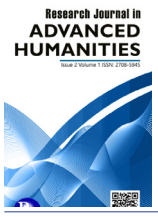

https://doi.org/10.58256/5p29dj87

Research Article

Section: Digital Humanities

Published in Nairobi, Kenya by Royallite Global.

Volume 5, Issue 4, 2024

Article Information

Submitted:4th October 2024
Accepted: 8th November 2024
Published: 3rd December 2024

Additional information is available at the end of the article

ISSN: 2708-5945 (Print)
ISSN: 2708-5953 (Online)

To read the paper online, please scan this QR code

How to Cite:

Abu-Rayyash, H. ., & Haider, A. S. . (2024). Introducing Transvisio: A customizable AI-powered subtitling tool. *Research Journal in Advanced Humanities*, *5*(4). https://doi.org/10.58256/5p29dj87

Introducing Transvisio: A customizable AI-powered subtitling tool

Hussein Abu-Rayyash[1] & **Ahmad S. Haider**[2]

[1]Kent State University, College of Arts and Sciences, Department of Modern and Classical Language Studies, Kent, Ohio, USA

[2]Department of English Language and Translation, Applied Science Private University, Amman, Jordan

*Correspondence: a_haidar@asu.edu.jo

https://orcid.org/0000-0002-7763-201X

Abstract

The growing demand for multilingual audiovisual (AV) content boosts expectations about efficient and diverse subtitling solutions accordingly. Traditional subtitling methods lack much-needed flexibility for catering to different audiences and various types of content. Therefore, this paper introduces TransVisio, an AI-driven tool that provides customized subtitling options with the use of Large Language Models (LLMs). It differs from other subtitle generators in that TransVisio allows users to customize subtitles to the content genre, target audience, language register, and even cultural nuances, to mention a few. Targeting professional translators, educators, and students, the tool combines artificial intelligence (AI) functions of the highest level with a user-centered interface in support of the subtitling process. Our paper describes the development and features of TransVisio and outlines some practical implications, ethical considerations, and future prospects to offer an overview of its contributions to audiovisual translation (AVT), AI language processing, and translation education.

Keywords: AI-powered tools, audiovisual translation, cultural adaptation, large language models, subtitling

Public Interest Statement

With the growing demand for multilingual audiovisual content, traditional subtitling methods struggle to meet diverse audience needs. This paper introduces TransVisio, an AI-powered tool that customizes subtitles based on content genre, audience, language nuances, and cultural context. Designed for translators, educators, and students, TransVisio combines advanced AI with an easy-to-use interface to enhance subtitle creation, making it more flexible and accurate. By improving accessibility and cultural relevance, TransVisio offers valuable contributions to audiovisual translation, AI language processing, and translation education.

1. Introduction

Media globalization has placed new demands on audiovisual (AV) content to cross languages and cultures (Saed, Haider, & Tair, 2023; Samha, Haider, & Hussein, 2023; Shuhaiber & Haider, 2023). Access to media by many viewers has made audiovisual translation (AVT), especially subtitling, an important tool for global accessibility (Al-Abbas & Haider, 2021; AlBkowr & Haider, 2023). AVT provides the key to unlocking media nuances with sensitivity across linguistic and cultural barriers, making it resonate with global audiences, promoting intercultural awareness, and providing better access to information and entertainment. This has become an urgent need in today's interlinked world. However, despite all the advantages, conventional subtitling does have a number of disadvantages concerning scalability in the cases of different genres and target audiences. Such issues may regard regional differences in languages and cultures, as well as content type specifics that can hamper the effective conveyance of a message.

This intricateness is the area most current AVT tools cannot cope with. Hence, it opens up a very important gap in the industry. Therein, there is an urgent need for creative solutions that could offer adaptable tools able to handle both linguistic and cultural complexities. Recent advances in artificial intelligence (AI) and large language models (LLMs) have brought new opportunities into view to automate and improve the subtitling process (Akasheh, Haider, Al-Saideen, & Sahari, 2024; Al-Salman & Haider, 2024; Farghal & Haider, 2024). However, Du and Lu (2024) suggested that the outputs of AI subtitles have yet to meet those of human translators. Therefore, while these AI-driven tools improved the efficiency of subtitling, they still require considerable human input to come up with high-quality outputs. Thus, Tardel (2023) proposes a workflow model of subtitling that considers a proper balance between purely manual and AI-assisted processes. This will include stages of transcription, synchronization, and editing, where the application of AI technologies like Automatic Speech Recognition (ASR) and Machine Translation (MT) is at the forefront without making human expertise superfluous. Hence, the emphasis is on making use of AI in such a way that one need not expend too much temporal and technical effort, without sacrificing the quality of the end product.

Contextual learning for adaptive real-time MT in subtitling has been carried out by Soudi et al. (2024). Their work demonstrated how dynamic retrieval of fuzzy matches combined with prompt composition, especially via more advanced LLMs such as GPT-3.5 Turbo, had the potential to grossly improve translation quality over and above static selection methods. However, sophisticated neural machine translation (NMT) and ASR systems have not been successful in addressing context-specific language demands. Such challenges are an illustration of the boundary in current AI systems that still are not at a high enough level in comprehension and adaptation to handle complex linguistic subtleties. Along these same lines, Arroyo Chavez et al. (2024) compared various forms of personalization of closed captioning for Deaf and Hard-of-Hearing (DHH) viewers by LLMs, ChatGPT, and AppTek's optimized models. In fact, most of the participants preferred to have verbatim captions; for some users, however, AI-edited versions were useful. This shows that subtitled solutions need to be personalized and

custom-fit to the users. Moreover, Briva-Iglesias, O'Brien, and Cowan (2023) found that users may be more inclined towards interactive post editing as it allows them to control the subtitles according to their needs and preferences. Additionally, Álvarez, Arzelus, and Etchegoyhen (2014) argue that accuracy should not be the only concern regarding the quality of AI subtitling. Hence, segmentation must be given attention to improve readability according to the specific needs of audiences and guidelines of companies. Therefore, subtitling requires customization options to balance these demands. TransVisio also develops user-adjustable tools for synchronization and segmentation intending to address these limitations and improve accuracy across a wide range of audiovisual genres. In this line, Soe et al. (2021) also present how the approach combining AI subtitling and human experience may help minimize errors and preserve cultural accuracy. As a result, with regard to professional subtitling, better collaboration between the AI system and humans through editing remains crucial.

Another model of Tardel (2023) points out the spectrum that may exist within subtitling: from completely manual to hybrid solutions that incorporate AI technologies to a greater or lesser extent. Based on this, TransVisio offers different customizable workflows according to the specific needs of subtitlers and AVT professionals. For instance, the subtitling workflow in TransVisio can be flexible, following either a direct model or one that is divided into stages, like transcription, translation, and post-editing. All this in specific concrete stages contributes to accuracy and provides an opportunity for specialization to operate at each step more easily. The subtitle quality produced this way is much more refined and contextually appropriate. Flexibility follows Tardel's proposed model, in which workflows were adjusted according to conditions that include transcript availability, the precision required in timing, and the genre of the content being translated.

This paper presents TransVisio, an AI-powered subtitling tool that is fully customizable. The current research aims to overcome the inefficiencies of the present AVT practice by effectively fostering efficient subtitling and ensuring the cultural and linguistic relevance of subtitles for diversified audiences. The consequences this could have regarding media localization, translation education, and accessibility make TransVisio an exceptionally remarkable innovation in AVT. This again goes to say that a balance must be achieved between the capability of AI and human oversight so that development in technology serves to increase, rather than diminish, the quality and authenticity of translated content.

2. Related Work

The field of AVT has recently undergone a large transformative change brought about by AI. Breakthroughs such as the development of Transformer-based models, including but not limited to GPT-4o, Claude, and Gemini 1.5, along with progress in ASR systems, have recently given AVT a push like never before. This makes AI increasingly swift and competent for more complex linguistic tasks. The transformer-based models presented herein hold an edge in capturing contextual information and accurately processing colloquial speech features for more culturally relevant natural translations (Dankers et al., 2022).

In the last two to three years, several AI-assisted subtitling systems have emerged that have become very useful in speeding up the process of subtitling. Soe et al. (2021) pointed out that AI-powered speech recognition devices have become an integral part of media production systems, automating much of the subtitling tasks, therefore reducing the amount of manual work required in transcription and synchronization. However, such systems regularly prove insufficient when approaching intricate creative texts with complex interpretative tasks well beyond the capability of the current machine learning (ML). Similarly, Jia and Sun (2023) suggest that the quality of the machine translation effects the speed, ease, and efficiency of the post editing process. This would presuppose that such limitations draw more human intervention within AVT to keep up with the level of complexity and richness required of such

sophisticated media.

Tardel's (2023) work on subtitling workflow models puts into light the changing face of AVT. In this regard, TransVisio was able to pick it up from here and included custom settings on tasks such as transcription and ASR use, based on which professionals will be able to refine the workflow further depending on given roles within AVT. Following Tardel, subtitling workflows should be flexible to fit the needs of different stakeholders: spotters, translators, and reviewers alike. In TransVisio, role-based customization lets professionals tune parameters according to their needs-be it directly creating subtitles, post-editing machine translations, or reviewing dialogue for cultural sensitivity.

Soe et al. (2021) discussed adaptive MT subtitling and general text translation in the context of LLM integration using Encoder-Decoder frameworks. Their results, quite interestingly, demonstrated that these LLMs were significantly more adaptable for application in translating generic texts instead of for subtitle-specific tasks, even in the case of language pairs such as English-to-Arabic and English-to-French. Dynamic retrieval methods and prompt-based solutions further improved translation quality, above all for the attainment of domain-specific needs. The approach has been applied in TransVisio, where the degree of automation for each subtitling task is granted with the respective customized work to produce quality. TransVisio illustrates the adaptive workflow underlining the balance of automation against human editing needs emphasized by Tardel (2023).

Equally illustrative is Matesub, a state-of-the-art subtitling tool using ASR and MT arranged in a cascade architecture. Unlike TransVisio, however, Matesub does not have rigorous customization which will enable changes depending on the genre of content or cultural adaptation. By contrast, TransVisio embodies the results of research about closed caption customization (Arroyo Chavez et al., 2024) and knowledge from SubGPT—a tool that uses OpenAI GPT for the generation of culturally adapted subtitles from source language subtitle template files (Abu-Rayyash & Al-Ramadan, 2024). While SubGPT performs the translations by using dialect-specific custom prompts and preserving the temporal relationships between the original subtitles, this would further optimize the efficiency and ease of access for AVT. On a similar note, by examining how specifications in the prompts provided to ChatGPT effect the qualities of its translations, Yamada (2023) found that customizations based on culture, purpose, and target audience provide more flexible results with better quality. TransVisio provides a simple and straightforward method to apply such modifications to enable users to achieve their desired outcomes with ease. In TransVisio, the user can define all parameters, such as linguistic complexity or demographic considerations. The level of granularity with TransVisio allows it to be a very multi-functional, sophisticated tool for a wide array of subtitling tasks that tools such as Matesub and SubGPT tend to fall short on.

Experiments on subtitle datasets, such as those of TED Talks by Soudi et al. (2024), show that larger datasets do not always give better translations, with particular evidence for the case of English-to-Arabic subtitles. This draws attention to the fact that content-specific quality rather than dataset size helps in achieving better MT performance-a guiding factor inherently built into the design of TransVisio. Recent research has focused on the subtitling customization aspect to entertain viewers. Arroyo Chavez et al. (2024) experimented with large language models in adapting captions; it was noted that most viewers preferred verbatim captions, while a few liked them simplified since it helped in reading. Gorman et al. (2021) showed adaptive subtitles that were aimed at real-time customization. This option increased the viewing experience manifold because it completely fit according to individual preferences. Arroyo Chavez et al. (2024) also highlighted the benefits associated with customizable AI-generated captions. For instance, whereas verbatim captions were usually preferred by respondents, some viewers preferred AI-driven edits that reduced complexity and increased reading speed. That would be consistent

with the approach of TransVisio since it offers a set of options for synchronization, linguistic style, and cultural appropriateness. Amin et al. (2021) investigated visual attributes that include high contrast and placement, emphasizing the fact that with no obstruction of placement and a capability for adjusting the subtitled visual settings, user satisfaction will be very well attained. Therefore, there is a suggestion from these studies that allowing adjustable parameters concerning the speed, density, and visual attribute of subtitles would yield high expectations for user satisfaction in every form of media, as imposed in the core TransVisio customization strategy.

While AI promises a great future, this field still faces quite a few challenges in terms of the satisfaction of users regarding the display of closed captions generated by AI. Arroyo Chavez et al. (2024) noted that even though LLMs like ChatGPT and AppTek were excellent at times in condensing the content, accuracy with nuanced meaning was not always well preserved, and hence there were instances where user expectation and actual performance were mismatched. These results point to the crucial contribution of human judgment in keeping contents culturally and linguistically appropriate—a principle on which TransVisio was designed.

Relating to ethicality, Mihalache (2024) pointed to critical stereotypes being perpetuated with the use of biased training data for AI-powered subtitling. To this end, diversity-oriented data collection, accompanied by serious validation, should be conducted so that these biases are appropriately managed. TransVisio animates these principles as foundational elements of its design philosophy. That makes the tool actively damp any probable biases, because it provides user-defined cultural adjustments. In this respect, TransVisio has managed to work continuously on the subtitled version to make it culturally refined, accurate, and sensitive for different viewers.

The last remark that can be made from this literature overview is that AI-powered tools hold a revolutionary potential to contribute to the quality, accessibility, and personalization of AVT. These developments equally make it clear, however, that human skill and judgment cannot be dispensed with, particularly in areas of cultural sensitivity and linguistic precision in AI-generated translations. A good example could be TransVisio, which marries AI efficiency and scalability with unmatched depth brought in by human editorial expertise. This harmonious balance caters to today's subtitling challenge with the user at the center while raising the quality standard of AVT to a higher level.

3. Overview of TransVisio

TransVisio is designed to provide an AI-powered, user-centric solution in subtitling, moving towards the tailoring of diverse content to varied audience needs. One key differentiator that sets TransVisio apart from competitors is broad personalization for genre settings, demographics, and culture adaptation. In contrast with other tools like Matesub that use a combination of ASR and NMT for subtitle generation, as pointed out by Perone (2023), the extra mile in TransVisio is in the workflow models, taking inspiration from Tardel (2023). In the workflows proposed by Tardel, the differentiation of roles and adaptive task division are emphasized, which easily blend into the user interface of TransVisio. This would mean flexibility at different levels for the user in subtitling: ranging from full automatic transcription and synchronization, right through to in-depth human-led adaptation of subtitled content.

Following the recent work of Soudi et al. (2024), TransVisio further incorporates prompt composition and dynamic retrieval of fuzzy matches in order to better adapt the subtitles. Prompt composition is a way of composing prompts tailored to the context of the content to get more accurate translations. That would mean the dynamic retrieval of fuzzy matches in TransVisio refers to previously translated segments that are similar to the present content, which leads to more accurate generation of subtitles with context awareness. A method like that would ensure domain-specific translations without

resource-intensive model retraining and would surely enhance efficiency with contextual accuracy. Therefore, TransVisio can deliver domain-specific translations more effectively without the overhead of extensive model retraining, hence maintaining efficiency and contextual relevance of the outputs.

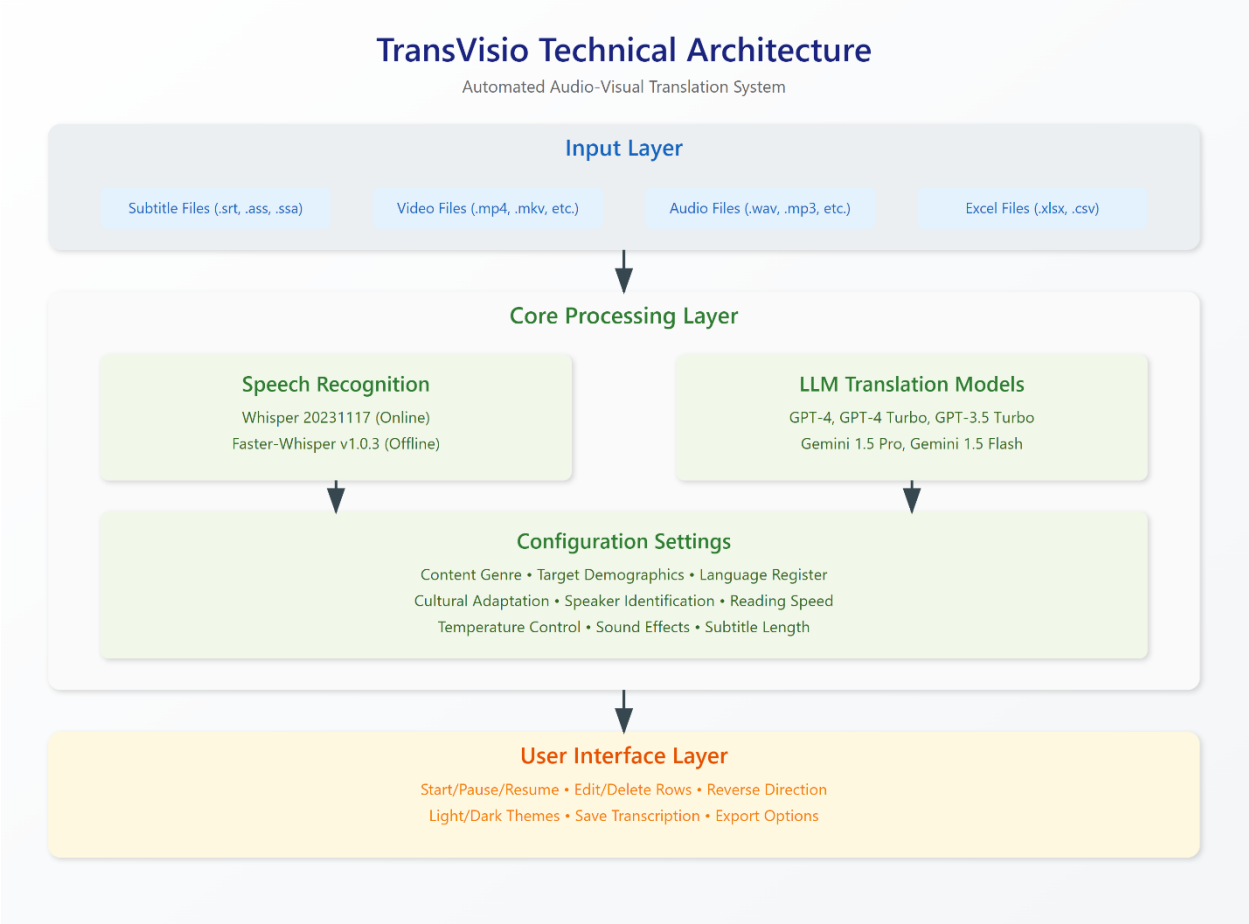

Figure 1. TransVisio Technical Architecture

Figure 1 illustrates the technical architecture of TransVisio, where several components such as ASR, NMT, cultural adaptation modules, and the post-editing interface work collaboratively towards an efficient subtitling pipeline. More precisely, the general architecture consists of three main layers: an input layer, the core processing layer, and a user interface layer; these collaborate effectively with each other to ensure a smooth workflow for subtitling. These layers process media inputs with ASR and NMT functionalities, further reworking them through cultural adaptation by a user-driven post-editing stage. This becomes a collaborative structure of AI and human translators, personifying automated efficiency with necessary human oversight. Thus, the integration reduces technical disruptions and times of processing in processing to an end-user experience that focuses on customization rather than troubleshooting. Also, the input layer is open to different types of media formats; for example, video, audio, subtitle files, and even Excel spreadsheets can be added. Therefore, this will add more flexibility to using TransVisio in different professional contexts.

OpenAI's Whisper handles this at the core processing layer for online processing, while Faster-Whisper v1.0.3 does so for offline applications, which giving flexibility to the extent that one's system capabilities or availability of an internet connection may allow. Whisper is a large model, factory-trained using large-scale weak supervision; whereas Faster-Whisper is a reimplementation of Whisper through CTranslate2 in order to offer efficient inference with Transformer models. LLMs are GPT-4, GPT-4

Turbo, and Gemini 1.5, all of which have deep connections with high-quality multilingual translation. More advanced configuration options for these models are going to allow users to modify features of content such as genre, emphasis on demographics, and degree of cultural accommodation. This will ensure that the subtitling output meets diversified needs of different contents and different audiences, enhancing the quality of subtitles and user satisfaction.

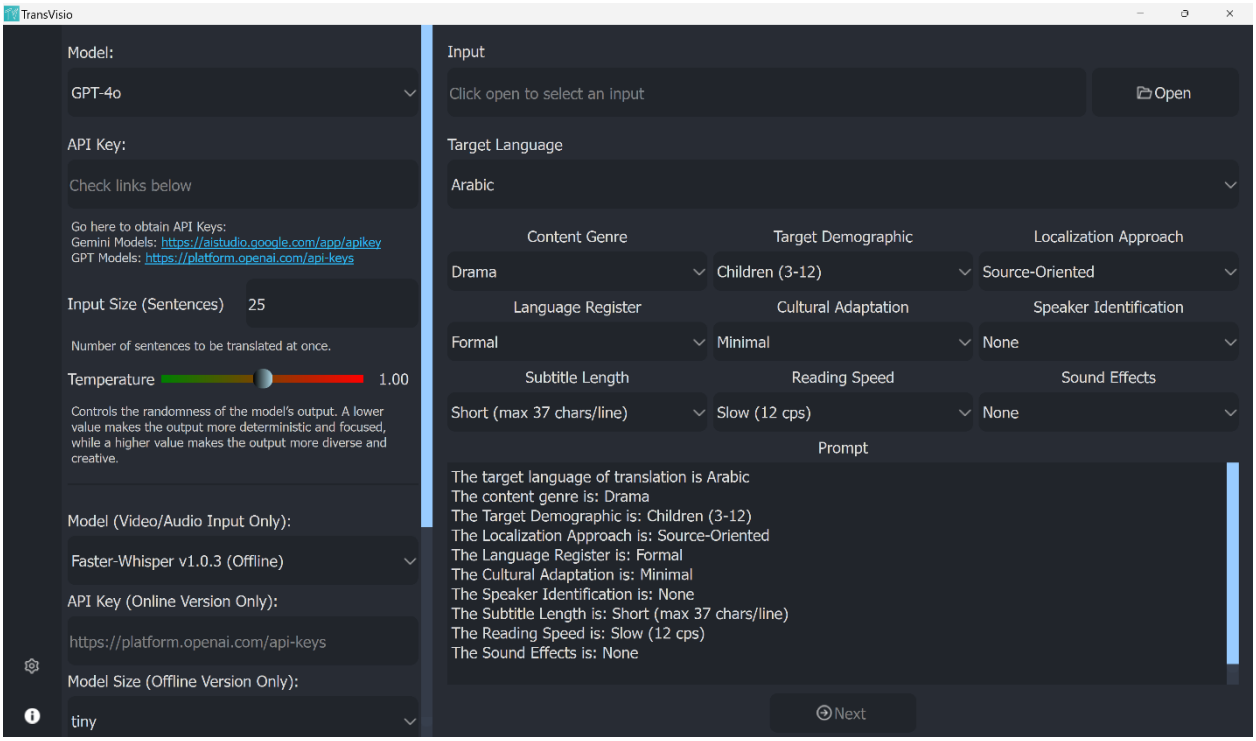

Figure 2. TransVisio User Interface

Figure 2 provides an overview of the general aspects of the TransVisio user interface, which was designed to be easy to use while keeping the cognitive workload as low as possible. In the interface, one notices first that dropdown menus for model selection, input types, and configuration parameters offer intuitive ways to navigate through the interface independently of any prior experience from the user. The interface is enriched by the Temperature Control slider, through which one gets closer or further away from randomness in AI outputs; that is, users can opt for either more deterministic or creative translations. The key sections of the interface are the model and language selection in the sidebar, the configuration of inputs, translation settings in the central workspace, and the advanced settings panel where users will be able to edit attributes of the subtitling, such as reading speed, length of subtitles, and synchronization preferences. That would be a way to make advanced AI capabilities available and accessible to the novice and experienced user alike.

Table 1. Customization Features of TransVisio

SETTING CATEGORY	OPTIONS AVAILABLE	DESCRIPTION
Target Language	Arabic, English, Spanish, German, French, Chinese (Simplified), Chinese (Traditional), Hindi, Japanese, Russian	Indicate the language to which translation is to be made.
Content Genre	Drama, Comedy, Action/Adventure, Documentary, News, Educational, Sci-Fi/Fantasy, Animation	Set the thematic style that will affect the tone and particularities of language applied in the content.

Target Demographic	Children (3-12), Teenagers (13-17), Young Adults (18-30), Adults (31-64), Seniors (65+), General Audience	Guide language complexity and cultural references and content appropriateness.
Language Register	Formal, Neutral, Informal, Colloquial	Apply the formality of language in translations which can affect the accessibility and appropriateness to the readers.
Localization Approach	Source-Oriented, Target-Oriented, Neutral, Transcreation	Set the closeness of translation to source or the adaptive tendency to the culture of the target.
Cultural Adaptation	Minimal, Moderate, Extensive, Explicative	Modify the level of cultural context and references integrated into translation.
Subtitle Length	Short (max 37 chars/line), Medium (max 42 chars/line), Long (max 50 chars/line)	Controls how much text appears on screen, influencing readability and timing.
Reading Speed	Slow (12 cps), Medium (15 cps), Fast (17 cps), Very Fast (20 cps)	Manages the speed at which subtitles flash up and disappear, influencing viewers' comprehension.
Speaker Identification	None, Dashes for Dialogue, Color-Coded, Speaker Names	Identify the identity of speakers within the content; therefore providing clarity to the viewer.
Sound Effects	None, Minimal (key sounds only), Descriptive (all relevant sounds), Enhanced (sounds and music)	Explains how sound effects and music cues are described in the subtitle-mostly for accessibility.

Table 1 details the different areas that will be customizable in TransVisio. The first option provided to users is the choice of language. Other choices related to linguistic variety include genre, which influences thematic style; language register, which relates to formality; and target demographic, in which language complexity is catered to the age of the audience. Other customization settings relate to culture. For example, localization approach relates to the degree of closeness to the source culture or adaptation to the target culture. Cultural adaptation controls the level of cultural references in the target text. Subtitle length and reading speed are customizable aspects that relate to the onscreen display of captions. Finally, TransVisio also offers options related to non-speech items, such as speaker identification and sound effects.

4. Practical Implications

AI-embedded subtitles like SubGPT and Matesub are driving evolution changes that influence media localization and workflows of professional translation. Far-reaching changes involve much greater automation, efficiency, and precision in handling linguistic nuances, hence allowing for increased turnaround speed and quality consistency of subtitles. Supportively, there is evidence by Soe et al. (2021) that semi-automated systems can increase productivity without sacrificing quality. In such situations, human oversight proves even more fundamental for a guarantee that subtitled texts attain both linguistic

and cultural appropriateness, refining AI-generated content in a genuinely effective manner to meet audience expectations.

In addition, this view can also be corroborated with respect to the studies concerning subtitling workflows by Tardel (2023), who discloses the importance of customizable production models, which enable the flexible integration of AI into different subtitling process stages depending on the requirements and complexity of the project. TransVisio is also no exception in such philosophy by underpinning human expertise with AI tools, making sure that the professional subtitlers retain the creative ownership while simultaneously leveraging machine efficiency and human intuition. The immediate effect of this kind of attitude is to make professional translators start working from AI-generated drafts while retaining just the critical aspects of post-editing and cultural adaptation in such a way as to be sure that the final output will correspond to the required tone and cultural context.

Going a step further, the workflow model from Tardel illustrates that to understand roles about the production, spotters, translators, and post-editors are involved, according to their contribution and impact on the quality of the subtitles. Besides, such roles are integrated by TransVisio, as it provides flexibility in users regarding workflow selection in choosing what kind of option would best suit the project objectives between quality control and speedy delivery. For instance, Malakul and Park (2023) conclude that subtitled systems can help make learning experiences even better by presenting complex information in a simpler light, thus lessening the mental strains for learners. Improved accessibility results in greater learner satisfaction, especially as compared to situations where subtitles have been manually edited or when no subtitles at all are available.

Furthermore, incorporating these AI-based tools into education allows students to learn more about state-of-the-art subtitling practices, which join technical automation with simple editing skills. As stated by Çetiner and İşisağ (2019), post-editing is necessary in an era where machine translation is prevalent while also showcasing little promise regarding quality. Additionally, it was found that providing post-editing classes improves students' capabilities and attitudes toward the practice. It is in this controlled environment that students are able to acquire hands-on experience with real-life examples of how to manage AI-generated subtitles. This helps them to make their decisions more critical while developing problem-solving capabilities related to any complex situation. It is in this way that the integration of technology, coupled with the contribution of human expertise, prepares the students not only for the changing dimensions of industry standards but also for the technical proficiency and cultural awareness of translation professionals.

Amini et al. (2024) refer to the recent significant AI translation trends that touch upon advances in neural and hybrid machine translation models. In this way, such developments will further enhance the efficiency but also the cultural appropriateness of the translations produced in AI, which therefore can feel more relatable to the audience. Moreover, such improvements make the requirement of appropriately balancing the capacity of AI with human intervention, so that subtitles are not only correct but also retain the necessary subtleties of the culture. This is where subtitled AI needs to be clear and apt for the requirements of the audience.

TransVisio serves as an invaluable resource for translation students by bridging the gap between the technical aspects of AI subtitling and creative, interpretive tasks so valued in high-quality translation. Besides that, TransVisio allows students to experiment with AI-assisted subtitling under controlled conditions. This controlled environment allows students to research how various subtitling configurations will determine viewers' comprehension and knowledge of cultural accuracy. Such work provides practical knowledge about what the variables are that have to do with subtitling. It is only through hands-on experiences like these that students will be able to develop appropriate skills in order

to handle the complications of everyday translation and localization. It ensures they become proficient in both technical processes and the understanding of the cultural complexities involved in making effective and meaningful subtitles.

Besides, the adaptability inherent in the tool is such that users, from professionals down to students, can benefit from customized workflows, thereby easing their work to tackle specific project needs or learning objectives. With TransVisio, one obtains a complete feeling for the interaction between AI-driven automation and human expertise, be it a tool for creating subtitles or for learning. It also allows the ability to toggle between different levels of automation and manual input, which is so important for flexibility both in educational and professional contexts. This provides users with the capability to craft experiences targeted at optimizing outcomes for varied audience needs.

5. Ethical Considerations

At the artifact level, ethical analysis focuses on the specific features of the technology that could raise moral concerns. For instance, Mihalache (2024) points out unintended consequences like the introduction of biases or ethical issues related to content manipulation. These challenges must be carefully managed to avoid negative impacts. TransVisio's design prioritizes features that facilitate transparency and allow users to maintain control over the subtitling process, minimizing the risk of misuse or bias.

A vital sector in MT is thus bias mitigation that is critical to foster fairness as well as bypass various discriminatory outcomes (Akinrinola et al., 2024; Mensah, 2023). The approaches toward it would include the implementation of Explainable AI, open sharing of data, and ethics-based AI frameworks in place (Akinrinola et al., 2024). Another important ethical issue in AI deployment is related to the protection of privacy (Floridi, 2024; Huriye, 2023). Transparency, accountability, and fairness are part of TransVisio's approach in warranting AI output to be ethically sound and ensuring protection for user rights.

The ethical considerations of using AI in subtitling also revolve around transparency, accountability, and cultural sensitivity (Huriye, 2023). Transparency is important for understanding AI decision-making processes, while accountability ensures responsible implementation (Mensah, 2023). Abu-Rayyash and Al-Ramadan (2024) emphasize the importance of transparency in the application of AI within translation processes, ensuring that end users are clearly informed about when and how AI systems are involved. TransVisio has performed well in this by providing complete clarity on AI involvement to maintain trust among creators, translators, and readers.

In evaluating AI-enhanced subtitling tools, Mihalache (2024) applies the Anticipatory Technology Ethics (ATE) framework, which categorizes ethical concerns into three levels: technology, artifact, and application (Brey, 2012). The same goes with technological levels, where another critical ethical issue has to do with the accuracy and reliability of AI in handling complex or ambiguous language, a variety of accents, and background noise. Mihalache (2024) concludes that inaccuracies in subtitles may lead to important misrepresentations of content that might cause harm or otherwise confuse viewers of such programs, particularly in contexts related to cultural sensitivity. For this reason, human oversight becomes very instrumental in preventing these risks and ensuring the subtitling outputs are correct and contextually appropriate.

Secondly, Ali (2024) maintains that there is a dire need to deal with biases that might be inherently part of AI systems, more so when cultural delicacies are an issue, not to mention dialectical variations. Stereotyping could easily occur through the perpetuation of biases in the AI-generated translations, which therefore calls for careful human oversight to deliver ethical AI output. Accountability too is an important aspect. Considering that the quality of AI-generated subtitles may vary, TransVisio embeds

user-friendly post-editing features to ensure that subtitles not only are linguistically correct but also relevant for the culture. An approach like that would be a big contribution to helping professional translators go over and refine AI-generated content to avoid unintentional misrepresentations.

Ethical reflection, at the level of the artifact itself, focuses on the particular traits of the technology that may raise issues of moral significance. In this sense, Mihalache (2024) identifies unintended consequences related to biases in introduction or ethical problems associated with content manipulation. Certainly, such challenges can be mediated carefully in order not to yield negative results. The design principles of TransVisio mainly concern features which favor transparency and enable users to retain control over the subtitling process to mitigate risks of misuse or bias.

Ethical analysis at the application level contemplates the usage of AI-enhanced subtitling. It weighs use within moral reason for better accessibility of people with disabilities against misuse of the technology for deceptive purposes (Mihalache, 2024). TransVisio allows human translators to give input about cultural settings; thus, they are capable of collaborating with AI to obtain content that is both culturally resonant and ethically sound. Ali (2024) further underlines the importance of data privacy and the reduction of bias in treating sensitive information. The ethical framework that would be developed by TransVisio would make sure respect for users' rights and interests is considered, foster fairness, and nurture inclusion in all audiovisual translation processes.

In a nutshell, ethical considerations of development and deployment for TransVisio balance increased efficiency and accessibility with responsibilities relating to cultural sensitivity, transparency, and accountability. Anticipatory practice will go a long way toward ensuring potential ethical issues are dealt with in advance that will support the integration of this tool into varied social and cultural contexts while mitigating adverse impacts.

6. Future Directions

TransVisio is a fundamental development for AI-powered subtitling, but it opens the door to even more innovative approaches to be developed. Future research will be very important in further integrating real-time machine learning algorithms that directly learn from user feedback while continuously improving the accuracy and adaptation of subtitles in context. Other key areas for improvement are increased treatment of cultural references, idiomatic expressions, and dialectal variability, which in the end will make the subtitles more intuitive and relevant in cross-cultural contexts. As Seyidov (2024) mentions, advanced adaptive learning will contribute much to the system's growing proficiency in subtitling genre-specific content, reducing dependence on extensive human intervention.

Future work, with respect to the research conducted by Soudi et al. (2024), would delve deeper into how, in real time, subtitles can be generated in cooperation between LLMs and Encoder-Decoder models in live broadcast applications. Their research finds that their combination has strong benefits with respect to both speed and accuracy; this suggests that an important direction for improvement of capabilities in live subtitling is available here. Moreover, widening the language capabilities within TransVisio is a key priority, especially for the less-resourced languages and dialects. Joint initiatives of linguists might supplement the language database even more, and as such increase the cultural particularity and precision of the system.

One very promising line of research is the integration of TransVisio with live broadcasting infrastructures and support for real-time subtitling. Such systems will pose unique challenges in terms of latency and synchronization. Again, edge computing combined with optimized network infrastructures may help to reduce latency further and improve the viewing experience for live content viewers. Because it does the processing closer to the origin of the data and does not rely on faraway cloud servers, edge

computing reduces the latency associated with generating subtitles. This would be of particular benefit in live events for news broadcasts and sports, where even the slightest latency in subtitles tends to disappoint audiences.

Future iterations of TransVisio could also benefit from more advanced user-feedback mechanisms designed to actively engage professional subtitlers in refining AI outputs. Tools that track and learn from user corrections over time could create a more interactive and adaptive learning environment. As highlighted by Abu-Rayyash and Al-Ramadan (2024), involving human subtitlers as co-creators rather than passive editors fosters a sense of agency and ensures that AI tools evolve in ways that align closely with professional standards and expectations.

Finally, with advanced features to support concurrent editing, AI and human editors could collaborate in a much more cooperative way that will substantially raise productivity and the quality of the output. The collaboration model would be much like many successful document editing platforms that encourage real-time collaboration. This would translate into adding an extra dimension to teamwork, making the subtitling process ever more efficient and enhancing the quality of the end product when applying these methods to audiovisual translation.

Funding

This research was supported by the Abdul Hameed Shoman Foundation (Grant Reference: 230800351) through the project "TransVisio: Automating Subtitle Translation – A Leap Towards Improved Audiovisual Content Accessibility"

Acknowledgment

We thank Abdul Hameed Shoman Foundation for funding this project. Also, special thanks to Ali F. Al-Ramadan for leading the development of TransVisio.

References

Abu-Rayyash, H., & Al-Ramadan, A. (2024). SubGPT: An Automated Tool for Subtitle-Template Translation. *International Journal of Linguistics, Literature and Translation*, *7*(6), 126–134. https://doi.org/10.32996/ijllt.2024.7.6.14

Akasheh, W. M., Haider, A. S., Al-Saideen, B., & Sahari, Y. (2024). Artificial intelligence-generated Arabic subtitles: insights from Veed. io's automatic speech recognition system of Jordanian Arabic. *Texto Livre, 17*, e46952-e46952. doi:10.1590/1983-3652.2024.46952.

Akinrinola, O., Okoye, C. C., Ofodile, O. C., & Ugochukwu, C. E. (2024). Navigating and reviewing ethical dilemmas in AI development: Strategies for transparency, fairness, and accountability. *GSC Advanced Research and Reviews*, *18*(3), 050–058.

Al-Abbas, L. S., & Haider, A. S. (2021). Using Modern Standard Arabic in subtitling Egyptian Comedy Movies for the Deaf/Hard of Hearing. *Cogent Arts & Humanities, 8*(1), 1-16. doi:10.1080/23311983.2021.1993597

Al-Salman, S., & Haider, A. S. (2024). Assessing the accuracy of MT and AI tools in translating humanities or social sciences Arabic research titles into English: Evidence from Google Translate, Gemini, and ChatGPT *International Journal of Data and Network Science, 8*(4), 2483-2498. doi:https://doi.org/10.5267/j.ijdns.2024.5.009

AlBkowr, N., & Haider, A. S. (2023). Subtitling for people with hearing impairments in the Arab world context: The case of the Blue Elephant 2 movie. *Online Journal of Communication and Media Technologies, 13*(4), e202347. doi:https://doi.org/10.30935/ojcmt/13602

Ali, D. H. A. (2024). Artificial Intelligence In Translation Studies: Benefits, Challenges, And Future Directions. مجلة القاهرة للعلوم, *24*(275). https://mrk.journals.ekb.eg/article_381236.html

Álvarez, A., Arzelus, H., & Etchegoyhen, T. (2014). *Towards customized automatic segmentation of subtitles*. Paper presented at the Advances in Speech and Language Technologies for Iberian Languages: Second International Conference, IberSPEECH 2014, Las Palmas de Gran Canaria, Spain, November 19-21, 2014. Proceedings.

Amin, A. A., Glasser, A., Kushalnagar, R., Vogler, C., & Huenerfauth, M. (2021). Preferences of Deaf or Hard of Hearing Users for Live-TV Caption Appearance. In M. Antona & C. Stephanidis (Eds.), *Universal Access in Human-Computer Interaction. Access to Media, Learning and Assistive Environments* (Vol. 12769, pp. 189–201). Springer International Publishing. https://doi.org/10.1007/978-3-030-78095-1_15

Amini, M., Ravindran, L., & Lee, K.-F. (2024). Implications of using AI in Translation Studies: Trends, Challenges, and Future Direction. *Asian Journal of Research in Education and Social Sciences*, *6*(1), 740–754.

Arroyo Chavez, M., Thompson, B., Feanny, M., Alabi, K., Kim, M., Ming, L., Glasser, A., Kushalnagar, R., & Vogler, C. (2024). Customization of Closed Captions via Large Language Models. In K. Miesenberger, P. Peňáz, & M. Kobayashi (Eds.), *Computers Helping People with Special Needs* (Vol. 14751, pp. 50–58). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-62849-8_7

Brey, P. A. (2012). Anticipatory ethics for emerging technologies. *NanoEthics*, *6*(1), 1–13.

Briva-Iglesias, V., O'Brien, S., & Cowan, B. R. (2023). The impact of traditional and interactive post-editing on machine translation user experience, quality, and productivity. *Translation, Cognition & Behavior, 6*(1), 60-86.

Çetiner, C., & İşisağ, K. (2019). Undergraduate level translation students' attitudes towards machine translation post-editing training. *International Journal of Languages' Education and Teaching,*

7(1), 110-120.

Dankers, V., Lucas, C. G., & Titov, I. (2022). *Can Transformer be Too Compositional? Analysing Idiom Processing in Neural Machine Translation* (arXiv:2205.15301). arXiv. http://arxiv.org/abs/2205.15301

Du, J., & Lu, J. (2024). *A Comparative Study on the Translation Quality between Human and Machine-Generated Subtitles*. Paper presented at the 2024 6th International Conference on Natural Language Processing (ICNLP).

Farghal, M., & Haider, A. S. (2024). Translating classical Arabic verse: human translation vs. AI large language models (Gemini and ChatGPT). *Cogent Social Sciences, 10*(1), 2410998. doi:https://doi.org/10.1080/23311886.2024.2410998

Floridi, L. (2024). The Ethics of Artificial Intelligence: Exacerbated problems, renewed problems, unprecedented problems-Introduction to the Special Issue of the American Philosophical Quarterly dedicated to The Ethics of AI. *Available at SSRN*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4801799

Gorman, B. M., Crabb, M., & Armstrong, M. (2021). Adaptive Subtitles: Preferences and Trade-Offs in Real-Time Media Adaption. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–11. https://doi.org/10.1145/3411764.3445509

Huriye, A. Z. (2023). The ethics of artificial intelligence: Examining the ethical considerations surrounding the development and use of AI. *American Journal of Technology*, *2*(1), 37–44.

Jia, Y., & Sun, S. (2023). Man or machine? Comparing the difficulty of human translation versus neural machine translation post-editing. *Perspectives, 31*(5), 950-968.

Malakul, S., & Park, I. (2023). The effects of using an auto-subtitle system in educational videos to facilitate learning for secondary school students: Learning comprehension, cognitive load, and satisfaction. *Smart Learning Environments*, *10*(1), 4. https://doi.org/10.1186/s40561-023-00224-2

Mensah, G. B. (2023). Artificial intelligence and ethics: A comprehensive review of bias mitigation, transparency, and accountability in AI Systems. *Preprint, November*, *10*. https://www.researchgate.net/profile/George-Benneh-Mensah/publication/375744287_Artificial_Intelligence_and_Ethics_A_Comprehensive_Review_of_Bias_Mitigation_Transparency_and_Accountability_in_AI_Systems/links/656c8e46b86a1d521b2e2a16/Artificial-Intelligence-and-Ethics-A-Comprehensive-Review-of-Bias-Mitigation-Transparency-and-Accountability-in-AI-Systems.pdf

Mihalache, I. (2024). Ethical Analysis of AI-Enhanced Audiovisual Translation Tools. *2024 IEEE 48th Annual Computers, Software, and Applications Conference (COMPSAC)*, 1558–1559. https://ieeexplore.ieee.org/abstract/document/10633381/

Perone, S. (2023). Matesub: The Translated Subtitling Tool at the IWSLT2023 Subtitling task. *Proceedings of the 20th International Conference on Spoken Language Translation (IWSLT 2023)*, 461–464. https://aclanthology.org/2023.iwslt-1.45/

Saed, H., Haider, A. S., & Tair, S. A. (2023). Gender, sexual, ethnic, color and Disability-related Epithets and labels across languages: Evidence from Arabic subtitling of English Movies and series. *Eurasian Journal of Applied Linguistics, 9*(1), 213-225. doi:http://dx.doi.org/10.32601/ejal.901018

Samha, F., Haider, A. S., & Hussein, R. F. (2023). Address forms in Egyptian vernacular and their English equivalence: A translation-oriented study. *Ampersand, 10*, 1-12. doi:https://doi.org/10.1016/j.amper.2023.100117

Seyidov, R. (2024). Importance and Role of Artificial Intelligence for Arabic Translators. *International Journal for Arabic Linguistics & Literature Studies (JALLS)*, *6*(2). https://www.refaad.com/Files/JALLS/JALS-6-2-4.pdf

Shuhaiber, R., & Haider, A. S. (2023). Strategies of Subtitling Egyptian Idiomatic Expressions and Proverbs into English. *Language Value, 16*(2), 60-99. doi:https://doi.org/10.58256/sptxps94

Soe, T. H., Guribye, F., & Slavkovik, M. (2021). Evaluating AI assisted subtitling. *ACM International Conference on Interactive Media Experiences*,96–107.https://doi.org/10.1145/3452918.3458792

Soudi, A., Hannani, M., Van Laerhoven, K., & Avramidis, E. (2024). Exploring the Potential of Large Language Models in Adaptive Machine Translation for Generic Text and Subtitles. *Proceedings of the 17th Workshop on Building and Using Comparable Corpora (BUCC) @ LREC-COLING 2024*, 51–58.

Tardel, A. (2023). A proposed workflow model for researching production processes in subtitling. *Trans-Kom*, *16*(1), 140–173.

Yamada, M. (2023). Optimizing Machine Translation through Prompt Engineering: An Investigation into ChatGPT's Customizability. *arXiv preprint arXiv:2308.01391*.