TRANS-VECTOR PLATFORM: OPTIMISED DISTRIBUTION OF VIDEO ASSETS ACROSS DIGITAL CHANNELS

R. Bocyte[1], L. Nixon[2], J. Oomen[1]

[1] Netherlands Institute for Sound and Vision, the Netherlands
[2] MODUL Technology GmbH, Austria

ABSTRACT

Digital video platforms and social media channels have opened new means for media organisations to interact with their audiences and generate more value from their video assets by republishing them online. However, the complexity associated with digital content curation, creation and distribution across these channels leaves many smaller organisations at a disadvantage. Digital content workflows require a deep understanding and constant monitoring of audiences' interests, video editing skills and ability to effectively communicate with audiences who are already overwhelmed by content choices. We introduce the Trans-Vector Platform (TVP), a modular technological infrastructure that semi-automates the distribution and optimisation of video assets at opportune moments. It demonstrates how AI-driven solutions for video analysis, topic prediction and personalisation can be combined to support a variety of content publication workflows in media organisations. We describe four front-end applications that illustrate the use of TVP components in different scenarios to support the selection, adaptation and distribution of media content optimized for different digital platforms and channels.

1. INTRODUCTION

The last three decades have brought significant disruption to the world of television, first with the emergence of the Web (1990s), then social media (2000s), then the third generation of wireless mobile telecommunications (3G) for streaming-ready mobile devices (2010s). The original model of linear TV broadcast on a single channel is evolving into nonlinear unicast on multiple digital channels. TV has become digital, personal and ubiquitous. The present decade was already forecast to see a further evolution of television in this direction, which has only been accelerated by the current COVID-19 pandemic, where consumers' time online - including time on social networks or consuming online media via web platforms and mobile apps - has increased significantly. At the same time, as demonstrated by Newman et al (7), the pandemic highlighted the threats of misinformation in the media and the lack of trusted, reliable sources of information on social media and online content sharing platforms.

Media organisations have large audiovisual collections available at their disposal that can match this demand for trustworthy content online. More importantly, digital publication channels present a unique opportunity for these organisations to generate more value out of their existing content by republishing it online and bringing it to new audiences (Bocyte and Oomen (3)). However, they need to overcome the increasingly complex challenge of attracting viewers and maintaining their engagement over time given the availability of choice and content across an abundance of digital platforms. They need to adapt their content publication strategies to audience expectations and media consumption habits -

specifically, the desire to find relevant content on the digital channels that they already use, the ability to consume media content on any digital device and the expectation for personalised experiences and interactive engagement (Westcott et al (12)). Hence it is even more important than ever before to be able to optimise video publication workflows and distribute content on the appropriate channels at optimal time - optimal in terms of achieving publication goals (reach, engagement). As media organisations emerge into a post-pandemic world where media consumption, among all other aspects of life, will have been permanently disrupted, optimised distribution of video assets across digital channels will continue to guide the digital media landscape.

The Horizon 2020 research project ReTV[1] addressed this challenge through the vision of *trans-vector publication* - the means to publish video assets across multiple digital channels (vectors) with the effort of one. The project developed technical components and services that form a trans-vector publication workflow. These different components and services have been implemented in a way to communicate with one another and compose a holistic and functional offering. Front-end Web interfaces have been built to give user access to the functionalities. The infrastructure which combines all of the components, services and interfaces into a single, modular platform is called the Trans-Vector Platform (TVP).

This paper demonstrates how video publication workflows can be optimised with the TVP. It argues that in order to create lasting impact on their audiences, media organisations need to implement data-driven strategies, monitor online trends and user engagement across platforms, and adapt their video offerings for (interactive) consumption on digital channels. The paper begins with the state of the art for digital content publication in media organisations (Section 2). Identifying the gaps, we describe the technical implementation of the TVP (Section 3) as a solution to fill those gaps with components for trend monitoring and prediction, video analysis, summarization and audience segmentation. Section 4 presents how these components are integrated into publication workflows through front-end Web interfaces to increase the visibility and competitiveness of digital media offerings, exemplified by three use cases developed with public media organisations and validated by usage tests performed with media professionals and consumers (Section 5).

2. STATE OF THE ART: DIGITAL CONTENT PUBLICATION

Media organisations employ marketing and editorial teams to decide which content from their collections is published where, when, and in what form. The importance of these content teams has been rising steadily as media consumption has shifted online and across a multitude of new digital channels. Social networks have become the biggest competition to both TV broadcast and the respective Web streams of video material. In particular, “short form” content has exploded due to the traffic that frequent publication of short clips brings to commercial platforms and the personalised curation strategies that segmented video clips permit (Wang (10), Wright (13)).

With exponentially increasing competition from both other professional producers and the global body of vloggers, editorial teams at broadcasters need to be much more agile and efficient in order to attract audiences to their video assets. Teams typically work with static content publication plans (e.g. manually created calendars with upcoming events relevant to their audiences or generic Web-based calendars) and use regular editorial meetings to discuss how the current news stories could be used as an opportunity to promote archival

[1] https://retv-project.eu/

material. In larger organisations, specialised professionals use video editing software to create different versions of the content optimised for publication on different digital platforms (different length, cropping ratio etc.). Web-based applications are used to schedule content publication in advance and they, or separate software, may also provide analytics subsequently on the publication success (traditionally measured in terms of *reach* and *engagement*). This complex workflow that requires numerous different tools and involvement of different teams prevents organisations from efficiently publishing audiovisual content, in particular older, archival collections, optimally and in time.

We argue for a need for greater automation and optimisation in this publication workflow in the fast-moving world of digital content. Typical genres of TV content - news, politics, sports - are of brief yet intense interest to audiences, yet the key to success is publishing the right piece of content at precisely the right time on the right channel. Since the inception of the ReTV project in 2018, the market for trans-vector publishing has expanded significantly. Over a dozen competitors have established themselves. Typically, they schedule text and image posts for social (re)publishing: *Later, Facelift, Tailwind, Buffer, hootsuite, Sprout Social* and *Falcon.io* are examples of well-established tools, demonstrating the viability of the market space. While some of the above tools provide monitoring capabilities (measuring audience engagement), none of them provide support for topic selection, the first and the most important step in content publication workflows. Similarly, tools such as *tellyo, grabyo, WSC Sports,* and *Wildmoka* have emerged to support video editing but they have gaps - we refer to the market need for rapid repurposing and archival video feature detection and retrieval. Overall, the existing tools focus on one specific step in the content publication workflow for social media channels only which means that media professionals are required to constantly switch between multiple applications in order to publish a single piece of content online.

3. IMPLEMENTATION: THE TRANS-VECTOR PLATFORM

The Trans-Vector Platform (TVP) is the set of components, services and interfaces which in combination provide the functionalities of trans-vector publishing to user applications (see Scharl et al (9) for a detailed technical description). Trans-vector publication is conceived as a workflow with seven steps (cf. Fig. 1 below):

- **Collect** selected content being published on digital channels;
- **Annotate** this content in terms of topics (sets of keywords and referenced entities) and use these annotations in the extraction of content metrics (e.g. relative frequency of mentions of certain topics over time on a channel);
- **Predict** what will be the future topics of interest to the audience based on these content metrics;
- **Enhance** a piece of media content in the organisation's media collection relevant to a selected topic to have the optimal form for publication on the target digital channel;
- **Schedule** the enhanced media content for publication on the target digital channel at a recommended date according to the prediction (when the topic of the content is likely to attract the most attention);
- **Publish** as organic content or paid advertising on the target digital channel;
- **Measure** the audience reach and engagement with the published content (and use these metrics to optimise the prediction for the target audience).

Figure 1: Concept of the trans-vector publishing workflow.

Responding to the needs of media professionals (Bocyte et al (5)), the TVP introduces novel capabilities that enable a full end-to-end trans-vector publication:

- Services that enable the discovery and tracking of topics in the online discourse through automatic crawling and topic-based annotation of video files, Web articles and social media (the `collect-annotate` step);
- Algorithms that can predict the future trending topics among the target audience, i.e., it is assumed content around a topic which is experiencing higher popularity at the time of posting should attract more reach/engagement (the `predict` step);
- Services that enable media owners to easily find (segments of) media assets in their archives or Media Asset Management Systems (MAMS) that relate to the chosen topic and the automated, topic-focused summarization of video and text (the `enhance` step);
- Services connected with the APIs of various digital channels for the scheduling and trans-vector publication of the prepared content (the `schedule-publish` step).

All the components of TVP communicate via APIs and can be used in different configurations to support various use cases and build different frontend applications for content repurposing. It combines functionalities in the form of components, services and interfaces (cf. Fig. 2 below) that:

- collect relevant Web and social media data via a **crawler** component that stores the resulting annotations (topic-based descriptions of the online content) in a **Metadata Repository**;
- collect the video assets of the organisation to make them available for re-purposing and re-use, combined with available audience metrics and EPG data through a **scheduler** component which queues assets for processing then stores the video file in a **Binary Repository**;

- process video assets for subsequent re-purposing and re-use by conducting **fragmentation** (into sub-shot/shot/scene), **annotation** (visual concept detection), **brand detection** (a specialization of object detection focused on brand logos) as well as **text-to-video embedding** (machine learning through a combination of visual concept features and text associated with the video - title and description - to represent both text and visual elements in the same embeddings layer of a neural network and thus be able to conduct multi-modal search where relevant videos may be found based on a textual query - even when the video lacks any textual metadata). All extracted metadata about the videos is stored in a **Video Feature Storage**;
- perform **event extraction** from online data sources such as knowledge graphs (Wikidata), iCals or through **temporal annotation** of online documents (i.e. extraction of date references inside social media postings or Web pages to associate that content with the extracted date). Extracted event knowledge is stored in a **Semantic Knowledge Base**;
- generate content-based success metrics from the large scale of collected document annotations by aggregating them in a **Metadata Search Index** (e.g. calculate frequency of mentions of a topic online over time on a certain channel) and provision these time-based success metrics to a **prediction service** to make machine learning-based forecasts for the future popularity of topics (Nixon (8)). Prediction is supported by the additional event knowledge retrieved from the Semantic Knowledge Base as a further feature for the hybrid prediction model;
- create infoviz (information visualisations) through a **visualisation engine** that takes data from the Semantic Knowledge Base and/or the Metadata Search Index and can be embedded in a front end interface, e.g. bar/line charts showing changing sentiment and emotions towards a TV program in social media (Weichselbraun et al (11)).
- perform **video adaptation and re-purposing**[2] according to the target digital channel and the content genre, with multiple pre-configurations (e.g. fast moving shot changes as an Instagram Story) as well as full manual configuration possible (Apostolidis (1));
- **summarize text** for accompanying a video posting with a focus on the optimal wording for audience attention (i.e. use of terms and topics forecast to enjoy more attention at the time of publication);
- **recommend** a video to publish according to the future trending topics of the prediction component and a matching video found through the text-to-video embedding, as well as **schedule** the prepared video for publication on the optimal date according to the prediction component.

[2] Software available at https://github.com/e-apostolidis/AC-SUM-GAN.

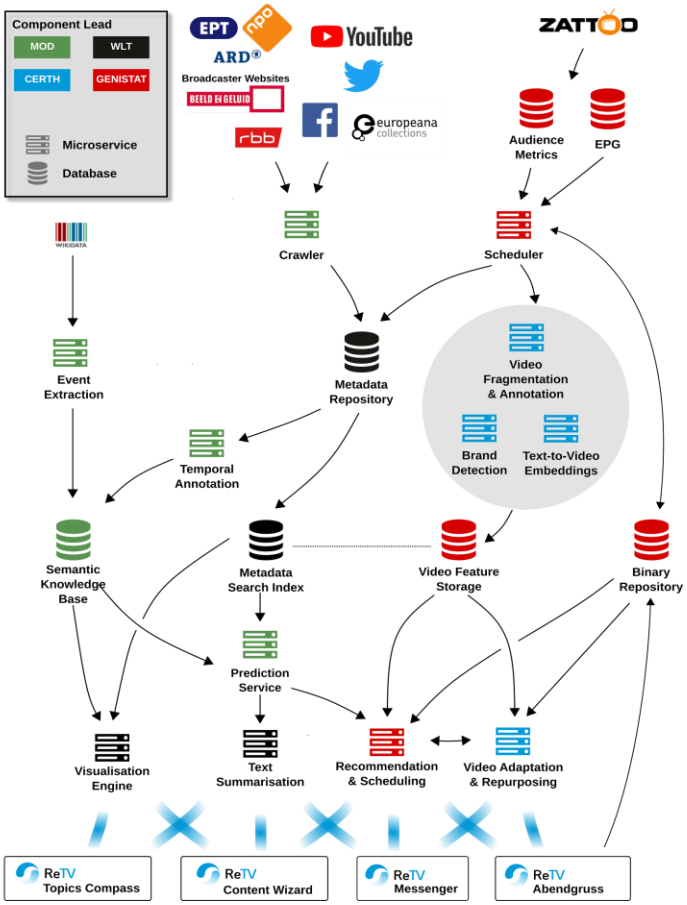

Figure 2: TVP architecture: components, services and user applications.

4. USE CASES: SCENARIOS AND TOOLS

The modular architecture of the TVP described above supports a wide range of use cases across digital publication channels. Specifically, in the ReTV project, we developed four applications:

- **Topics Compass** - a data analytics visualisation dashboard that monitors media stories as they evolve in real-time and analyses future trends that will be of interest to audiences online. The tool helps editorial teams to identify opportunities and ideas for new content creation and to understand the context of stories in the media (see Bocyte et al (4) for detailed description);
- **Content Wizard** - a social media publishing platform that acts as a single platform for media content publication across social media channels;
- **4u2 Messenger** - a tool that automatically distributes a personalised feed of video content to individual users via messaging applications;
- **4u2 Smart Speaker Skill Abendgruß** - an application for smart speaker devices that enables users to create their own videos from existing archival material using voice commands (see Zwicklbauer et al (14) for detailed description).

This section focuses on two of these prototypes, Content Wizard and the 4u2 Messenger to demonstrate the flexibility and potential of the TVP.

Content Wizard

Given the large number of social media accounts that media organisations manage, creation of social posts is one of the most demanding and time consuming tasks for editorial teams. The Content Wizard application brings together the majority of TVP services to provide a seamless, semi-automated workflow for audiovisual content publication across social media platforms. It presents a unique end-to-end solution that supports each step of the trans-vector publication workflow.

For topic selection, the user can open the Planning Calendar which displays upcoming contemporary and historic events from the TVP's Semantic Knowledge Base. In contrast to commonly used Web-based content planning calendars, the events here can be fully customised to specific user needs. To ensure meaningful results that do not overwhelm the user with too many suggestions, a template is used to define what events would be of interest to each user of the Content Wizard and their specific audience (e.g. a Berlin-based broadcaster could see historic events that happened in Berlin area, death and births of famous Berliners, etc.). Alternatively, users can select topics based on the Trending Stories feature for social listening (see Fig. 3). It monitors and visualises trends across social media platforms and Web articles to highlight stories which are likely to attract audience attention in the future. While the Planning Calendar focuses on fixed events, Trending Stories allows users to respond to stories that emerge spontaneously on professional news channels as well as topics popular amongst ordinary users. Various data visualisations are used to display the results in a way that allows users to understand the context of a story and zoom in on a particular topic.

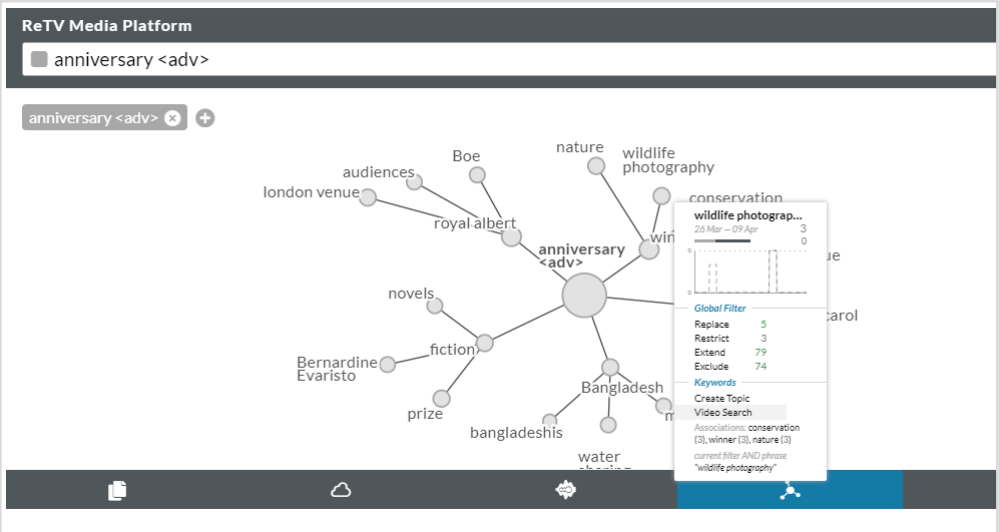

Figure 3: Content Wizard's Trending Stories feature.

To ensure that media organisations reuse and republish potentially older and more diverse content and leverage the long-tail of their collections, Content Wizard supports the text-to-video search (provided by the text-to-video embeddings component of the TVP). The search retrieves videos related to a given search query (a title of an event from the Planning Calendar or keyword(s) selected from the Trending Stories feature) based solely on the visual aspects of the videos. This is particularly useful to identify segments of high relevance when there is little metadata available.

Once media content has been selected, the video summarization feature automatically shortens a selected video into a summary of a specific length optimising it for different social

media platforms, while keeping the key parts of the full video. The editorial interface (see Fig. 4) allows users to configure video summaries according to their needs - users can choose from preset profiles which have been defined to generate summaries of optimal length for consumption on different platforms (including Twitter, Facebook Feed, Instagram Stories, TikTok), adjust cut frequency (the speed at which the shots in the summary change) and define the range from which the summary should be generated. It was important to ensure that users maintain control over the creative decisions therefore the video editor displays individual segments that the video summary is composed of, and allows users to make further edits (e.g. change the order of those segments, delete/add segments, etc.).

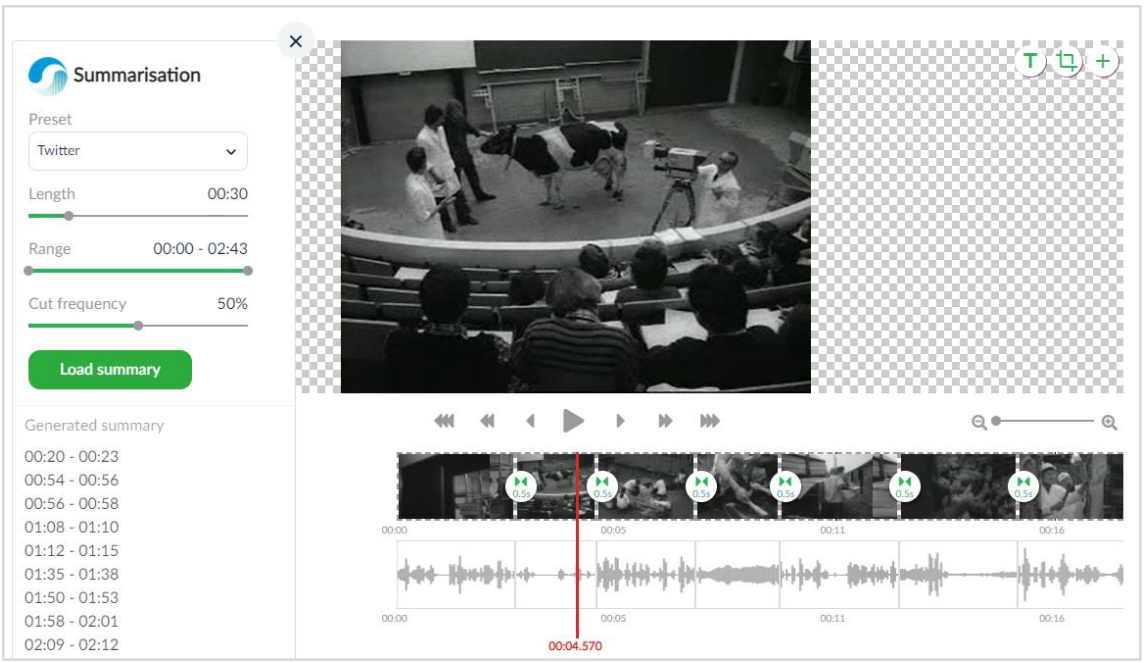

Figure 4: Content Wizard's video summarization page.

The next step in the workflow is creating a compelling text to accompany the video. Building on the social listening capabilities, the text editor (see Fig. 5) provides advanced functions to optimise the text based on real-time analysis of related texts published across digital sources. Responding to the need for short messages that immediately capture reader's attention, this feature also automatically shortens the text to a desired number of sentences.

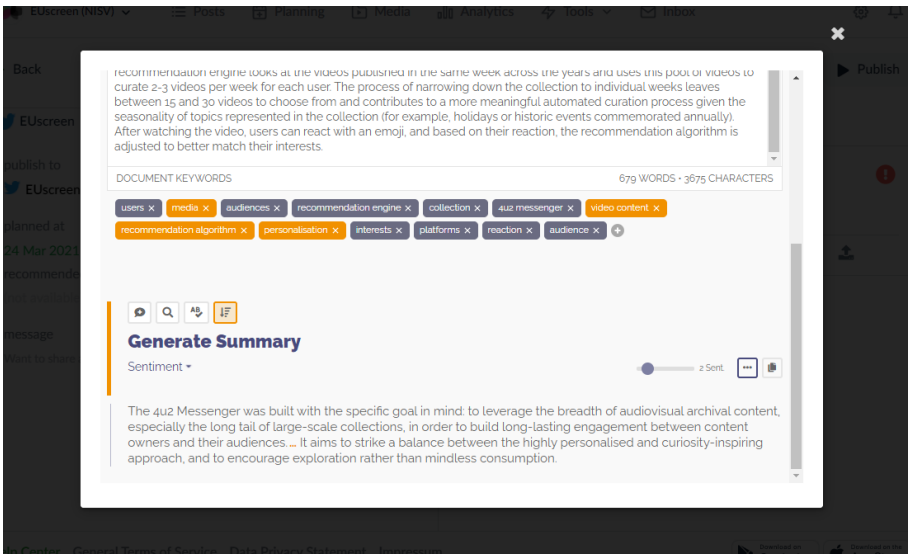

Figure 5: Content Wizard's feature for text optimisation.

The final step in the workflow concerns choosing the appropriate date to publish the social media post on the selected topic. Content Wizard provides a recommendation based on the text created in the previous step, i.e. the TVP's prediction service suggests a date over the next 30 days when the combination of keywords used in the text is most likely to be mentioned based on the analysis on past data.

4u2 Messenger

Responding to the demand for unique experiences and personalised engagement with media content, the 4u2 Messenger application was conceived to tailor broadcaster content to individual users and bring it to them via channels that audiences already use. It automatically distributes a personalised selection of videos directly to users via personal messaging applications. This allows media organisations to push their content directly into the information stream of their audiences which is likely to keep users engaged for longer. Unlike the content publication on social media platforms that address large audience segments, a messaging application establishes a connection with individual users and tailors the video offering specifically to them.

During the project, the 4u2 Messenger was implemented using the Facebook Messenger for the ReTV consortium partner Netherlands Institute for Sound and Vision.[3] Each user receives 2-3 videos each week and is invited to provide feedback using one of the emoji reactions (see Fig. 6). Based on these explicit reactions, the recommendation service is continuously adjusted to better match individual user interests. Additionally, the application monitors when each user watches the video to determine the best time to distribute messages in the future. TVP capabilities enable automation of the entire workflow - the messages are created automatically based on available metadata, video selection is based on the video analysis and text-to-video matching, the recommendation and scheduling component automatically distributes the respective message to each user.

[3] https://www.facebook.com/4u2messenger/

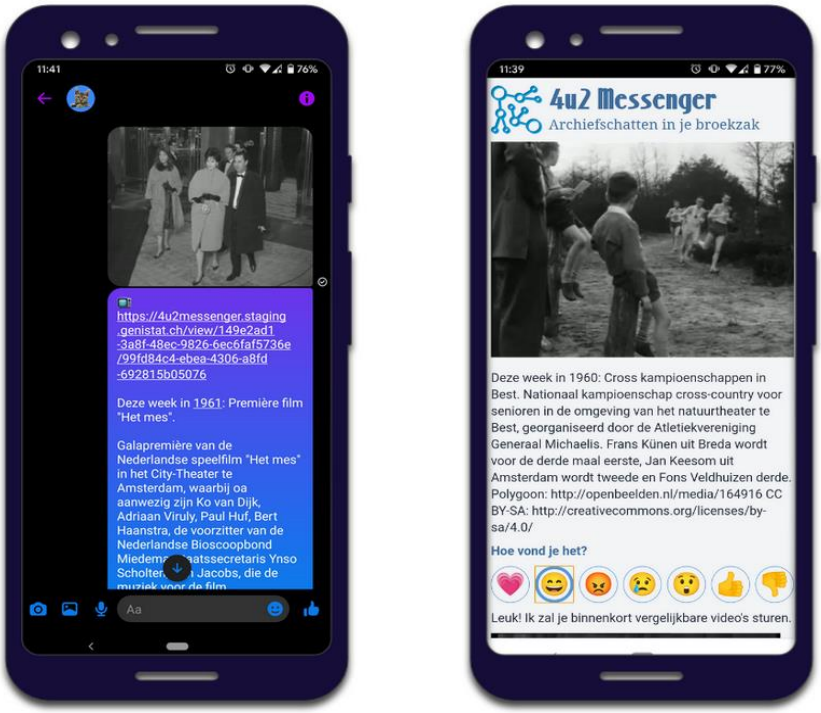

Figure 6: 4u2 Messenger implementation for Facebook Messenger.

5. EVALUATION: USER TESTS

To evaluate the usability, performance and impact of the TVP, ReTV conducted longitudinal testing with professional users and consumers for the four applications. Detailed results are reported in Bocyte et al (2) and Lamm et al (6) and the key findings and lessons learned are summarized below:

- Professional users particularly warmly welcomed solutions for topic selection as it seems to be the most time demanding task in their workflows. In particular, participants highlighted the potential of predictive capabilities, a feature that is not part of their currently used tools but that would significantly support content planning efforts. For this task, users prefer narrative-based visualisations that give them immediate ideas.
- The customizability of data sources that ReTV supports was highlighted as a unique feature. In particular, it helps to filter out data that overshadows search results (for example, users chose to exclude dominant topics (e.g. COVID-19) or particular data sources (e.g. English-language news)). However, it should be noted that such customisation is time consuming and requires a close collaboration with the users and a deep understanding of their audiences and content publication strategies. The introduction of configuration templates in Content Wizard has helped to improve and scale this process.
- Transparency and control in application interfaces are key to professional users when working with AI-powered tools. Users did not trust the recommendations and suggestions provided by Content Wizard unless they understood why those recommendations were made. In this context, the video editing interface was rated positively as it gave users control over the creative decisions and allowed them to experiment with different parameters of the video summarization algorithm.
- In terms of personalisation, users indicated that it was often not clear that their experience was different from that of other users. More research is needed in this direction to make personalisation mechanisms more transparent. It should also be noted that users

did not want to receive content that was personalised to perfectly match their interests - on the contrary, all participants noted that the element of surprise when receiving an unexpected piece of content kept them engaged for longer.

- Most importantly, while the ReTV project started with the vision to automate and optimise content publication workflows as much as possible, engagement with different stakeholders revealed that this is not always the desired scenario. While AI-driven tools can produce content that matches the high-quality standards of professional organisations, media professionals prefer to perform some of these tasks and decisions themselves - even if this requires significant time and effort. Therefore, when developing similar tools for the media sector, it is important to ensure that automation enhances rather than replaces manual efforts and creativity.

6. CONCLUSION

The TVP is ready to be deployed to customers who seek professional-grade tools for live and archive video discovery and re-publishing across multiple media vectors with the effort of one - the USP of ReTV. A typical challenge for results from a funded research project like ReTV is their continued maintenance after the funding period ends. In our case, we have built our technologies into existing commercial offerings that continue to be supported. Therefore, we are able to offer access to both individual components (as Web services or as containers for local installation) as well as packages of functional offers (optionally with front end interfaces) which can be customized to different clients. Trans-vector publishing will be an increasingly critical functionality for media organisations as traditional linear media loses attention and viewers are ubiquitously connected, consuming non-linear media in a format and time interval that corresponds to their preferences and varies with the specific context. ReTV provides the professional-grade feature set for broadcasters and archives to grow their audience and succeed commercially in the digital space.

REFERENCES

1. Apostolidis E., Adamantidou E., Metsai A., Mezaris V. and Patras I., 2020. AC-SUM-GAN: Connecting Actor-Critic and Generative Adversarial Networks for Unsupervised Video Summarization. IEEE Transactions on Circuits and Systems for Video Technology, accepted for publication.
2. Bocyte, R., Atsidis, D., Philipp, B., Lamm, W., Wilson, A., Krämer, K., Nixon, L. and Scharl A., 2021. Deliverable D5.3 Second Validation of Engagement Monitoring Prototype.
3. Bocyte, R. and Oomen, J., 2020. Content Adaptation, Personalisation and Fine-grained Retrieval: Applying AI to Support Engagement with and Reuse of Archival Content at Scale. Proceedings of the 12th International Conference on Agents and Artificial Intelligence, vol. 1, pp. 506-511.
4. Bocyte R., Nixon L., Oomen J. and Scharl A., 2020. Online News Monitoring for Enhanced Reuse of Audiovisual Archives. Digital Libraries for Open Knowledge. TPDL 2020. Lecture Notes in Computer Science, vol. 12246.
5. Bocyte R., Tzialli, E., Philipp, B., Zwicklbauer, M., Lamm, W. and Nixon, L., 2019. Deliverable 5.2 First Validation of Engagement MonitoringPrototype.
6. Lamm, W., Wilson, A., Krämer, K., Bocyte, R., Atsidis, D., Philipp, B., Kündig, M., Ciesielski, K., Apostolidis, K. and Mezaris, V., 2021. Deliverable 6.3 Second Validation of Personalized Prototype.
7. Newman, N., Fletcher, R., Schulz, A., Andı, S. and Nielsen, R., 2020. Reuters Institute Digital News Report 2020.

8. Nixon, L. 2020. Predicting your future audience's popular topics to optimize TV content marketing success. Proceedings of the 2nd International Workshop on AI for Smart TV Content Production, Access and Delivery, October 2020.
9. Scharl, A., Göbel, M., Böhm, K., Philipp, B., Apostolidis, K.,Apostolidis, E., Gkalelis, N., Galanopoulos, D., Adamantidou,E., Mezaris, V., Nixon, L., Steixner, J. and Avdullahi, A., 2020. Deliverable 4.3 Trans-Vector Platform,Final and Optimized Version.
10. Wang, Y., (2020). Humor and camera view on mobile short-form video apps influence user experience and technology-adoption intent, an example of TikTok (DouYin). Computers in Human Behavior, vol.110.
11. Weichselbraun, A., Steixner, J., Braşoveanu, A.M.P., Scharl, A., Göbel, M. and Nixon, L., 2021. Automatic Expansion of Domain-Specific Affective Models for Web Intelligence Applications. Cognitive Computation, vol. 13.
12. Wescott, K., Arbanas, J., Downs, K., Arkenberg, C. and Jarvis, D., 2020. Digital media trends, 15th edition. Courting the consumer in a world of choice.
13. Wright, C., 2017. Are beauty bloggers more influential than traditional industry experts? Journal of Promotional Communications, vol. 5(3).
14. Zwicklbauer, M., Lamm, W., Gordon, M., Apostolidis, K., Philipp, B. and Mezaris, V., 2020. Video Analysis for Interactive Story Creation: The Sandmännchen Showcase. Proceedings of the AI4TV workshop at ACM Multimedia, October 2020.

ACKNOWLEDGMENTS

This work is part of the ReTV project, funded by the European Union's Horizon 2020 Research and Innovation Programme (Grant Agreement No 780656).