

Machine translation as support for epistemic capacities: Findings from the DECA project

Maarit Koponen¹, Nina Havumetsä¹, Juha Lång¹, Mary Nurminen²

¹University of Eastern Finland, ²Tampere University,

¹firstname.lastname@uef.fi, ²firstname.lastname@tuni.fi

Abstract

The DECA project consortium investigates epistemic capacities, defined as an individual's access to reliable knowledge, their ability to participate in knowledge production, and society's capacity to make informed, sustainable policy decisions. As a tool both for accessing information across language barriers and for producing multilingual information, machine translation also plays a potential role in supporting these epistemic capacities. In this paper, we present an overview of DECA's research on two perspectives: 1) how migrants use machine translation to access information, and 2) how journalists use machine translation in their work.

1 Introduction

The ability to access relevant, reliable knowledge and information is an essential part of epistemic capacity (Werkheiser, 2016). From the perspective of individuals, equal availability and accessibility of information can be seen as a fundamental epistemic right (Nieminen, 2024). While modern societies are increasingly multilingual, information produced by societal institutions like public administration and media is generally limited to dominant local languages and to a lesser extent some lingua franca (e.g. English). Information is therefore not equally accessible and intelligible to all members of society.

Questions of epistemic capacities and epistemic rights are at the core of the research carried out by the project consortium Democratic epistemic capacity in the age of algorithms (DECA) formed by University of Helsinki (consortium coordinator), University of Eastern Finland, Tampere University, Aalto University and the Finnish Youth Research Society. It is funded by the Strategic Research

Council established within the Research Council of Finland (2022–2025). Different work packages of this multidisciplinary consortium¹ address various aspects of epistemic rights and capabilities from the perspective of institutions, infrastructures and individuals.

This paper focuses on WP3 *Linguistic barriers, algorithms and epistemic capabilities*, which investigates linguistic accessibility and the role of machine translation (MT) in realising epistemic rights and capacities. The work conducted by the research team at the University of Eastern Finland in collaboration with the Finnish Youth Research Society has two main lines of research. The first line focuses on the use of MT by migrants as support for finding and accessing information. The second investigates how journalists use MT as part of their work. The next sections outline the current status of the work conducted in these lines of research, as well as directions for future work.

2 MT for linguistic accessibility

MT is an important tool for migrants, particularly vulnerable migrants like refugees and asylum seekers (Vieira, 2024a). DECA has aimed to investigate in more detail how migrants who do not speak Finnish use MT to find, access and use information about Finnish society. We have focused on people who have arrived in Finland from Ukraine or Russia since 2022.

Data collection started in spring 2024 with focus group discussions (n=35) mapping the participants' information needs, information sources and linguistic barriers they encounter. Participants were then invited for individual interviews (n=29) complemented with simulated information search experiments. The experiment consisted of tasks where the participants were asked to find information related to aspects of life in Finland (e.g. library services,

© 2025 The authors. This article is licensed under a Creative Commons 4.0 licence, no derivative works, attribution, CC-BY-ND.

¹See <https://www.decatutkimus.fi/research>

public transport). The purpose of the task was to observe their strategies for searching information and their use of MT. Screen recordings and think-aloud were recorded for analysis.

A preliminary analysis shows that most participants attempted to find information in their first language (Ukrainian or Russian), which provided relatively little information. The most effective use of MT was first translating the search phrases into Finnish to find relevant pages and then using MT to read these pages. However, this strategy was used by only a minority of the participants, mainly young adults who showed good information skills. A more detailed analysis of the interviews and tasks is currently ongoing.

3 MT in journalism

Journalism plays a vital role in providing relevant societal information (Watson, 2018; Nieminen, 2024). While previous work has developed MT tools for use in journalism (e.g. van der Kreeft et al., 2022), the perspective of journalists using MT remains relatively unexplored. DECA has therefore aimed to investigate more closely the ways in which journalists in Finland use MT in their work.

The first stage of research starting in 2023 focused on journalists at the Finnish national broadcasting company Yle producing news in English, Russian and Ukrainian. We conducted interviews (n=7) exploring how the journalists see their role in promoting the epistemic capacities of different language groups and how they use MT as part of their work (Havumetsä and Nurminen, 2025). Findings from the interviews show that MT use varies in the different departments. The Ukrainian news are mainly produced by translating Finnish news, and the work relies on an MT tool provided by Yle. In the other departments, MT use is less common, and the attitudes of journalists toward MT vary. Benefits identified by the journalists include making their work faster and allowing them to access more diverse sources in different languages. On the other hand, they expressed some doubts related to the quality of MT and ethical aspects.

In 2024, the work has been extended to MT use by journalists producing news in Finnish. We conducted interviews (n=10) and a survey (n=69) focusing on the use of MT, which was circulated to journalists working at Finnish media houses. Analysis of the data is currently ongoing.

4 Future work

As next stage of the work, we aim to further investigate the use of MT in Finnish public institutions, particularly immigration and integration services. A survey inspired by prior work carried out in the UK (Vieira, 2024b) and discussions with stakeholders in the field is being planned for 2025. Additionally, we are producing materials aimed at increasing MT literacy among different user groups, such as journalists, public administration or teachers, to be published in 2025.

Acknowledgments

This work is part of the DECA project funded by the Strategic Research Council established within the Research Council of Finland, funding agreements 352577 (University of Eastern Finland) and 352557 (consortium coordinator University of Helsinki).

References

- Nina Havumetsä and Mary Nurminen. 2025. (Kone)kääntäminen Ylessä ja muunkielisten episteemiset oikeudet. *Mikael: Finnish Journal of Translation and Interpreting Studies*, 18(1):7–22.
- Hannu Nieminen. 2024. Why We Need Epistemic Rights. In Minna Aslama Horowitz, Hannu Nieminen, Katja Lehtisaari, and Alessandro D’Arma, editors, *Epistemic Rights in the Era of Digital Disruption*, pages 11–28. Springer International Publishing, Cham.
- Peggy van der Kreeft, Alexandra Birch, Sevi Sariisik, Felipe Sánchez-Martínez, and Wilker Aziz. 2022. GoURMET – Machine Translation for Low-Resourced Languages. In *Proceedings of the 23rd Annual Conference of the European Association for Machine Translation*, pages 339–340, Ghent, Belgium. European Association for Machine Translation.
- Lucas Nunes Vieira. 2024a. Machine translation and migration. In Brigid Maher, Loredana Polezzi, and Rita Wilson, editors, *The Routledge Handbook of Translation and Migration*, pages 221–234. Routledge, London.
- Lucas Nunes Vieira. 2024b. Uses of AI Translation in UK Public Service Contexts: A Preliminary Report. Technical report, Chartered Institute of Linguists CIOL.
- Lani Watson. 2018. Systematic Epistemic Rights Violations in the Media: A Brexit Case Study. *Social Epistemology*, 32(2):88–102.
- Ian Werkheiser. 2016. Community Epistemic Capacity. *Social Epistemology*, 30(1):25–44.