Co-funded by the Erasmus+ Programme of the European Union

AI Avatars and Automated Translation in E-Learning: Lessons from the International Educational Project IIDEMIC (Inclusive and Innovative Digital Education for Migrant Community in Kenya and Somalia)

Only a few years ago – and certainly before the COVID-19 pandemic – developing a professional online course built around video required not only subject-matter and pedagogical expertise but also access to an entire production infrastructure: a recording studio, appropriate equipment, and a team responsible for editing and post-production. The person delivering the content – in instructional design terminology, the subject matter expert (SME) – had to command the material thoroughly and, in addition, was expected to perform on camera: to present well, articulate ideas fluently, manage stress and read from a teleprompter. Where a course was intended for audiences speaking different languages, further challenges arose: translating the materials, recording new language versions, producing foreign-language subtitles and synchronising them with the footage, or preparing audio description in another language. Today, a considerable share of these tasks can be performed with the support of artificial intelligence tools. Does this mean, however, that creating multilingual educational materials in e-learning format has become genuinely simple?

The experience of the international Erasmus+ project IIDEMIC (Inclusive and Innovative Digital Education for Migrant Community in Kenya and Somalia) shows that AI can substantially facilitate the production of instructional content, yet the effectiveness of such solutions does not depend on technological capability alone. On the contrary, learners' needs, cultural context and human involvement in the design and verification of materials remain equally important – if not still the most important – factors.

Not So Fast, Not So Easy

Within the IIDEMIC project, ten asynchronous online courses on inclusive education were developed for teachers working with refugee and migrant communities. The materials were authored by experts representing diverse educational settings and the project partner countries: Poland, Türkiye, Kenya and Somalia.

As early as the initial course design and work-planning stage, it became evident that the production model for asynchronous courses developed over years of experience at WSB Merito University in Poznań would not suit every partner institution. One obstacle was that not all content authors had access to a professional recording studio. Nor did all experts have

experience of performing on camera[1]. Recording a lecture involves far more than reading out a prepared text. It requires ease of expression, an appropriate speaking pace, sustained eye contact with the lens, and suitable facial expression and body language. For those who had never worked in such conditions, lecture recordings could become a source of stress and a reason to look for ways to withdraw from them. Moreover, every slip of the tongue or error may mean that a take has to be repeated. Video recording for online courses also distributes working time differently: recordings for a single multi-module course often have to be broken down into working hours and scheduled across several days. Maintaining concentration on constructing one's delivery is an enormous intellectual effort and, not infrequently, a physical one as well. Several hours a day standing in front of a camera, shots from different angles, constant monitoring of one's body (Am I swaying? Am I shifting from foot to foot?), squinting at the camera when the teleprompter text is hard to read – these are just some of the pitfalls awaiting educators with little on-camera experience.

So That the Camera Is Not a Barrier

To overcome reluctance towards video lecture recordings, the project team offered some of the authors an alternative: digital avatars generated by artificial intelligence from the SMEs' own likenesses. The HeyGen software was used for this purpose. It is not, however, a perfect remedy for the difficulties associated with reluctance to record lectures in person. Why? HeyGen requires the explicit consent of the person whose image and voice are used to create the avatar. Each author gave this consent by submitting a three-minute sample recording, which served as the basis for building their avatar. Consenting to the use of one's likeness also entails other risks and responsibilities that the person lending their image must accept and agree to. The terms of service for HeyGen's paid plans[2] explicitly state that users grant HeyGen a broad, royalty-free, worldwide and irrevocable licence to use uploaded content, including for the development of its services and AI models. HeyGen further declares that Enterprise customers' data are not used for model training by default; on lower-tier plans, however, whether such use occurs depends on the applicable data-processing rules and consents, and users may lodge an objection. In practice, this means that one may request the deletion of an avatar, but deleting the material does not necessarily reverse all prior data-processing operations. It is therefore an alternative – but one that comes at a price, and not merely the

[1] Problems related to access to a stable internet connection and to devices enabling asynchronous learning are discussed in my article "Designing E-Learning under Conditions of Limited Internet Access: Lessons from the IIDEMIC Project", published at: https://iidemic.org/?p=10768 (accessed 13 October 2026).

[2] See: https://www.heygen.com/terms (accessed 13 October 2026).

cost of a subscription. Ultimately, eight of the ten authors opted for a digital avatar. Two experts from Poland recorded their presentations in the traditional way, in a studio.



Source: screen from course in Swahili language

Three Minutes of Footage Instead of a Day-Long Session

So how does it work? The technology analyses any recording supplied by the expert and subsequently enables the generation of further videos in which the avatar delivers a prepared script. The avatars were created at the turn of 2025/2026. At that time, approximately three minutes of sample footage sufficed to produce a digital copy of the expert. Authors did not need to deliver the actual lecture or memorise a script; they could speak freely on any topic, providing the system with the material needed to replicate their appearance, facial expressions and manner of speaking. Before recording, the experts received clear guidelines from the production team on how to prepare such a sample correctly. Technical details were essential: the recording had to be free of background noise and echo; the sample could not be filmed in an empty room or in poor lighting; and the footage had to be properly framed, well lit and – importantly – not handheld, unlike typical social media videos. The submitted samples were then reviewed in individual meetings, and corrections were made (in most cases these concerned fine-tuning of colour). To finalise the avatar, the expert lending their likeness had to authenticate themselves in the tool; the system verified that this was the same person who appeared in the uploaded recordings. Finally, the author was asked to read aloud a statement of consent to the use of their image during authentication. Without completing this process, the tool does not allow the expert's avatar to be used. From that point on, the work was straightforward. The project team uploaded the finished scripts to HeyGen, occasionally made

minor adjustments – such as adding a pause or correcting the pronunciation of individual words – and the videos were ready. The advantages of this solution included the relatively low cost of creating the avatars and the ability to generate new, revised videos quickly, without authors having to return to the video studio.

This workflow enabled the project team to separate two tasks that are closely intertwined in traditional video production: developing valuable instructional content and presenting it on camera. Authors could concentrate on what they knew best – developing the substantive content of the courses – while the technology took over part of the work involved in producing the recordings. It did not replace the experts; rather, it allowed them to take part in the project without having to overcome additional technical and organisational barriers.

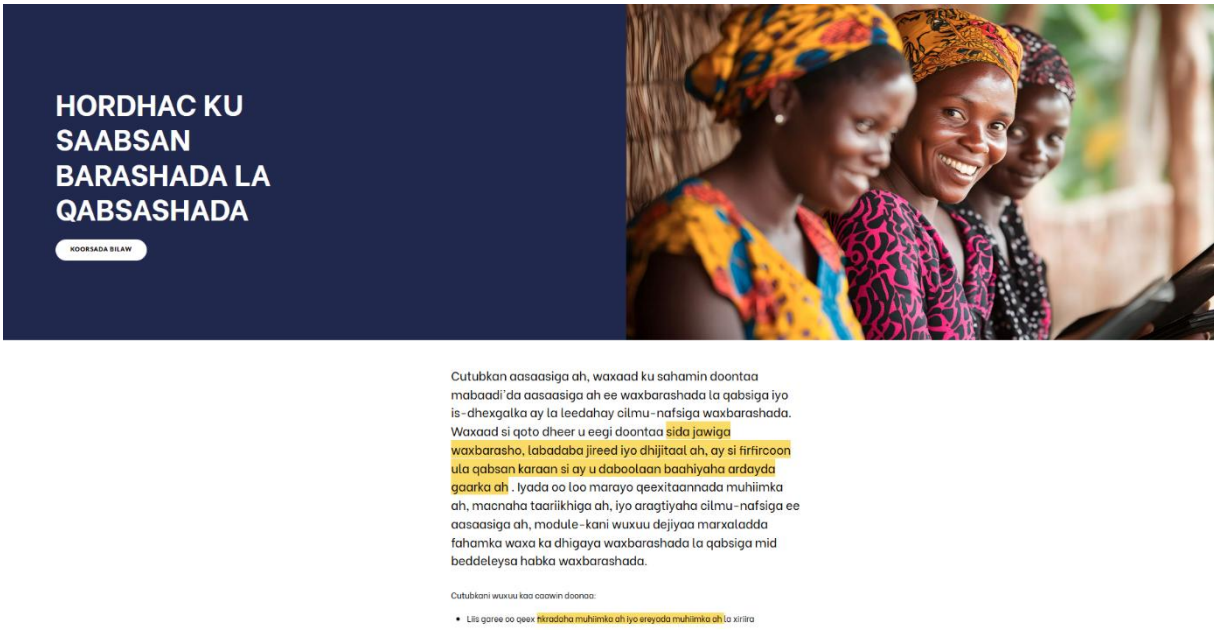
Is This a Solution for Everyone?

The benefits of using AI avatars are not limited to streamlining the production of educational materials. Research suggests that digital presenters may also positively influence how content is received and how engaged course participants become. In one study involving 150 e-learning participants, materials delivered by virtual speakers were rated positively, particularly when provided in the learners' native language. In a pilot study with 15 engineering students, avatars endowed with specific personality traits were observed to encourage longer interaction, with some sessions lasting as long as 40 minutes. These findings point to the technology's potential for fostering engagement, although they do not yet establish its long-term impact on learning outcomes. It should also be borne in mind that the appeal of a technology is not synonymous with its pedagogical effectiveness. A realistic avatar may capture attention, but it may equally become a distraction if its presence does not serve specific learning objectives. The decision to use a digital presenter should therefore stem from learners' needs and the nature of the content, rather than solely from what the tool makes possible. In the IIDEMIC project, the use of AI-generated avatars proved to be a win-win solution.

One Course, Several Languages

A further design challenge was the multilingual nature of the audiences for the asynchronous courses. The materials prepared for the IIDEMIC project were intended for teachers from different language communities. Although English was the main working language of the project, it was not the most natural medium of learning for all participants. Translation into local languages – Swahili and Somali – proved indispensable. To accomplish this, the team turned to another edtech tool, Smartcat, a platform that, in its own marketing, promises AI-powered translation of content into any language. Translation into the languages of the migrant

communities covered every course component: syllabuses, the reading handbooks, the exercises in the Articulate Rise multimedia handbooks, the final tests and the video subtitles. The only element that remained unchanged was the avatars' spoken narration: owing to project constraints, it was decided not to produce avatars speaking Swahili or Somali. Course users can, however, freely switch subtitles between English, Swahili and Somali. Automating translation made it possible to accelerate the preparation of successive language versions without involving the authors in further recordings. This did not mean, however, that machine translations were treated as finished teaching materials. The project team, represented by the Kenyan and Somali universities, asked members of migrant communities who speak these languages to check whether the automatically translated content sounded natural, was comprehensible and conveyed the intended meaning. Some of the translations were also reviewed by academics from Kenya and Somalia. Interestingly, the Somali reviewers noted that the automatic translations were not always flawless (Somali being a highly complex language that is difficult to translate) but were nonetheless adequate. In addition, each author checked the translation of their own course into their native language. This was a particularly important stage: linguistic correctness does not always equate to communicative effectiveness. A translation may be grammatically correct yet contain phrasing that is unnatural, overly formal or culturally inappropriate. In education – especially education addressing such sensitive issues as inclusion and work with diverse communities – such differences can affect how the message is understood.



Source: screen from course in Somali language

The importance of this approach is confirmed by research on the use of language models to translate educational materials. An analysis of machine translations of the Force Concept Inventory, a test used to assess understanding of basic physics concepts, showed that AI models can correctly render the general meaning of a text while making errors in specialised terminology, numbering or item structure. In assessment materials, even a slight shift in meaning can alter what a given item actually measures. The problem therefore concerns not only linguistic accuracy but also the validity of the educational process. At the same time, advances in translation tools open up new possibilities for personalising content. Research on such solutions indicates that readers may prefer different ways of phrasing a translation, depending on their needs and the intended register of communication. A single, stylistically uniform translation for all users is thus not always the best solution. In international education, the ability to adapt language not only to learners' proficiency but also to their social and cultural context becomes particularly significant.

The IIDEMIC experience illustrates this relationship well. Having the translations verified by members of the language communities made it possible to view the teaching materials from the perspective of their future users, not only that of the authors and course designers. This was particularly important for content on inclusive education, where the way phenomena, social groups and educational needs are named matters not only linguistically but, above all, socially.

Accessibility Is More Than a Change of Language

The use of AI also brought benefits in terms of accessibility. Based on the content of the recordings, the project team was able to prepare text scripts to accompany the videos. This gives learners an additional way of engaging with the materials: they can not only watch and listen to a lecture but also read its content, return to a selected passage, or rely on the text when video playback is difficult. Indeed, the design of the IIDEMIC courses had to take into account some learners' limited access to stable internet and electricity, as well as the fact that many teachers access educational materials primarily on their phones. A text alternative to a recording does not solve every accessibility problem, but it increases the flexibility with which a course can be used. This reflects an important principle of universal design: not assuming that every learner will learn in the same way and under identical conditions. It is worth distinguishing, however, between the digital accessibility of materials and their actual pedagogical effectiveness. Simply offering content in several languages, using an avatar or providing a transcript does not in itself guarantee that learners will better understand the issues presented. Research on educational applications of artificial intelligence yields promising

results regarding user satisfaction and engagement, but there is still insufficient evidence to determine conclusively the long-term impact of such solutions on learning outcomes. Another problem is the uneven maturity of the available technologies. Not all tools are developed with the principles of universal design in mind, and solutions that work well in marketing or business communication do not automatically meet the needs of education. Particularly in settings with limited internet access, a visually attractive format may prove less useful than simpler, technically lightweight text-based material. The choice of technology should therefore be preceded by an analysis of the conditions in which learners will actually use the course. From this perspective, artificial intelligence can support the implementation of universal design principles in teaching and learning, but it does not replace the design process itself. It is the instructional designer who must decide which forms of content presentation are needed, which barriers should be reduced, and how to ensure that learners can use the materials on a range of devices and in diverse socio-technological conditions.

Technology Does Not Replace Responsibility

The IIDEMIC experience reveals not only the possibilities but also the limitations of technology and of automating instructional work. An avatar can convincingly deliver an expert's message, but it cannot by itself evaluate its pedagogical value. A translation system can process text quickly, but it will not always correctly interpret specialised terminology, cultural context or nuances of meaning. Ethical considerations also remain important. Creating a digital likeness of an author requires their informed consent and a clear specification of how the recordings and generated materials will be used. Learners should also be made aware that they are watching an AI avatar rather than a conventional recording of a person. Nor should the risk of over-reliance on technology be underestimated. A realistic image and fluent delivery may create an impression of professionalism, but they are no guarantee of substantive accuracy. Responsible use of AI therefore requires maintaining substantive, linguistic and pedagogical oversight at every stage of materials development – and that oversight should always be exercised by people. In the IIDEMIC project, this was a significant responsibility shared by the entire project team.

The protection of personal image is, however, only one of the ethical challenges associated with the use of AI in education. Research on machine translation also points to the risk of perpetuating biases and stereotypes present in the data used to train the models. An algorithm may not only mistranslate a particular concept but also strip a statement of its social or cultural context. In international educational projects, particularly those aimed at migrant and refugee

communities, such distortions may lead to the unintended exclusion of learners rather than supporting their participation in education. Another significant risk is excessive trust in automatically generated content. A realistic avatar, fluent delivery and a professional-looking recording can create an impression of credibility even when the material presented contains errors. Likewise, a translation that sounds natural does not necessarily convey the meaning of the original faithfully. For this reason, specialised content, assessment materials and socially sensitive topics required particular scrutiny from the project team. Responsible use of AI thus requires not only authors' informed consent and the protection of their data but also transparency towards learners. Course participants should know that they are watching a digital representation of an expert and that the content was prepared with the aid of automated tools. Equally important is a clear designation of who is responsible for the accuracy of the materials and who approves their final version before publication. Automating the production process must not lead to a diffusion of responsibility for its outcome.

Removing Barriers, Not Replacing People

In the IIDEMIC project, artificial intelligence was not an end in itself. It became a tool for accomplishing tasks that would have been considerably more difficult to organise under a traditional production model. Avatars made it possible to involve experts without arranging professional recording sessions. Automated translation facilitated the preparation of multilingual materials, and text scripts broadened the ways in which the educational content could be used. At the same time, it was the members of the project teams at the partner universities who remained responsible for the quality of the outcomes: authors for the subject-matter content, instructional designers for the way it was conveyed, and representatives of the language communities for assessing the comprehensibility of the translations. This is perhaps the most important lesson for those designing digital education. The value of AI lies not in eliminating people from the course development process, but in removing the obstacles that limit their participation in education. In international projects implemented under diverse technological, linguistic and cultural conditions, such an approach may be of particular significance[3].

Agnieszka Goncerzewicz
WSB Merito University in Poznań
Marcin Czyżuk
WSB Merito University in Toruń

[3] Editorial revisions and the English translation were prepared with the assistance of an artificial intelligence tool: Claude Opus 5.5 (accessed 26 October 2026).

References:

S. Miranda and R. Vegliante, “Leveraging AI-Generated Virtual Speakers to Enhance Multilingual e-Learning Experiences,” Dec. 2024, doi: 10.20944/preprints202412.1037.v1.

J. Miquel-Vergés, “El doblaje automático de vídeos educativos universitarios mediante inteligencia artificial con Heygen,” *European public & social innovation review*, vol. 9, pp. 1–21, July 2024, doi: 10.31637/epsir-2024-358.

T. Kim, D. P. Agarwal, J. Ackerman, and M. Saha, “Steering AI-Driven Personalization of Scientific Text for General Audiences,” Nov. 2024, doi: 10.48550/arxiv.2411.09969.

G. Majgaard and M. S. Andersen, “From Code to Character: Investigating Personality in Generative AI-Driven Educational Avatars,” vol. 19, no. 2, pp. 543–551, Sept. 2025, doi: 10.34190/ecgbl.19.2.4002.

M. Babayeva, J. C. Dunlap, M. Snětinová, and R. Widenhorn, “Translating the Force Concept Inventory in the age of AI,” Aug. 2025, doi: 10.21203/rs.3.rs-7410179/v1.

M. Fink, S. A. Robinson, and Bernhard, “AI-Based Avatars Are Changing the Way We Learn and Teach: Benefits and Challenges,” Mar. 2024, doi: 10.35542/osf.io/jt83m.

C. Vallis, S. Wilson, D. Gozman, and J. Buchanan, “Student Perceptions of AI-Generated Avatars in Teaching Business Ethics: We Might not be Impressed,” *Postdigital Science and Education*, June 2023, doi: 10.1007/s42438-023-00407-7.