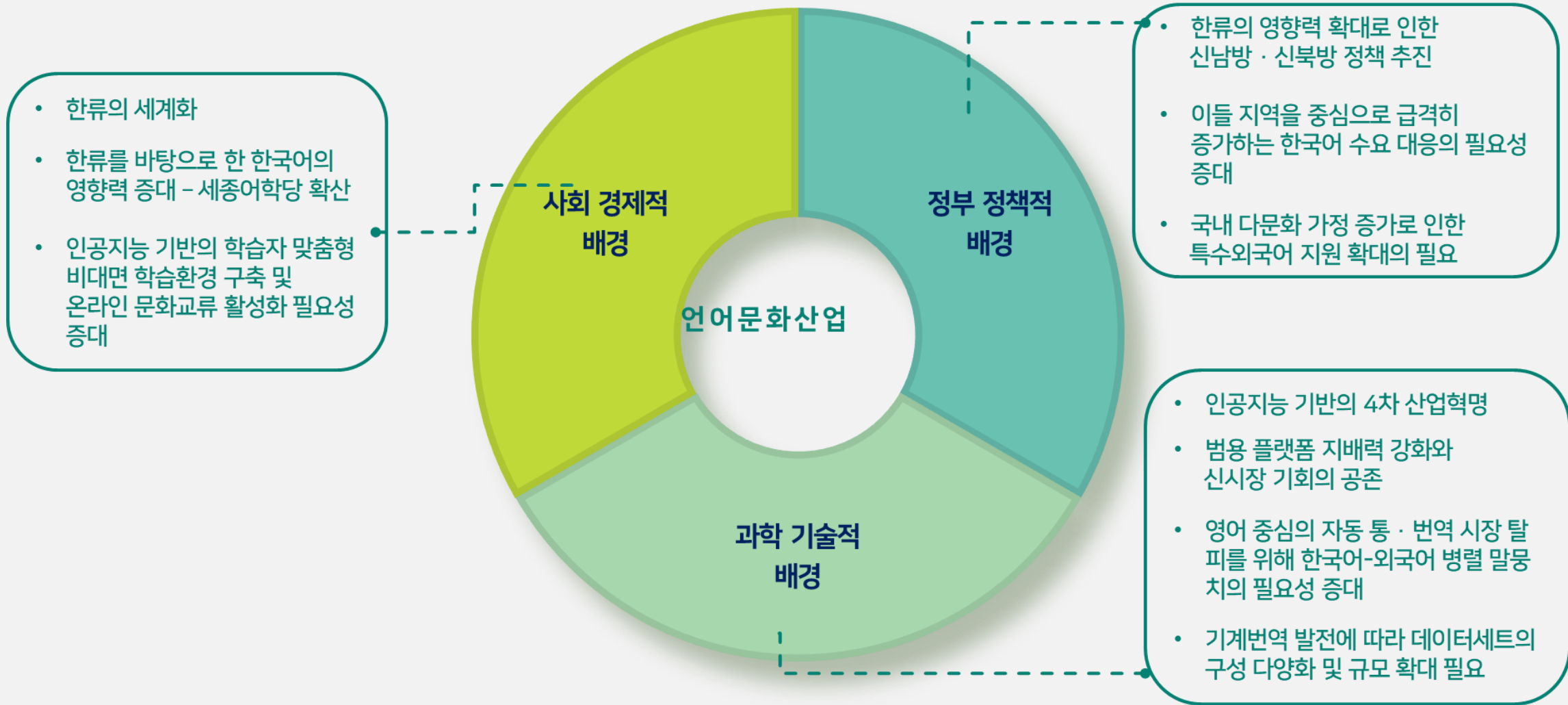
한승희, 한국외국어대학교 통번역대학원

Seunghee Han, *HUFS GSIT*

1. 논의 배경
2. 기계번역 고도화 사업의 당면 과제
3. 저자원 병렬말뭉치 구축관행의 문제
4. 국내 번역 산업의 사례 분석
5. 전문 감수·평가 교육과 정립된 포스트에디팅 지침의 중요성

1. 논의 배경

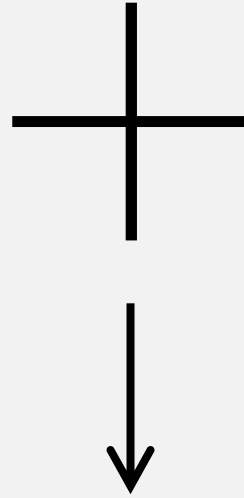
한국어·한국문화의 세계적 확산



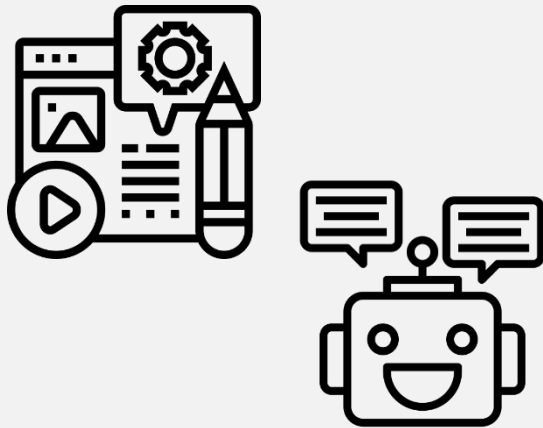
4차 산업혁명과 언어문화산업의 진화



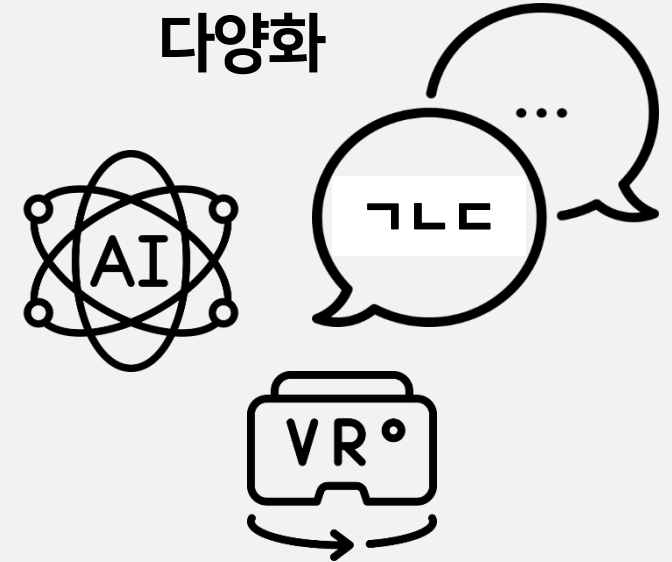
정보통신 및 융합기술의 고도화



글로벌 콘텐츠 유통 플랫폼의
다양화



'디지털'과 '콘텐츠'를 양 축으로 하여
언어문화산업의 전세계적 확산





2. 기계번역 고도화 사업의 당면 과제

기계번역 고도화의 당위성

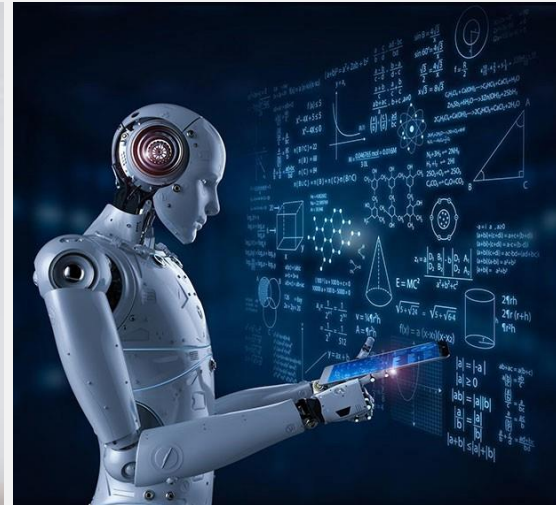
콘텐츠 산업



언어 교육



다문화 사회



중소기업 활성화



미디어용
자막 번역기

한국어 및 특수외국어
학습자용 인공지능 튜터

다문화 가정 지원용
챗봇 및 기계번역

해외진출 중소기업용
언어 지원

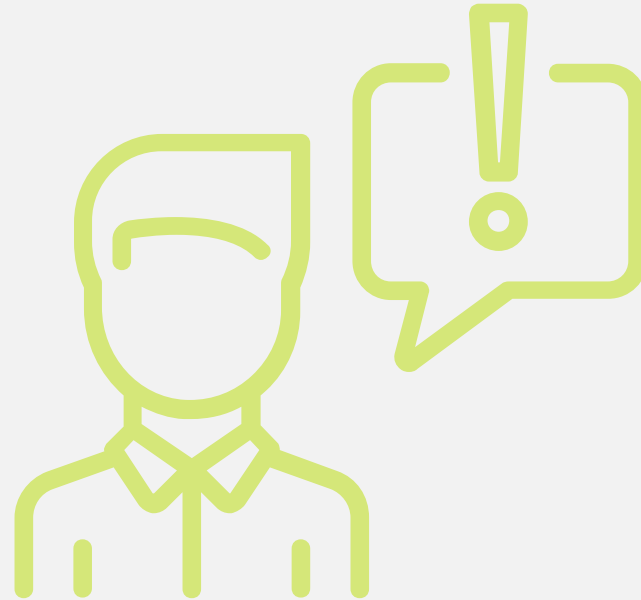
기계번역 고도화 사업의 당면 과제

1

특수외국어 병렬말뭉치 구축 자체가 희소

2

저자원 언어 병렬말뭉치 구축 시 데이터 증강기법 사용에 따른 문제



3

시장 내 특수외국어 병렬말뭉치 구축 관행상의 문제

3-1. 한국어 이해가 부족한 국내원어민 (다문화 가정원) 활용 시 문제점

3-2. 일반 사용자 번역을 웹크롤링하여 기계학습 수행 시 문제점

3.3 검증되지 않은 현지원어민 구인 관행 절차의 위험성

4

특수외국어 기계번역의 품질 문제

3. 저자원 병렬말뭉치 구축관행의 문제

특수 외국어 병렬말뭉치 구축 자체 희소



저자원 언어 말뭉치 희소

2007 ~ 2018, ETRI에서
베트남어, 러시아어 구축
(낮은 접근성, 소규모, 유료)

특수 외국어 병렬말뭉치 전무

신남방, 신북방 언어의
병렬말뭉치가 대규모로
구축된 사례 전무



유럽어 병렬말뭉치 희소

2021년 하반기, NIA에서
유럽언어 병렬말뭉치를 언어당
30만 문장 씩 구축할 예정

저자원 언어의 데이터 증강기법 사용에 따른 문제



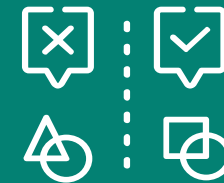
데이터 증강기법 사용 시, 모델로 생성된 'Pseudo Data'라는 한계로 인해 기계번역의 성능 향상이 어려움



페이스북의 CCMatrix, Flores-101 데이터세트는 대용량 웹 크롤링 데이터로 유사 대역문을 기계적으로 추출하여 단시간에 대량의 데이터를 모았으나 검증이 이루어지지 않아 정확도가 떨어질 가능성이 큼



50여 명의 저자가 공저한 웹 크롤링 방식의 다국어 데이터 세트 품질 분석 논문 (캐스웰 외, 2021)에서 205개 언어 데이터 중 저자원 데이터에서 수용 가능한 품질 수준은 50% 미만이라는 결과 제시



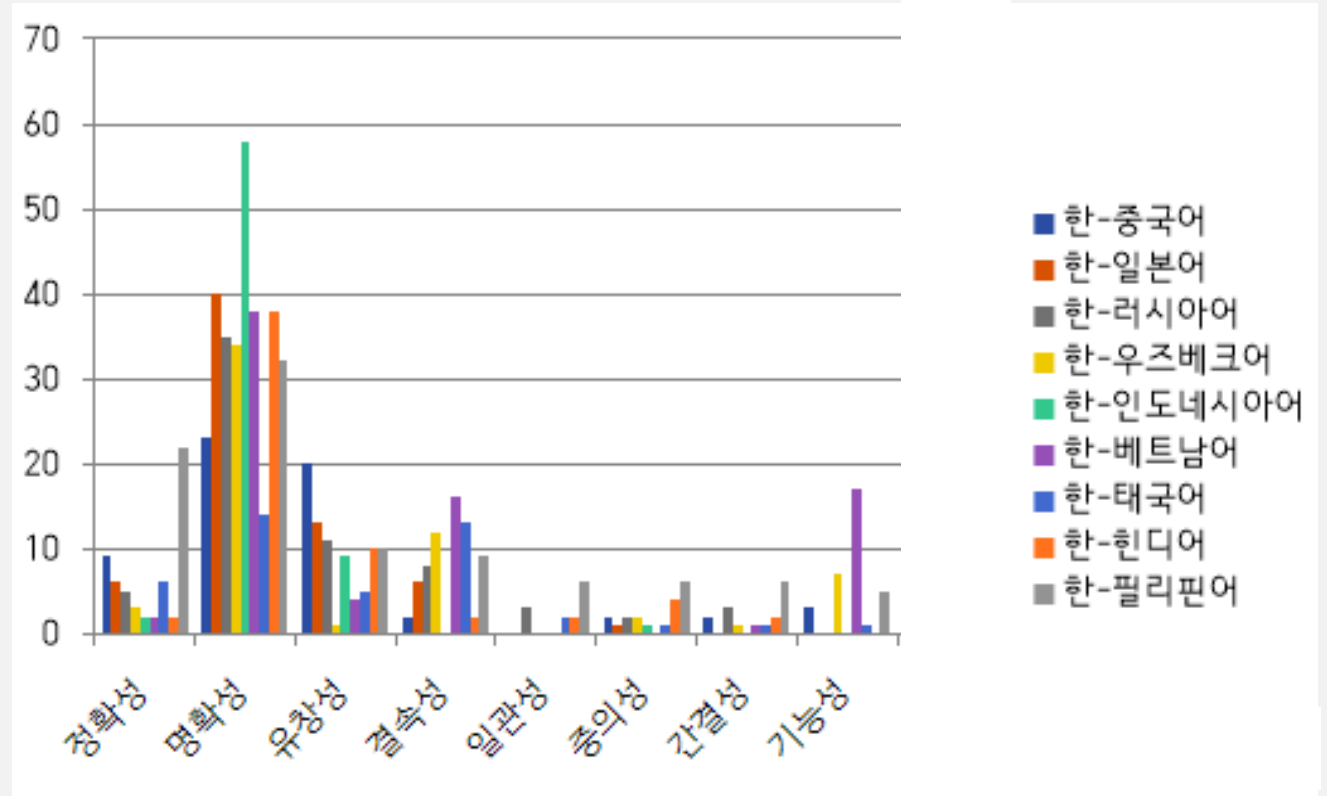
웹 크롤링 방식으로 대규모 데이터 구축 시 품질 확보를 위해 반드시 데이터 검수와 정제가 필요

4. 국내 번역 산업의 사례 분석


특수외국어 병렬말뭉치 부족으로 기계번역품질 저조



오류율 (%)



<특수외국어의 기계번역 고빈도 오류 유형>



가. 한국어 이해가 부족한 국내 원어민(다문화 가정) 활용

나. 일반 사용자의 번역을 웹 크롤링하여 기계학습 수행

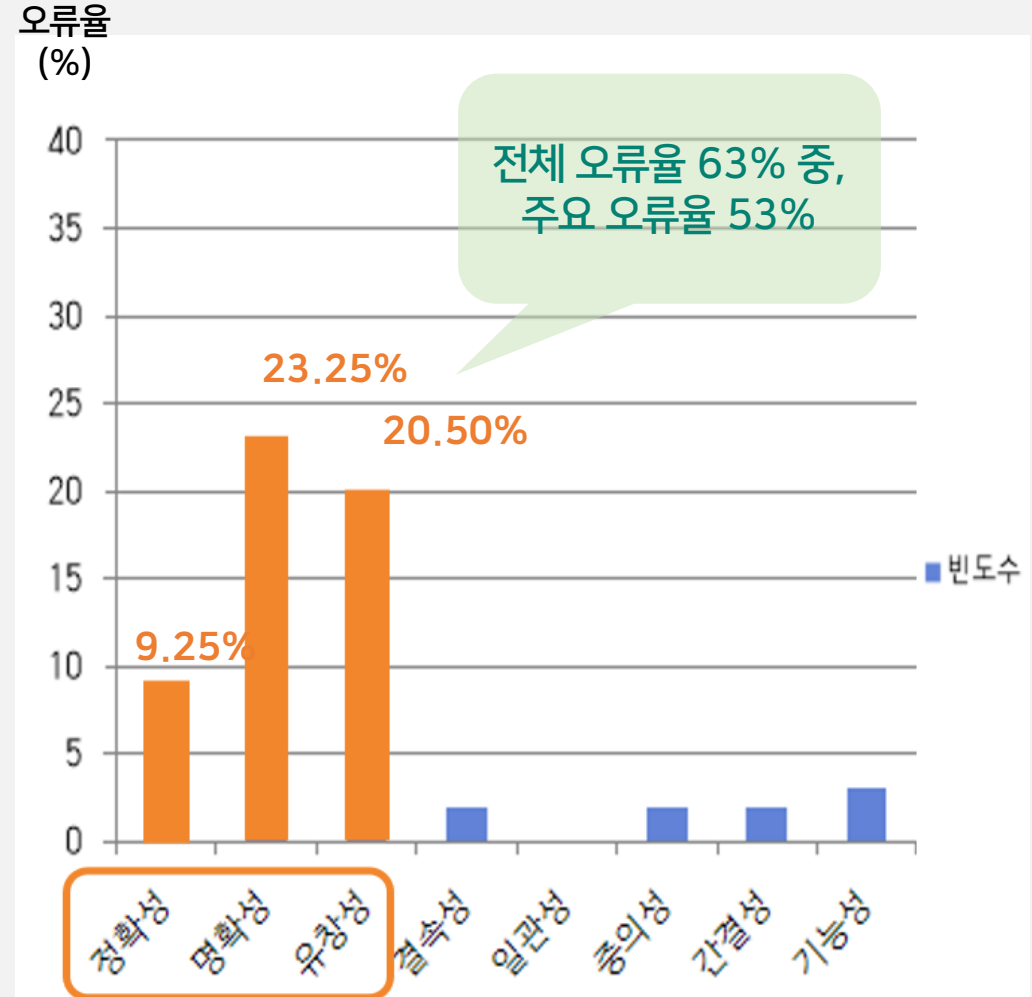
다. 검증되지 않은 현지원어민 구인 관행 절차의 위험성

가. 한국어 이해가 부족한 국내 원어민 활용

한국지능정보사회진흥원은 2021년부터 병렬말뭉치 품질 감수 및 검수 기준을 대폭 강화, 클라우드워커 고용에 앞서 번역 능력 평가 시스템 구축을 입찰 조건으로 제시

<다문화 가정 활용 번역의 품질평가 결과>

비전문가인 다문화 가정원들이 주로 구축한 한국어-중국어 병렬말뭉치에 대해 번역을 품질평가함

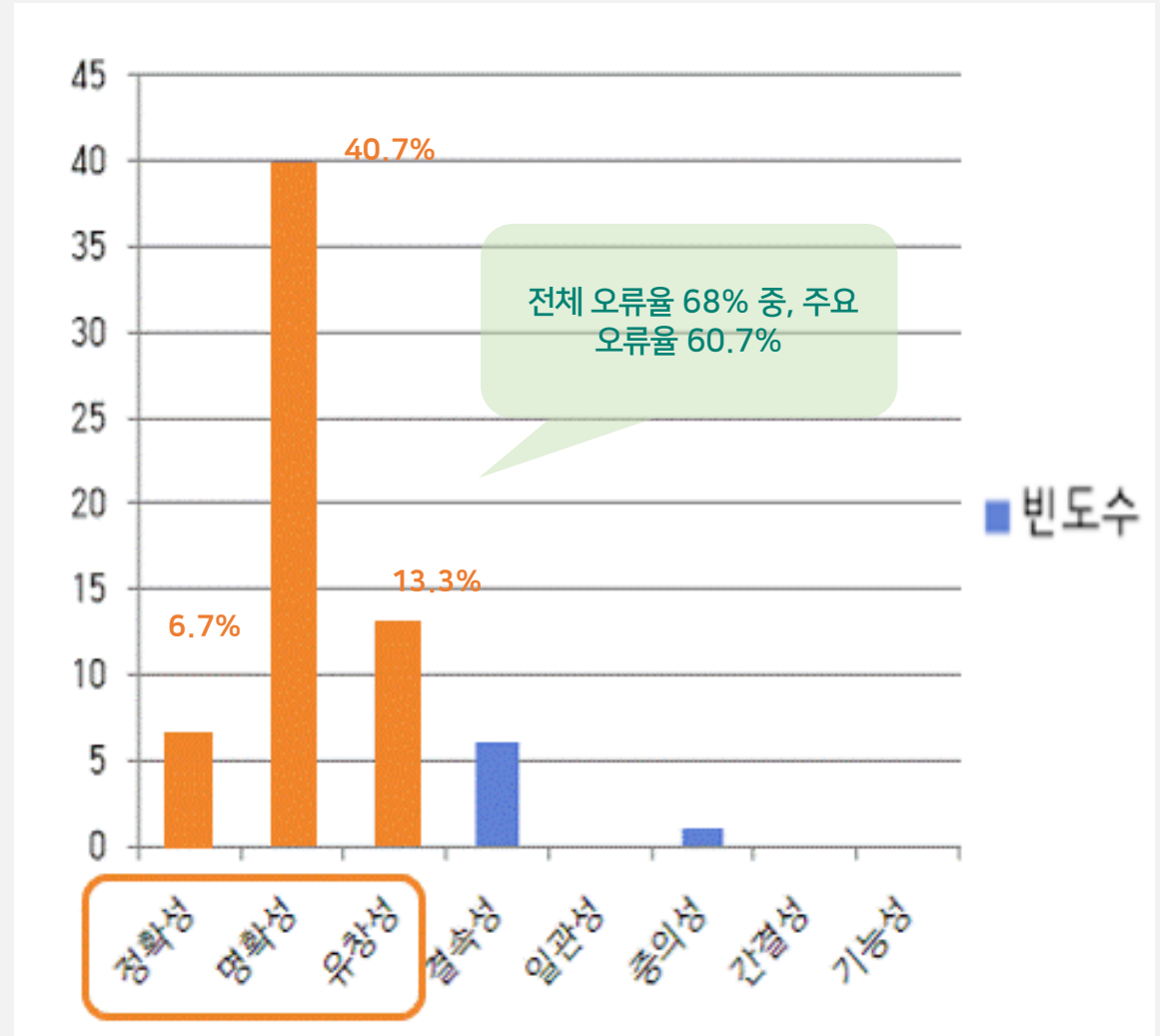


나. 사용자 번역을 웹 크롤링하여 기계학습 수행

클라우드소싱한 번역 데이터를 번역 전문가의 오류수정 없이
바로 학습용 데이터로 사용하면
기계번역 품질이 저하될 위험

<사용자의 앱 로그 번역 품질평가 결과>

네이버 사용자의 앱 로그 번역 데이터 품질평가



다. 검증되지 않은 현지 원어민 구인 관행 절차



비전문가 중심의
클라우드워커 시스템 활용

기계번역을 활용해 영어로 증역한
결과물을 해당 언어권 내 번역사에게
보내 최종 결과물 획득

번역 능력 검증 미흡으로 인해
번역품질에 타격

원문의 정보 손실로 인해 번역문에서
명확한 의미 전달 및 문화적 표현 요소
누락

5. 전문 감수·평가 교육과 정립된 포스트에디팅 지침의 중요성

기존 기계번역 품질검증의 한계

BLEU 점수	해석
10점 미만	거의 의미 없음
10~19점	핵심을 파악하기 어려움
20~29점	요점은 명확하지만 많은 문법적 오류가 있음
30~39점	이해할 수 있는 양호한 번역
40~50점	고품질 번역

[사례 분석] 네이버 번역 지침 변경 사례

1. 형식은 첨부된 파일을 따른다.
2. 문장의 행을 잘 맞추고 문장 병합 등 편집을 하지 않는다.
3. 원문의 문장 성분을 최대한 유지하여 번역한다.
4. 어휘에 대해 사전식으로 설명하지 않는다.
5. 중의적 단어, 문장은 가장 적합한 의미만 표기한다.
6. 문장은 20단어를 초과하지 않는 단문 위주로 번역한다.

개정 전



개정 후 - 추가

1. 형식은 첨부된 파일을 따른다.
2. 문장의 행을 잘 맞추고 문장 병합 등 편집을 하지 않는다.
3. 원문의 문장 성분을 최대한 유지하여 번역한다.
4. 어휘에 대해 사전식으로 설명하지 않는다.
5. 중의적 단어, 문장은 가장 적합한 의미만 표기한다.
6. 문장은 20단어를 초과하지 않는 단문 위주로 번역한다.

번역 및 검수 절차에 대한 국제표준인 ISO 17100

3.1.4 Translator qualifications

The TSP shall determine the translator's qualifications to provide a service conforming to this International Standard by obtaining documented evidence that the translator can meet at least one of the following criteria:

- a) a recognized graduate qualification in translation from an institution of higher education;
- b) a recognized graduate qualification in any other field from an institution of higher education plus two years of full-time professional experience in translating;
- c) five years of full-time professional experience in translating.

기계번역 사후편집 (Post-editing) 절차에 대한 국제표준인 ISO 18587

5.2 Qualifications

The TSP shall determine the post-editor's qualifications to provide a service conforming to this document by obtaining documented evidence that the post-editor can meet at least one of the following criteria:

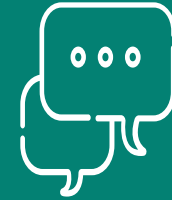
- a) has obtained a degree in translation, linguistics or language studies or an equivalent degree that includes significant translation training, from a recognized institution of higher education;
- b) has obtained a degree in any other field from a recognized institution of higher education and has the equivalent of two years of full-time professional experience in translating or post-editing;
- c) has the equivalent of five years of full-time professional experience in translating or post-editing.

전문가 집단의 체계적 품질 검증

다면적 검증 절차



전문적 평가 체계



고품질 번역말뭉치

감사합니다