Journal of Audiovisual Translation

Human Agency in the Age of Technology

Nina Reviers ✉
University of Antwerp

Gert Vercauteren ✉
University of Antwerp

Joselia Neves ✉
Independent researcher

Reviers, N., Vercauteren, G., & Neves, J. Human Agency in the Age of Technology. *Journal of Audiovisual Translation*, *7*(2), 1–5. https://doi.org/10.47476/jat.v7i2.2024.

Published: December 19, 2024.

Abstract

The landscape of audiovisual translation (AVT) is rapidly evolving due to AI and language technologies that are fundamentally reshaping production dynamics, workflows and professional profiles. This special issue examines the impact of these advancements on the media localization industry and, in particular, the professionals involved. Contributions in this special issue explore challenges faced by subtitlers in template-driven environments and the effectiveness of automatic speech recognition (ASR) tools, underscoring the need for human intervention to maintain quality. Additionally, the articles address broader social movements emphasizing representation and inclusion in media accessibility. Collectively, this issue illustrates that while technology offers efficiency, the irreplaceable role of human expertise in creativity and contextual adaptation remains crucial. The future of AVT will depend on a collaborative approach, where technological advancements complement human insight to enhance the quality and accessibility of translated media.

Keywords: audiovisual translation, technology, artificial intelligence.

✉ nina.reviers@uantwerpen.be, https://orcid.org/0000-0003-0271-6662
✉ gert.vercauteren@uantwerpen.be, https://orcid.org/0000-0001-6711-2005
✉ go@joselianeves.com, https://orcid.org/0000-0002-4406-8379

Introduction

The field of audiovisual translation is undergoing significant transformations, primarily driven by the emergence of cloud-based systems and an AI-powered media localisation industry. Technological advancements are heightening the demand for expedited content delivery, enabling simultaneous releases across multiple languages and fundamentally altering production dynamics.

These technological advancements are rapidly reshaping translation workflows and AVT services. Audiovisual translation processes are now facilitated by the increased integration of machine translation (MT) and other AI technologies. Speech recognition and synthetic speech enable automated subtitling and captioning as well as realistic synthetic voices for dubbing, voiceover and audio description purposes. Automated translations into multiple languages are produced at the click of a button. Real-time translations in live settings are streamed globally in an ever-growing list of languages. Workflows are accelerating, turnaround time and production costs are decreasing, and quality is "satisficing" (Simon, 1956). It has become difficult for human translators to compete. Neural machine translation (NMT) models, as those used in Google and DeepL's products, are becoming increasingly capable of handling the translation of metaphors, idiomatic expressions and other culturally sensitive or context-specific content. What could only be done with written text is now possible in the context of film, in which image, sound and words impinge upon each other for meaning-making. The ability to decode images is achieved by using AI to support blind individuals by describing visual elements like scenes, actions, and expressions, for instance. It might not be possible, just yet, to produce audio descriptions that are fluidly intertwined with the original audiovisual text, but it certainly supports the audio describers' work, should they use AI in the scripting phase.

The contributions of AI-driven communication mediation have their advantages. It allows for customization and personalization. Machine learning models and viewer profile analysis enable personalized tailoring of outputs, which can be particularly useful for individuals with specific needs. This promises a significant impact on accessibility services, which are bound to increase in availability and reach.

Technological developments are certainly impacting the landscape of professional AVT services and the translators' roles. Recent developments such as LLMs and GenAI give rise to new professional profiles in the language service provider (LSP) industry attributable to skills required by artificial intelligence, such as prompt engineering or, custom LLM development. Furthermore, the European Federation of Audiovisual Translators 'Machine Translation Manifesto underscores the necessity of ethical considerations and collaboration among diverse stakeholders in this evolving landscape.

This is the topic of the first article in this special issue entitled "Placing Human Agency in the AI-powered Media Localisation Industry". In this article, Orrego-Carmona examines the positions of key stakeholders in the media localization industry, elucidating both commonalities and conflicts among their interests. He notes that while the industry has adopted AI as a means to address the accelerated growth of media production and compensate for a perceived shortage of professionals, professional

associations have issued stark cautions against the implications of post-editing and the integration of AI in media localization. These associations emphasize the adverse impact on working conditions that arise from automated processes, which may ultimately compromise quality standards. The article advocates for mutual recognition and understanding among stakeholders, intending to ensure that the agency of human translators is acknowledged within this rapidly evolving landscape.

The second article in this issue also addresses the role of human AV translators and expresses concerns about challenging working conditions, specifically for subtitle professionals. In "Are templates killing the subtitling principles? Subtitlers' perception of the impact of subtitling templates on the media localization industry in the era of streaming", Skwarek, Szarkowska and Szkriba conducted semi-structured interviews with professional subtitlers to explore their perspectives on working with modern template-based workflows and to evaluate the templates' influence on both the translation process and the quality of the end product. The findings of this contribution highlight the critical need for a shared commitment from all stakeholders in the industry to achieve high-quality outcomes.

Quality in subtitling is also the topic of the third contribution to this special issue, in which Davitti et al. explore the effectiveness of automatic speech recognition (ASR) tools to produce automatic subtitles for broadcast. Recent developments in speech recognition technologies have had a dramatic impact on the provision of live subtitling in media, and broadcasters worldwide are starting to deploy automatic subtitling in their workflows. However, reliable quality control measures are crucial to ensure quality. This paper compared the performance of two tools that produce automatic subtitles from English into Italian, focusing on accuracy and readability. They conclude that the output does not meet quality requirements yet and that ASR can, therefore, only be considered an intermediate step, emphasizing the role of human agency in AV workflows.

The added value of human translators also becomes clear in Korybski et al.'s contribution to respeaking, entitled "Human Agency in Live Subtitling through Respeaking: Towards a Taxonomy of Effective Editing". They explore the work of respeakers in live subtitling, through an analysis of effective editions, including condensation, re-expression, and compensation. They highlight the types of editions human respeakers perform and evaluate the impact of the live editing process on their performance and the output quality. Their analysis reveals the pervasive nature of such edits, which also emerge as significant predictors of respeakers' performance accuracy. They hope their taxonomy can facilitate a more equitable and pragmatic assessment of subtitle accuracy but also holds the potential for refining (semi-)automated subtitle accuracy evaluation systems.

The fifth contribution to this special issue continues to prioritize quality concerns in an AI-powered industry but switches the focus to Audio Description - a domain in which AI technologies have so far only had a very limited impact. However, as technologies that could be relevant for AD production, such as image recognition and voice synthesis, develop rapidly, understanding the quality thresholds to uphold is all the more important. Lyberg-Åhlander, Holsanova and Johansson focus on the study of narrative specificity and voice quality in Audio Description. In an experimental study, they

investigate the impact of narrative specificity and voice quality on imageability and comprehension in both sighted and non-sighted populations. The results of this experiment underscore the importance of voice quality and narrative specificity in AD and the implications for both professional audio describers and the development of automated AD systems.

The concern for human agency in Audiovisual Translation Studies and Media Accessibility Studies goes beyond the impact of technology alone. As illustrated by the final three articles in this special issue, the development of our field of study in the 21st century is also marked by impactful social movements for human rights, diversity and inclusion. Such movements have contributed to the enormous growth of AVT and, in particular, MA services over the past decade, under the influence of global and local regulatory processes. However, this development has also raised concerns about the philosophical and ideological assumptions underlying audiovisual translation and media accessibility. Representation and participation of users are key concerns in this context, and such ethical considerations become even more crucial in a rapidly evolving technological landscape.

De Ridder addresses the issue of representation of minoritized language communities in the AVT landscape, focussing on revoicing practices on streaming platforms, in her article entitled “Revoicing to give a voice to minoritized language communities”. She argues that “while the users of minoritized languages are often proficient in the dominant language of their country, content must be available in their smaller — and often endangered — languages as well.” De Ridder discusses the role AVT can play in this respect by using the Swedish Public Broadcaster as a case study.

Dangerfield switches the focus to Media Accessibility in her contribution entitled “Decolonising media accessibility: Alternative approaches to media access for film and live events”. The critical concern in her article is that of participation in every step of access creation processes: what is “the role and responsibility that we, as humans, have in determining how collaboration and participation can be successful within the research process, in bringing creativity and difference to the content that is created, and in challenging biases and discriminatory practices?” Linking this question to broader societal debates, this article questions the balance of ideology and art and explores how questions of justice lie behind the access that is provided.

Taking a similar philosophical stance, Romero Fresco takes the access creation process of his documentary Where Memory Ends (2022) as a case study to explore the ideas behind alternative access. By offering a critical view on compliance-based access practices and comparing them to alternative approaches, this article introduces notions such as access intimacy and disability justice into Media Accessibility Studies, offering new inroads into the study, practice and training of access services.

The works collected in this special edition show the urgency of embracing technology, and AI in particular, as an ally in audiovisual translation. They also show us that there is, and will always be, space for human engagement. Translators can leverage technology to enhance their work efficiency, quality, and reach.

Over time, technology has developed to support and ease human tasks. The Industrial Revolution introduced machines to take over the tedious, repetitive tasks of manufacturing. In the present context, technology and AI have also come to support white-collar workers in their industries. We can leave the initial work of background research and drafting to the machine, to apply our energy to refining and adapting texts to context and audiences. Translation memories will make repetitive tasks lighter and guarantee consistency across the board. By working with adaptive AI systems, turnaround time can be improved. Humans will always be required to enhance Neural Machine Translation (NMT) and in post-editing and quality assurance tasks. And even there, AI-driven tools can assist by identifying errors and inconsistencies that are difficult to spot manually, but the bottom line is that in the end human translators will always (have to) be responsible for the final product.

The call is now for adaptation and constant updates. Integrating technology and AI in existing workflows is a given. The future is one of close collaboration between technological and human intelligence. At present, human expertise overcomes that of AI in creative, complex and context-specific tasks. Machines work with existing knowledge, building on it to create new knowledge. Humans will make that knowledge more humane and will adapt it to the needs of real people, in real contexts.

References

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review*, *63*(2), 129–138. https://doi.org/10.1037/h0042769