The Trans-Vector Platform for optimised Re-purposing and Re-publication of TV Content

Lyndon Nixon
MODUL Technology GmbH
Vienna, Austria
nixon@modultech.eu

Miggi Zwicklbauer
Rundfunk Berlin Brandenburg
Berlin, Germany
miggi.zwicklbauer@rbb-online.de

Lizzy Komen
Netherlands Institute of Sound and Vision
Hilversum, Netherlands
lkomen@beeldengeluid.nl

Basil Philipp
Genistat AG
Zürich, Switzerland
basil.philipp@genistat.ch

ABSTRACT

This submission presents a first prototype for a new and innovative TV content analysis and publication system we call the Trans-Vector Platform (TVP). The TVP derives its added value by the aggregation of TV data from different sources into a Metadata Repository. Its value for TV viewers and TV content publishers alike is reflected in a number of concept applications to be built on top of the TVP. The TVP and the associated applications will now be tested in the EU H2020 funded project ReTV (www.retv-project.eu) with end users to validate the added value of the TVP to TV viewers and the organizations that provide them with media content.

CCS CONCEPTS

• Information systems~Data analytics • Information systems~Multimedia content creation • Applied computing~Publishing • Applied computing~Service-oriented architectures • Applied computing~Digital libraries and archives

KEYWORDS

Content management; Media asset management; Personalized TV; TV recommendations; video summarization; Dynamic Ad Insertion; Content automation; Audience metrics

1 INTRODUCTION: WHO NEEDS A TVP?

All media organizations today face the modern challenge of becoming multi-channel content publishers. Their digital content may still be distributed on linear broadcast TV yet also on IP streams (Live and catch-up TV, VoD, Web archives), mobile apps, third party video platforms and social media. Content creation and publication today is typically manual, supported little by data analytics, and separated in organizations by channel (e.g. the VoD portal is managed separately from the Facebook page).

Today's media stakeholders may lack content-specific analytics on the performance of their content on one or more of these channels and certainly miss a combined view of the publication success across all channels. An innovative tool is missing that can continuously measure and predict the success of digital content (according to its topics) across vectors (the individual publication channels) and thus recommend which content should be re-purposed according to the most appealing topics and then re-published when on which vector.

This tool derives its added value functionalities by aggregating TV data from different sources into a Metadata Repository. This data includes EPG metadata, TV programming descriptive annotations and transcripts, including keyword and entity extraction, video analysis leading to structural and conceptual descriptions, audience metrics, TV content success metrics for Web and social media channels, as well as an event knowledge base. Services for prediction, summarization and recommendation are built on top of this data.

ReTV is an EU Horizon 2020 funded project (www.retv-project.eu) where a collaborative effort is made to develop a first working concept of such an innovative tool, termed the Trans-Vector Platform (TVP). This paper presents the work in progress to deliver first prototypes of several applications which instantiate the TVP (as a set of online interconnected components) and

will allow evaluation of its functionality and usability with both the editors in media organizations as well as the end user: TV content consumers.

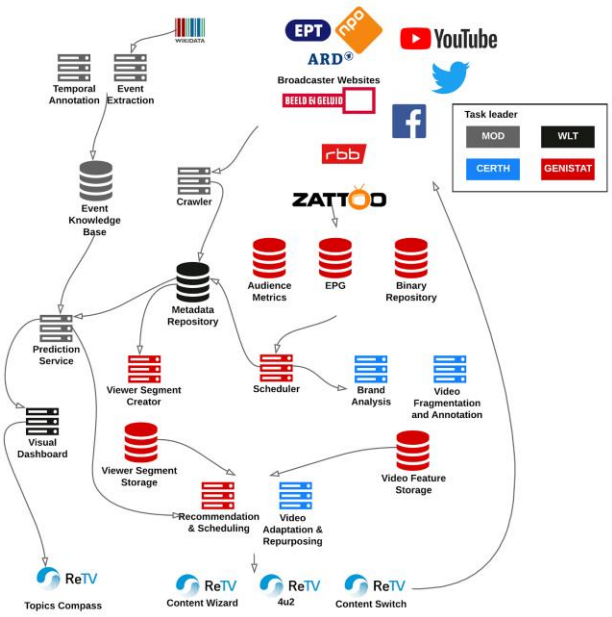

Figure 1: Simplified architecture of the TVP components[1]

2 CONCEPT: A PLATFORM FOR TRANS-VECTOR PUBLICATION

Figure 1 shows the simplified architecture of the TVP. The Trans-Vector Platform (TVP) in its implementation will be a system of modular online components that can be used together (in a Service Oriented Architecture (SOA)) in different configurations and workflows by digital content stakeholders to recommend and re-purpose their media semi-automatically and distribute it optimally across multiple publication vectors.

MOD, CERTH, WLT and GENISTAT refer to the four technical partners of the ReTV project who implement the TVP components. TV program-related data is collected from the content partners (RBB, Sound and Vision and Zattoo) and is aggregated with EPG data as well as open data (e.g. public content on social networks) in a Metadata Repository. TV program video is fragmented and analyzed to produce metadata descriptions of program visual content (in the Video Feature Storage), with a specialized component for detection of well-known brands. Performance of published content is measured by Audience Metrics (viewership of a program on one channel) as well as online publication success metrics (calculated reach and impact of online content). Analysis of the relationships between TV programming (content) and its performance when published on different vectors informs a Prediction Service, e.g. to suggest which topics can have the greatest communication success on a vector at a certain future time. Prediction is corrected according to the effects of external events (e.g. a major sports championship tends to change viewing patterns according to which channels are showing the sports events) through machine learning supported by an Event Knowledge Base. The functionalities of prediction of topics' communication success are combined with TV content descriptions according to visual concepts (which communicate topics) to enable a content Repurposing and Recommendation service (i.e. suggest a repurposed content item that emphasizes the topics which will have the greatest communication success on which vector and which time). Eventually Viewer Segments will allow the suggested content publication to be adapted by viewer types as well. As indicated by the arrowed line going up from the very bottom of the architecture to the top, the TVP-scheduled content is published and thus data about the content of the publication and its publication success can then be fed back into the analysis and metrics components of the TVP in order to provide further learning about the optimization of the content to achieve maximal reach and impact.

3 APPLICATIONS: TESTING WITH PUBLISHERS AND CONSUMERS

Four scenarios have been defined to guide the design and implementation of first prototypical applications on top of the Trans-Vector Platform (TVP) - they are listed at the bottom of the architectural diagram (Fig. 1). Two scenarios consider the functionality and interface needed on the organizational side (professional user). The other two scenarios complement the professional scenarios in considering the functionality and interface needed on the user side (TV viewer). First implementations of each scenario have been prototyped to allow for user evaluations in Summer 2019. Here we will briefly introduce each scenario, the purpose of the application being implemented, show the first results in

[1] ReTV Deliverable 4.1 "Trans-Vector Platform Technology Roadmap and Initial Prototype", available from http://retv-project.eu/deliverables/

terms of design & comment on how this scenario can address a need of media organizations (evaluation criteria).

3.1 Topics Compass

The Topics Compass is intended to guide editors at media organizations to identify trending topics in their own as well as others' content, organized by time (past popularity as well as predictions of future popularity) and vector. It provides an analytical summary that a user can interpret and make informed decisions about what topics (to be covered in some published content) are most likely to attract the highest reach on a certain vector if published at a certain time.

The basis for the Topics Compass will be the data analysis and visualisations of the webLyzard Web Intelligence platform, which aggregates news and social media documents in a Natural Language Processing (NLP) pipeline to extract frequent keywords, associations, sentiment towards topics and geographical source and target, among other things. Named Entity Recognition (NER) is also used to align annotation values to entities defined in public Knowledge Graphs like Wikidata . A simplified Visual Dashboard will present editors with trending topics and filtering by time and vector, so that they can select topics to "target" in content publications (the content preparation comes in the 'Content Wizard', see the next section). For example, Figure 2 shows a sample of associations and frequency of mentions in German language social media of the RBB news program Abendschau over the month of February. There is a clear peak in social media mentions of the Abendschau on 21 February. After examination it can be found that on this day the program was strongly associated with "Boris Palmer" (a German politician and member of the Green Party). An interview with the politician broadcast in the show on that day generated the social media discussion. Such insights can be used by the broadcaster to decide which content to publish (e.g. posts on social media regarding the Boris Palmer interview can be expected to garner more engagement that other topics) or to promote (e.g. the RBB mediathek constantly provides access to the last 7 days of programming, but which program to highlight on the front page? The Abendschau interview with Boris Palmer may be of more interest to the audience than other topics.) A future Topics Compass will use trends in keyword frequency and detection of references to future events to not only provide insights into past publications of content but to predict the comparative future interest in different topics to suggest what content should be published on vectors on future dates.

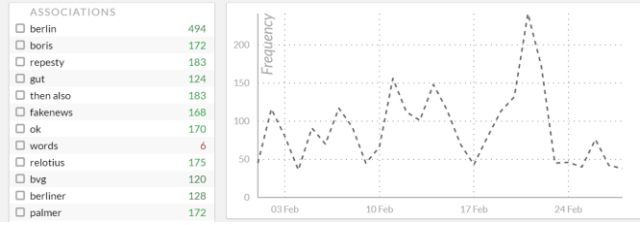

Figure 2. Frequency of mentions and sample of associations with RBB evening news program Abendschau for the month of February 2019

A Topics Compass prototype will be tested with professional users only. This is the only ReTV scenario that is used only by professional users like broadcast editors or heritage professionals. The first tests were done in September 2018 at Sound and Vision and RBB. Both partners have made interviews with the testers to figure out the needs and wishes they have regarding using a tool like the Topics Compass. The next steps in testing will be to build a permanent testing group out of the focus group to have smaller test iterations. The tests will be done via a mix of online questionnaire and face2face meetings. The outcome of each evaluation will be used to optimize the scenario, having a quick and smooth agile iteration of testing and development. The participants should be integrated in the most iterations to build a tool that demonstrates how data analysis and visualizations for online publications about TV can help to optimise their own content publication strategy.

3.2 Content Wizard

As mentioned above, the Content Wizard is also addressing professional users and offers the TVP-powered functionality to select, repurpose and schedule content for publication. A starting point may be the predicted popular topics identified in the Topics Compass, or it may be the selection of video clips directly from within the Content Wizard interface (e.g. as a specific TV program is to be promoted, so available content from that program must be the starting point). From a set of selected videos, the purpose of the Wizard

is to semi-automatically select, given the intended vector and time of publication, the video content with the maximal reach potential (based on predictions of popularity of topics with the vectors audience), re-purpose it according to the vector (emphasizing the more popular topics, but also e.g. social media may have limitations on the duration of a published video) or user's own requirements, and schedule it for re-publication, with the editor able to make a final check and any further adjustments. Figure 3 shows a wireframe UI prototype for the content publication (using Adobe XD). The editor selects the video clips they want to use; based on how they are annotated (topics across fragments) and the predicted popularity of those topics, and according to the vectors chosen for re-publication, a recommended publishing time is suggested. A draft can be saved where the content may be re-purposed (e.g. to highlight clips with the predicted more popular topics), and the final content scheduled for re-publication on each vector.

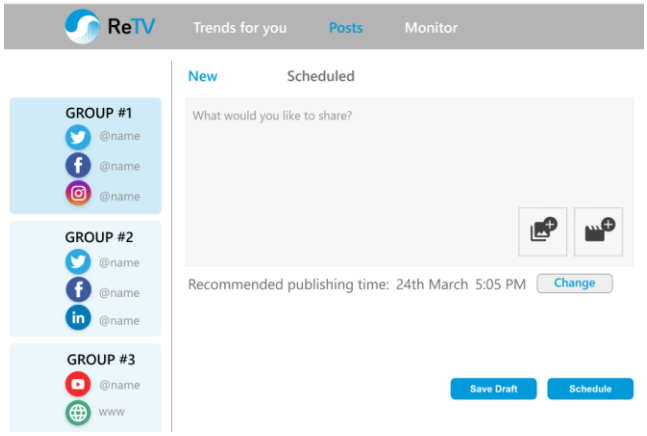

Figure 3. Content Wizard Prototype interface for selecting content and scheduling it for publication

The Content Wizard prototypes will be tested with professional users but also with consumers, as we want to ensure the acceptability of the video summaries produced by the tool. As we initially tested the first ideas of the Content Wizard together with the other scenarios, we learned what requirements were to be prioritized. Sound and Vision built clickable wireframes and interviewed a small group of professional users regarding the planning of social media posts and the creation of video summaries. RBB talked to a limited group of consumers to present them the first ideas of how the video summaries could be used for different purposes on different platforms. The next steps will be to update the wireframes with the outcome of the tests and present the scenario to at least five editorial departments at media organizations.

The outcome of the consumer tests will be concepts on how to integrate video summaries into the TV consumption experience. There will be different kinds of video summaries that must be tested with the consumers: (1) summarization of one or more videos to get an overview of a specific topic and (2) a kind of trailer that sparks interest to the consumer. The trailer could be shown dynamically inserted into a media stream (see 'Content sWitch' below) or used on a Web page, embedded as a mouse-over pop-up for recommended content (see '4u2' below).

3.3 Content sWitch

The content sWitch scenario considers how the TVP-powered annotated and analyzed TV content could enable a broadcaster to personalize their content delivery to viewers in a similar manner to the ideas of Dynamic Ad Insertion (DAI) for using personalization technologies to show different TV viewers of IP streamed TV different interest-targeted ads during breaks. In this case, for example, trailers for this evening's TV programming could be switched in the IP stream to different program previews depending on the classification of the viewer to an audience segment. This would be based on top of the TVP Viewer Segment Creation and Recommendation components. Figure 4 shows a part of a video demo of the content sWitch with the original IP stream of TV content (top left) and three different choices of TV content trailer being shown to three different viewer types in the other frames. Video Adaptation & Re-purposing will also be used as often the content chosen to be shown to the viewer may not be the same length as the available slot for content insertion in the media stream, making appropriate video summarization necessary.

Figure 4. Content sWitch demo - same TV channel, 4 different contents switched into the IP stream

A Content sWitch prototype will be tested with a larger consumer group at Zattoo. We plan to use the RBB media stream to test the Content sWitch in that Zattoo will dynamically replace trailers for RBB programming according to the viewer profile. This will be a large scale test with at least 1000 testers.

3.4 4u2

The 4u2 scenario applies TVP-powered recommendations to TV content on a broadcaster's website. Many Content Management Systems (CMSes) publish sections beneath articles which are automatically filled with links to related content. The linked content here is mostly pointing to other articles, which either are manually edited or linked automatically by basic text matching (between the textual content of the source article and the target article, often just matching with the title or keywords). Click path analysis on the website back up the presupposition that often the recommended articles are not relevant to the reader. The 4u2 scenario implements a CMS plugin that could make significantly better contextual recommendations based on the TVP's content analyses, matching the subject of the TV content which is being presented on the webpage to the subject of TV content being presented on other webpages (the target webpages could therefore be pages of the catch-up TV offer, the archives, VoD content or EPG entries highlighting content being broadcast in the future). Figure 5 shows a mock-up of 4u2 on the broadcaster website, automatically suggesting related content to the current webpage article which can include a (TVP-generated) video summary of the target TV content as a mouseover for the webpage visitor as a form of quick preview (cf. 'Content Wizard' above).

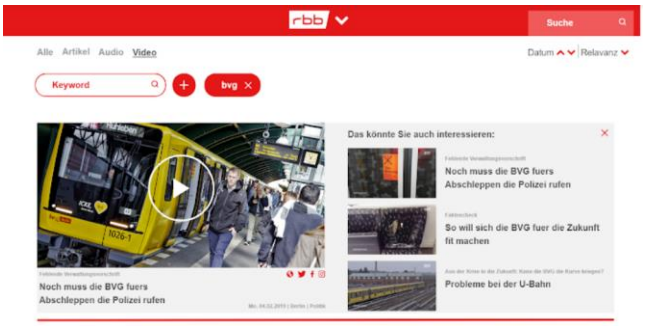

Figure 5. Mock-up of the 4u2 result on a RBB webpage: meaningful recommendations of related TV content to the video article on the left (in the right hand section "Das könnte Sie auch interessieren").

A 4u2 prototype will be tested with content coming from the RBB website. We are planning to test various types of recommendations. This could consider different types of recommended targets (other articles, TV programming in the Mediathek, EPG entries for future programs) and also the integration of social media. The outcome of the 4u2 evaluation will be to enhance the recommended content section of the website in a smarter and more personalized way. The benefit for the professional users will be to have an automatic insertion of recommended content through the TVP that also learns from the changes that the editor may do. This should encourage a higher percentage of click-throughs on the recommended content and website visitors staying longer on the site, viewing more pages. Also the consumers will profit from the smarter recommendations to get references to relevant related content according to their preferences.

4 CONCLUSION: NEXT STEPS AND OUTCOMES

In this paper, we have presented our concept for an innovative new platform for media stakeholders to analyze the performance of their content across multiple vectors and apply that analysis to reach-optimized publication of re-purposed content across the same vectors. We call this platform the Trans-Vector Platform (TVP) and are about to complete a first integrated implementation of all of its components, as shown in the TVP architecture (Section 2). The platform derives its added value functionalities from aggregating TV data in a Metadata Repository. Services for prediction, summarization and recommendation use this data to drive their capabilities. Since a platform only makes sense if there are applications using it, we

have introduced the four scenarios we have settled upon in the ReTV project to showcase the functionality of the TVP for media organizations and for their customers, the TV viewers (Section 3). All four scenarios are being prototyped with the first releases of the different TVP components, and user interfaces have been designed to give the professional user at the media organization access to the resulting TVP data and functionality. We expect to show how the TVP can enable media stakeholders to improve their content publication success across vectors, including with personalized content insertion in media streams and more relevant content recommendations on their websites. This will be validated by technical and user evaluations in the summer of 2019, with the lessons learnt from the evaluations feeding back into a new round of improvements in the components and the user interfaces. The goal of ReTV is a final TVP release with the applications that use the TVP also available to interested media organisations during 2020. Updates on project results and opportunities to test and use the TVP and its applications will be found at the project website (www.retv-project.eu) as well as social media channels.

ACKNOWLEDGMENTS

This work is funded by the EU Horizon 2020 research and innovation programme under grant agreement H2020-780656 ReTV.