TransPipe

A Pipeline for Automated Transcription and Translation of Videos

Joseph Bethge[1], Sebastian Serth[1], Thomas Staubitz[1], Tobias Wuttke[2], Oliver Nordemann[2], Partha-Pratim Das[2], and Christoph Meinel[1]

Hasso Plattner Institute
Potsdam, Germany
[1]firstname.lastname@hpi.de
[2]firstname.lastname@student.hpi.de

Online learning environments, such as Massive Open Online Courses (MOOCs), often rely on videos as a major component to convey knowledge. However, these videos exclude potential participants who do not understand the lecturer's language, regardless of whether that is due to language unfamiliarity or aural handicaps. Subtitles and/or interactive transcripts solve this issue, ease navigation based on the content, and enable indexing and retrieval by search engines. Although there are several automated speech-to-text converters and translation tools, their quality varies and the process of integrating them can be quite tedious. Thus, in practice, many videos on MOOC platforms only receive subtitles after the course is already finished (if at all) due to a lack of resources. This work describes an approach to tackle this issue by providing a dedicated tool, which is closing this gap between MOOC platforms and transcription and translation tools and offering a simple workflow that can easily be handled by users with a less technical background. The proposed method is designed and evaluated by qualitative interviews with three major MOOC providers.

1 Introduction

Massive Open Online Courses (MOOCs) have set out to disrupt, revolutionize, and democratize the educational world starting about ten years ago. However, neither have predictions come true that universities will become obsolete within a few years nor has everybody in the world the same access to high-quality education. This leads many critics to the conclusion that MOOCs have failed altogether. On the other hand, the number of MOOC platforms, course providers, and participants all over the world is still growing. Particularly, the current pandemic boosted the participant numbers in so far unknown heights. We are one of Europe's oldest

and largest MOOC platform providers. We are operating the platform for our own purposes but are also providing further instances of the platform for customers and partners often with different needs and purposes.

The common factor that unites all platform instances independent of their purpose and motivation: they heavily rely on videos as a major component to convey knowledge. Videos as learning material, however, exclude aurally handicapped participants and those who do not understand the lecturer's language. Both groups can be easily included by adding subtitles and/or interactive transcripts to the videos. Interactive transcripts, furthermore, come with additional advantages. They provide an additional option to navigate within a video and they enable search engines to include the videos' contents into their search results.

Creating multilingual subtitles for MOOC videos requires two steps: 1) the spoken words within a video need to be transformed to written text, so-called transcripts, and 2) these transcripts need to be translated into the required target languages. Tool support for both processes exists for several years. Speech-to-text software extracts the spoken word into text, enriched with timestamps so that these texts can be synchronously displayed with the video stream. This generally works well for lecturers who speak English fairly with neither dialect nor accent. The stronger the dialect or accent, the more effort is required to manually correct these texts. If the original language is not English, the quality of automated transcripts is often too low to be used, thus requiring manual transcription. Furthermore, the transcript has to be translated if subtitles in other languages are desired. In this case, a high quality of the transcript is even more important as errors would be transferred to translations. Similar to transcriptions, quality control and manual fixes are often necessary, even though the quality of AI-based translation has improved significantly within the last years – at least for some languages. Even assuming the best-case scenario, creating multilingual subtitles so far is still a quite time-consuming and tedious process. First, the video needs to be retrieved from the video hosting platform and passed to the transcription tool. The transcription itself takes some time, and once it's done, the transcript needs to be quality-checked and possibly improved. The transcript then needs to be downloaded and uploaded to the translation tool, processed, downloaded again, quality-checked, improved, and uploaded again to the platform.

These separate steps often have to be completed by different people. Many of them need access to different tools, which creates new challenges: security risks, possible loss of data, additional costs, etc. Furthermore, the transcripts and translations often have a particular format and character encoding, which tends to be messed up when people with a non-technical background are involved, e.g. by opening text files in word processors rather than text editors. In the worst-case scenario, several more manual steps are involved, such as a manual peer review. Finally, the workflow to create these subtitles greatly differs between the different

platform instances, as the quality requirements to publish these subtitles differ a lot from partner to partner. For example, some partners directly publish machine-generated transcripts and ask the participants to contribute by improving them, others would not publish any text until it is perfect in all supported languages. Complicating matters, the time between video recording and course start is often very limited. Studio capacities and particularly the available time of the lecturers often result in just-in-time video production. Therefore, a very streamlined transcription and translation workflow is required to enable adding subtitles before a course starts. Currently, many videos on our partner platforms are still missing subtitles or the subtitles are only added once the course is completed.

We have, therefore, set out to create an application that closes this gap and provides an end-to-end workflow to add subtitled videos to the courses of our platform partners. Hence, it has to connect the dots between the different transcription and translation services. Our application needs to be configurable to allow combinations of different service providers. Finally, it has to fulfill the different requirements and reflect the different workflows of our platform partners. The goal is to simplify the transcription and translation process to allow the platform providers to add transcripts and translations to more videos, to speed-up the process so that transcripts and translations can be added earlier than previously so that subtitles are available before the course has started. This enables the platform partners to offer a better platform experience to their participants with fewer accessibility issues and the option to address an internationally broader audience.

We conducted extensive interviews with our platform partners to learn about their particular needs, requirements, and workflows. We talked to our service providers to learn about their future plans and ways to integrate their services into our application. From here, we will refer to this application by the name "TransPipe", which stands for transcription and translation pipeline.

2 Related Work

Automated transcription and translation to enhance online learning videos, with subtitles or interactive transcripts has been an elementary part of several research projects in the recent past. A system to create such automated transcripts was designed as a part of the *transLectures* project (2011–2014). The project focused on improving the transcription quality by adapting from the speakers' previous lectures and adding the speakers' slides as additional input. The system supported human supervision by highlighting those parts where the system had little confidence in its results. The results were field-tested on the *VideoLectures.net* platform [5, 8].

One of transLectures' partners, the Universidad Politécnica de Valencia (UPV) took the results further and formed the *Machine Learning and Language Processing Group (MLLP)*. MLLP claims that their automated speech recognition (ASR) and machine translation (MT) systems are among the most competitive at an international level [9]. MLLP also operates *poliTrans* a commercial tool allowing customers to create automated transcripts and translations for their videos[1]. The tool also features a user interface to manually correct the machine's mistakes. The MLLP Group contributed to several other projects and published several papers in the area of automated speech recognition and machine translation [11, 3, 10, 7, 16].

One of these projects, the *European Multiple MOOC Aggregator (EMMA)* (2014–2016) aimed to showcase excellent innovative teaching methodologies on a large-scale multi-lingual MOOC platform. Since 2019, the platform receives additional funding to host (mostly English-only) courses in the content of another project: *ASSET (A Holistic and Scalable Solution for Research, Innovation and Education in Energy Transition)* [4]. To access the older multi-lingual courses, you have to create an account and sign in. Still only very few courses are really multi-lingual. Often only the written texts are translated. Video subtitles are rare and often only exist in one language. Sometimes the videos are completely missing when the user switches to another language. These observations are based on random samples on the platform[2]. There might be better examples, but they indicate that multilinguality was at least not achieved on a broader level. Further random samples on VideoLectures.net revealed that there as well subtitled videos are rather the exception than the rule.

The *TraMOOC* project (2015–2018) aimed to reduce language barriers in MOOCs by developing a high-quality translation tool for all types of texts used in MOOCs including video subtitles. The machine translation engine was provided mainly by the University of Edinburgh [15, 12, 14, 13]. During this project, openHPI served as one of the field test platforms. The openHPI platform's video player was extended to support the display of subtitles and interactive transcripts in multiple languages. The translation engine is marketed by one of the members of the consortium[3] under the name of *translexy*[4]. Although the transcription and translation quality were generally perceived quite well by course participants and instructors, until today, the field test is the only course that offers subtitles in all supported languages.

The main reason for this is the missing support for integrated end-to-end-workflows including automated transcription, manual quality assurance, automated translation, further quality insurance, and an option to upload the subtitles

---

[1] https://politrans.upv.es/
[2] https://platform.europeanmoocs.eu/
[3] Knowledge4All Foundation [5]
[4] https://www.k4all.org/project/translexy/

to the platform. The transcripts generated by *poliTrans* are all lowercase and lack punctuation. *TraMOOC* translates the subtitles time-stamped chunk by chunk, these chunks are often not even complete sentences and, therefore, deteriorate the translation quality. One solution to these issues is the method developed by Che et al. to add punctuation and upper-case letters to the transcripts, and to merge the chunks into a proper text that can be translated in context [2].

Transcription and translation engines are available by many academic and commercial providers. Next to the already mentioned *poliTrans* and *translexy*, the commercial players such as *Google Translate*, *Amazon's Media Insight Engine* (AWS MIE), or *DeepL* are well known. Each of the existing tools does a good job in what it does. However, none of the tools does the job completely on its own. Often several tools need to be combined. Some platform partners might be contractually bound to certain providers or company guidelines prohibit a platform partner to work with a certain provider. The workflows also might differ depending on the course's original language: English generally works quite well, non-native speakers, however, might have a bad effect on the transcript quality due to their accents. Automated transcripts for German videos are much worse[5]. In total, this results in a heavy workload for the teaching teams, which only a few can afford. We, therefore, offer this now as a service to the teaching teams. Still, the process is time-consuming, error-prone, and tedious. To address this, we decided to work on a configurable, flexible tool to simplify and support this process.

3 Requirements Analysis

We analyzed the requirements of three of our platform partners with the user-centered Design Thinking approach [6]. Consistent with this approach, we started by interviewing the process owners of our platform partners, who are the potential users of our application. We gathered detailed insights about their individual workflows to create, translate, and provide subtitles for videos in MOOCs. More specifically, we asked about the different stages in their process, the providers used during each stage, and how they integrated different tools.

We summarized the challenges the users faced during their workflow and our derived requirements in Business Process Management Diagrams. Based on these findings, we developed strategies to solve the various challenges the users faced in their current workflows. In line with the Design Thinking approach, we im-

---

[5]We quick-checked for literature to confirm this statement but only found more general papers addressing issues with under-resourced languages, e.g. [1]. For now, this statement is, therefore, only based on our experience concerning the quality of automatically generated German transcripts.

plemented wireframe prototypes to verify our ideas in a subsequent round of interviews with our users. We incorporated the feedback during the second round of interviews for the final prototype described in section 4. In the following, we show the results of both rounds of interviews with our project partners.

3.1 Partner A

In this section, we summarize our findings from interviews with the digital innovation department of our Partner A (see Figure 1). Their workflow for video transcription and translation has three kinds of participants: administrators responsible for coordination, volunteers for processing transcripts and translations, and linguistic experts responsible for quality assurance.
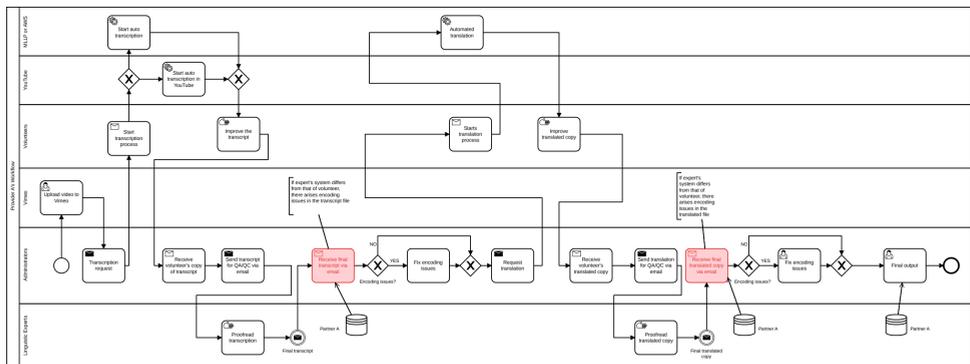
Figure 1: Process model diagram of our partner A's workflow for the transcription and translation task. Best viewed digitally with zoom.

The process starts with the administrator uploading the given video to the Vimeo platform. They then delegate the transcription task to one of their volunteers. The volunteer uses the video transcription service from either MLLP or YouTube to automatically generate the initial transcript from the video. Then, they download the machine-generated transcript to manually improve and proof-read it. Upon completion, they send the transcript to their respective supervisors. Afterward, experts are commissioned to perform quality assurance on the volunteer's copy of the transcript. The finalized copy is sent back to the administration. Next, the final transcription is sent to volunteers for translation into the required target languages. They use the *DeepL* service to translate the transcript automatically and later improve it where necessary. Subsequently, the translated copy is sent

back to the supervisor who then contacts the linguistic experts to perform quality assurance. Finally, the experts send the translations to the administrators, who publish both the original transcription and the desired translations on the platform.

During this process, there is no shared environment for editing subtitle files (neither for transcripts nor their translations). Instead, emails are used for communication and file exchange between the participants. Since files are edited locally, different operating systems used in different regions sometimes result in encoding errors between the different participants. If such errors occur, the transcripts and/or translations need further examination and reformatting before proceeding to the next step. This issue can cause delays which could easily be solved with a common editing platform. Furthermore, the most recent version of subtitles is not always clear to administrators, which sometimes leads to more manual effort to determine the correct files.

As described above, our Partner A's process has some points that lead to delays and manual effort. In summary, we identified the following process issues:

1. status of the videos is unclear to the participants and needs to be tracked manually by the supervisors

2. the most recent version of subtitles is unclear due to a lack of a versioning system for subtitles

3. several process participants collaborate in the creation process of subtitles but there is no role-based system to control the process

4. files are shared via email causing delays and accessibility issues

5. local file editing frequently leads to encoding and formatting issues in the transcripts (and translations)

3.2 Partner B

In a similar manner, we present our findings from interviews with our Partner B in the following (see Figure 2).

Their process starts with student assistants searching for the given video on Vimeo. The video is then temporarily set to be publicly available so that the MLLP service can be used to generate the transcription automatically. After the transcription is finished, they reset the video in Vimeo to private and download the generated subtitles. After manually improving the transcript, they send it to the production team for publishing. Furthermore, the student assistants use the *DeepL* translation service to improve the MLLP translations and hand them over to the production team for publishing. An intermediate manual improvement of
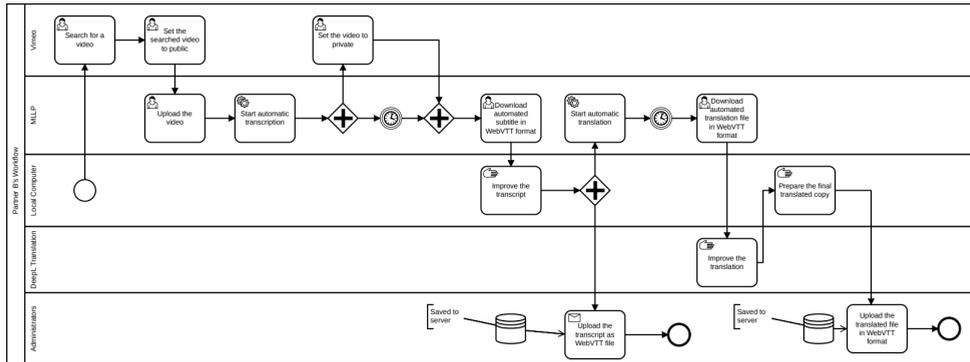
Figure 2: Process model diagram of our partner B's workflow for the transcription and translation task. Best viewed digitally with zoom.

the translations sometimes is done between this, but it lacks accuracy or is entirely skipped if the student assistants do not speak the target language.

Similar to the process of Partner A (but less frequent), encoding issues can occur and delay the process (in subsection 3.1). There is no quality assurance by linguistic experts, but an external agency is commissioned to create manually made, high-quality transcripts for videos in German (due to the lack of automatic tools with reasonable quality as mentioned in the previous section). The quality of these transcripts is superior to automatically generated ones used for videos recorded in other languages, e.g. English.

In summary, we identified the following process issues:

1. no common file management or sharing system creates communication overhead

2. manual work to use multiple, different applications for transcription and translation

3. several process participants collaborate in the creation process of subtitles but there is no role-based system to control the process

4. reverting to previous versions of subtitles is impossible due to a lack of a versioning system for subtitles

5. local file editing frequently leads to encoding and formatting issues in the transcripts (and translations)

3.3 Partner C

Finally, we interviewed the process owners of our Partner C (see Figure 3).
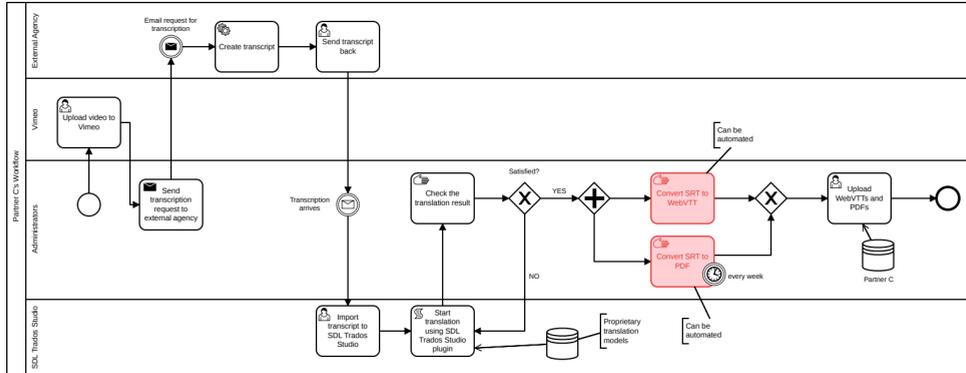
Figure 3: Process model diagram of partner C's workflow for the transcription and translation task. Best viewed digitally with zoom.

The process begins with the administrators uploading a given video to Vimeo. They employ a third-party agency to create a high-quality transcript. Once the file is ready, the agency hands it over to the administrators where the transcript is imported into a proprietary software for high-quality and efficient translation. Upon completion, the translated files are inspected for accuracy. If the administrators are satisfied, they convert the transcript and its translated copies into the SRT format. The SRT files are converted into PDF documents to generate weekly PDFs for each course week. Finally, those subtitle files are published. During these steps, files are transferred via email between the agency and the administrators.

In summary, we identified the following issues:

1. considerable manual effort is needed to generate a weekly course overview and the subtitle summary and for file conversion

2. files are shared via email which needs manual intervention and causes delays

3. subtitles need to be uploaded to the platform manually

3.4 Derivation of Requirements

The process model diagrams in the previous sections show the complex and manual processes involved in creating transcripts and translations for courses on our partners' platforms. While the partners share a common goal of adding subtitles to their videos within MOOCs, the involved parties (volunteers, student assistants, and external agencies) and the quality requirements (ranging from no improvements to a review by linguistic experts), differ. Using this understanding of the process, we formulate the following eight requirements needed for our solution to serve the needs of the end-users (we briefly refer to the corresponding issues of our partners with the identifier followed by the issue number, e.g. B2 is the second issue of partner B):

R1: Common Environment. Requirement for a common environment where all project participants collaborate as a basis for requirements R2, R3, and R6.

R2: Status Overview. Requirement for an accessible overview page to list all the available courses and their status (A1).

R3: Shared Editing Platform. Requirement for a single shared platform on which the subtitle files (for transcriptions and translations) can be edited (A4, A5, B1, B5, C2).

R4: Versioning. Requirement for a versioning system of subtitles (A2, B4).

R5: Roles and Permissions. Requirement for a platform where different users can have different permissions to edit subtitles (A3, B3).

R6: Service Integration. Requirement to integrate different external services into one platform to remove the necessity to switch between different platforms (B2).

R7: Summary Generation. Requirement to provide an automatically generated summary of the subtitles for a course week for download (C1).

R8: Subtitle Publishing. Requirement for an option to publish the edited subtitles directly on the corresponding MOOC platform (C3).

We created several wireframe prototypes for each platform partner and validated the prototypes with the individual process owners. Afterward, we merged the prototypes into one generalized version by mapping the single process steps of the first prototypes into one process that fits all identified requirements, which is described in the following section.

4 TransPipe Prototype

This section describes how we translated the derived requirements from the wireframes into a prototype which we call "TransPipe". *TransPipe* integrates into our MOOC system with external machine learning systems for transcription and translation. This results in a single shared platform to manage the generation, editing, and quality assurance of subtitles for all of our MOOC platform partners. At the same time, *TransPipe* must be highly configurable to choose the preferred transcription and translation services and should provide extensibility for new services.

The generalized pipeline for a given video consists of the following steps: (1) generating an initial automatic transcription, (2) improving the transcription (optionally followed by quality assurance), (3) generating an initial automatic translation of subtitles for each target language, and (4) improving the translated subtitles (again with optional quality assurance). The integration into our MOOC system allows us to trigger this process directly by fetching a video from the corresponding MOOC platform and publishing the final transcribed or translated subtitle files.

The prototype mainly consists of the following views (web pages): a course overview, a course details page, a course section overview, and a video details page showing the available subtitles for the video, to fulfill the requirements R1 (Common Environment), R2 (Status Overview), and R4 (Versioning). Furthermore, a page is needed to modify and review transcripts, and another page to modify translations for steps (2) and (4) of the pipeline, fulfilling requirement R3 (Shared Editing Platform). These pages consist of a side-by-side view of the input video and a text editor for the transcript (see Figure 4). For the translation tasks, two text areas are shown, which contain the original transcription and the current translation, respectively.

We integrated the external machine learning services used by our partners for the automatic generation steps (1) and (3) of the pipeline to fulfill requirement R6 (Service Integration). Each step is configurable, i.e. different services can be offered to the users of our platform partners, who can then choose the service to use. Automatic status updates on these externally running services ensure the status is up to date in our system and can be viewed by the users. Furthermore, finished subtitle files (either containing a transcription or its translations) are downloaded to our system once ready and can subsequently be edited. Even though the initial prototype included mock-up buttons for requirements R7 (Summary Generation) and R8 (Subtitle Publishing), the actual implementation was missing at the time of evaluation. Similarly, the requirement R5 (Roles and Permissions) was mocked for the evaluation interviews.

Figure 4: The prototype page displaying subtitle details coupled with a text editor and video player for checking and editing transcripts

5 Evaluation

Throughout the development of *TransPipe*, we regularly collected feedback from the different MOOC providers involved in the project. Our goal was to create a first working version to showcase the possibilities of an automated pipeline for transcription and translation. We based our prototypical implementation on the unified design described in section 4 and used it to evaluate the intended workflow with our partners in qualitative, unstructured interviews. While the MOOC providers valued the possibility to streamline the rather manual process of creating subtitles by replacing most steps with *TransPipe*, they provided additional feedback on the desired integration of their workflows.

Most importantly, the MOOC providers wished to have user roles better represented in *TransPipe*. Most have specialized staff for different languages or optionally request an additional review from language experts before subtitles are published. Due to the prototypical implementation, these roles were not yet supported nor was the status of subtitles fully reflected in the system. Supervisors also criticized the missing visualization of progress and a lack of required information to decide which courses or videos need attention for a timely release.

The technical integration of the transcription and translation services along with the MOOC platform fulfilled most of the requirements outlined in section 3. A new technical challenge discovered with the working prototype was our partners'

request to combine multiple services for a single video. When the initial translation was performed with another provider, a slightly modified input is required for the services. Similarly, we did not add dedicated support for custom terminology to *TransPipe* so far. While the providers generally use all previously created terms, users can add new entries for the custom terminology only through the website of the service. We plan to add both features upon request in a later stage.

5.1 Workflow Improvements

The most criticized part of the prototype was the user navigation which resembled the initial wireframes but did not sufficiently support our partners' workflow. Based on the collected feedback and the initial prototype, we iteratively created a workflow-based design. It defines seven different states for each subtitle file ranging from *not available* through *auto-generated*, *manually edited*, *in review*, *changes requested*, *approved* to *published*. If a language expert decides not to approve a subtitle but rather rejects it, the requested changes can be described with a textual comment. Each workflow step requires a user to have an appropriate role that administrators can assign to users at any time to reflect customer demands.

A new course overview page lists all videos of a MOOC ordered by their position with the lecturer's oral language and the desired translation languages in separate columns of a table. Each cell indicates the current state for the given language and video combination and thus allows checking the progress of a course with a glimpse. Representatives from our partners stated that the chosen visualization eliminates the need for another project tracking solution or additional synchronization across the involved staff. If desired, users can execute bulk actions on several languages or videos simultaneously.

We also tightened the coupling to the MOOC platform by making *TransPipe* the default editor for subtitle changes. A deep linking between the course or specific videos allows easier editing of subtitle files in *TransPipe's* two-pane view with built-in versioning. By disabling editing features on the MOOC platforms, we prevent possible merge conflicts.

5.2 Future Work

With the current implementation of *TransPipe*, we were able to resolve many of the initial shortcomings and show the feasibility of building a pipeline. While our application includes the required steps to automate the processes as much as possible, the correction of machine-created subtitles is still unsophisticated with a simple text editor. Hence, we plan to integrate an advanced subtitle editor that optionally synchronizes with the video and allows a more user-friendly change in

timestamps as we have for now. Doing so will also require additional syntax checks and editing support for the used WebVTT file format to prevent illegal changes.

Some transcription and translation services, such as AWS Media Insights Engine, provide advanced machine learning features, e.g. maintaining a custom terminology to improve future subtitle generation based on previous corrections. In the current status of *TransPipe*, no information about edits is shared with the original service rendering this feature almost useless without manual maintenance of the custom terminology. However, the open architecture and the availability of dedicated APIs by the providers make it possible to integrate support for these as well. Depending on the specific service, we will investigate how to offer support for advanced machine learning features within *TransPipe* or how to extend our application with other providers.

In addition to publishing transcripts and translation for videos, course administrators can use the generated subtitles to create a written summary of the video lecture. Hence, MOOC participants can use them when watching the video is not possible or desired and thus value them in various situations. By adding a PDF export functionality of the subtitles (without time information) to *TransPipe*, we can fulfill the learners' request for written information without increasing the workload for teaching teams or platform providers. Therefore, we plan to add an export job and extend the MOOC platform integration so that *TransPipe* users can automatically attach the PDFs to the course. Finally, we will revalidate our assumptions and the prototype with our partners and adjust the system where necessary.

6 Conclusion

In this work, we discussed a solution on how to reduce the effort of creating subtitles in the hopes of improving the accessibility of MOOC course videos for foreign or handicapped participants. Having subtitles in different languages also leads to other beneficial effects, e.g. indexing and retrieval by search engines and improved navigation within videos.

To reach an effective solution, we conducted interviews with three major MOOC providers, which are platform partners using our MOOC system. We presented our findings of their particular needs, requirements, and workflows and designed our *TransPipe* prototype to solve the challenges found. The main aspects tackled are communication issues and delays, reducing the need for manual effort and intervention in the process, and providing a sufficiently general, yet configurable platform to be suitable for all MOOC providers.

Although, we were not able to quantitatively determine the effectiveness, e.g. the share of videos with correctly (translated) subtitles, the evaluation through

qualitative interviews showed the potential and usefulness of *TransPipe*. However, they also revealed opportunities for future improvements, and we expect this process of gathering feedback and adapting the current application to the future plans of our platform partners to be continued in the future.

References

[1] L. Besacier, E. Barnard, A. Karpov, and T. Schultz. "Automatic speech recognition for under-resourced languages: A survey". In: *Speech Communication* 56 (2014), pages 85–100. ISSN: 0167-6393. DOI: https://doi.org/10.1016/j.specom.2013.07.008.

[2] X. Che, S. Luo, H. Yang, and C. Meinel. "Automatic Lecture Subtitle Generation and How It Helps". In: *2017 IEEE 17th International Conference on Advanced Learning Technologies (ICALT)*. 2017, pages 34–38. DOI: 10.1109/ICALT.2017.11.

[3] M. A. del-Agua, A. Giménez, N. Serrano, J. Andrés-Ferrer, J. Civera, A. Sanchis, and A. Juan. "The transLectures-UPV toolkit". In: *Proceedings of VIII Jornadas en Tecnología del Habla and IV Iberian SLTech Workshop*. IberSpeech 2014. Las Palmas de Gran Canaria, Spain, 2014.

[4] EMMA. *About*. URL: https://project.europeanmoocs.eu/about/ (last accessed 2021-01-21).

[5] Knowledge4All Foundation. *TransLectures*. URL: https://www.k4all.org/project/translectures/ (last accessed 2021-01-21).

[6] C. Meinel and J. von Thienen. "Design thinking". In: *Informatik-Spektrum* 39.4 (2016), pages 310–314.

[7] J. D. V. Miró, J. A. Silvestre-Cerdà, J. Civera, C. Turró, and A. Juan. "Efficiency and usability study of innovative computer-aided transcription strategies for video lecture repositories". In: *Speech Communication* 74 (2015), pages 65–75. ISSN: 0167-6393.

[8] MLLP. *TransLectures*. URL: https://www.mllp.upv.es/projects/translectures/ (last accessed 2021-01-21).

[9] MLLP. *TransLectures*. URL: https://www.mllp.upv.es/progress/ (last accessed 2021-01-21).

[10] A. Pérez González de Martos, J. A. Silvestre-Cerdà, J. D. Valor Miró, J. Civera, and A. Juan. "MLLP Transcription and Translation Platform". In: *Proceedings of 10th European Conference on Technology Enhanced Learning*. EC-TEL 2015. Toledo, Spain, 2015.

[11]   S. Piqueras, M. A. del-Agua, A. Giménez, J. Civera, and A. Juan. "Statistical text-to-speech synthesis of Spanish subtitles". In: *Proceedings of IberSpeech*. IberSpeech 2014. Las Palmas de Gran Canaria, Spain, 2014.

[12]   R. Sennrich and B. Haddow. *Linguistic Input Features Improve Neural Machine Translation*. 2016. arXiv: 1606.02892 [cs.CL].

[13]   R. Sennrich, B. Haddow, and A. Birch. *Edinburgh Neural Machine Translation Systems for WMT 16*. 2016. arXiv: 1606.02891 [cs.CL].

[14]   R. Sennrich, B. Haddow, and A. Birch. *Improving Neural Machine Translation Models with Monolingual Data*. 2016. arXiv: 1511.06709 [cs.CL].

[15]   R. Sennrich, B. Haddow, and A. Birch. *Neural Machine Translation of Rare Words with Subword Units*. 2016. arXiv: 1508.07909 [cs.CL].

[16]   J. D. Valor Miró, J. A. Silvestre-Cerdà, J. Civera, C. Turró, and A. Juan. "Efficient Generation of High-Quality Multilingual Subtitles for Video Lecture Repositories". In: *Proceedings of 10th European Conference on Technology Enhanced Learning*. EC-TEL 2015. Toledo, Spain, 2015, pages 485–490.