Use of machine translation in healthcare settings: a symptom of a wider problem?

Centre for Translation Studies
UNIVERSITY OF SURREY

BACKGROUND

- Healthcare staff MT literacy is understudied compared with more general populations and lower-stakes situations.
- The use of MT with patient medical records is understudied compared with in-person situations and patient-oriented written communications.
- Inadequately standardised medical abbreviations have been shown to result in an increased risk to patient safety even before throwing MT into the mix, thus serving as a use case.

RESEARCH QUESTIONS

RQ1:

To what extent are healthcare staff aware of the potential risks posed to patient safety when using machine translation (MT) to translate patient medical record information?

RQ2:

What control mechanisms, if any, can be put in place to help mitigate the risks posed by the use of MT with patient medical record information?

Eleanor Taylor-Stilgoe
e.j.taylor-stilgoe@surrey.ac.uk

METHODOLOGY

- **Data collection:** use of authoritative clinical corpora in French (CAS Corpus, 2021, 4,900 cases) and Spanish (CodiEsp Corpus, 2020, 1000 cases) comprising a mixture of real, deidentified clinical cases and those published in scientific literature and legal and medical training resources.
- **Data preparation:** development of custom software with which to identify and filter relevant use case abbreviations in context, implemented using regular expressions based on specific inclusion and exclusion criteria. Mass translation of relevant source language results into English, again obtained via custom software with Google Translate API integration.
- **Data analysis:** analysis of resultant parallel French and Spanish to English corpora in Sketch Engine to identify and select MT use case data examples for review by healthcare staff during interview study.
- **Pilot study:** small-scale methodological pilot study conducted among healthcare staff to determine best approach to future interviews in terms of length, question format, MT use case data example presentation (contextualised according to role and presented blind), and data example confidence measurement format.
- **Key pilot study finding:** all healthcare staff participants, including a medical professional with 20+ years' clinical expertise, had difficulty identifying potentially problematic phenomena in the MT output.

La NFS montrait une thrombocytopénie à 87 000/ml.

Google Translate: NFS showed thrombocytopenia at 87,000/ml.

NFS translates in English to CBC (complete blood count) or FBC (full blood count).

The key term remains completely untranslated using MT.

FUTURE WORK

- Semi-structured interview study with 20+ healthcare staff performing diverse professional roles, including contextualised, blind MT use case data examples, through which to explore healthcare staff MT literacy and identify potential risks to patient safety posed by MT use.
- Interview data analysis to identify salient themes and patterns relevant to healthcare staff MT literacy and the potential risks posed by MT use in healthcare.
- Ongoing analysis of use case abbreviations and/or other potentially problematic phenomena identified as posing a risk when combined with MT use and with which to develop potential risk mitigation control mechanisms.

EXPECTED IMPACT

- Insight into the extent of MT literacy among healthcare staff and the potential impact of this on patient safety as applicable to real-life, higher-stakes settings.
- Development of potential educational and/or training recommendations to improve healthcare staff MT literacy.
- Identification of potential risks and determination of control mechanisms with which to mitigate the risks posed by the use of MT with patient medical record information.

IMPORTANT REFERENCES

Bowker, L. and Buitrago Ciro, J. (2019) *Machine Translation and Global Research: Towards Improved Machine Translation Literacy in the Scholarly Community*. Bingley (United Kingdom): Emerald Publishing Limited, pp. 111. ISBN 978 1787567221.

Vieira, L. N., O'Hagan, M. and O'Sullivan, C. (2020) 'Understanding the societal impacts of machine translation: a critical review of the literature on medical and legal use cases'. *Information, Communication & Society*, 24(11), pp. 1515–1532. doi: 10.1080/1369118X.2020.1776370.

Vieira, L. N., O'Sullivan, C., Zhang, X. and O'Hagan, M. (2022) 'Machine translation in society: insights from UK users', Language Resources and Evaluation, (0123456789). doi: 10.1007/s10579-022-09589-1.

This research has been funded by Research England, through the Expanding Excellence in England (E3) award

UKRI Research England