



## **THE STORYMAKER – SUPPORTING STORYTELLING WITH USER-GENERATED VIDEOS THROUGH OBJECT RECOGNITION**

Sandy Claes, Rik Bauwens, Klaas Baert, Chaja Libot

VRT innovation, Brussels, Belgium

### **ABSTRACT**

The routines of social media are increasingly integrated in our daily life, including the way we consume and produce audio-visual media. As a result, public service media (PSM) are investing in participatory formats that allow the media consumer to share content for their professional production, thereby causing an overwhelming amount of audio-visual materials that need to be processed manually. In contrast to professional footage, content creators do not have production sheets, scripts or storyboards that indicates the action and camera angles of consumer-produced footage. In this paper, we report on the design and development of Storymaker; a tool that deploys video object recognition to label and categorize video content in order to facilitate the actual storytelling process. As a case study, we present how a first version of this tool was used to create a television item for the Flemish public broadcaster VRT. Then, we present the second iteration and its deployment in a journalistic context. Finally, we reflect on future work within our own organization and beyond.

### **INTRODUCTION**

Today, consumers of traditional media as television and radio have many more tools than before for interacting with that media, making their voice heard, circulating media content within their networks and producing their own content. Audiences are now used to not only receiving content, but to produce, circulate and share content 'Jenkins et al (1)'. A participatory turn is happening in the way audiences engage with media content and media companies. While only a few years ago, user generated video content (UGV) mostly resided on platforms such as Youtube - today, young audiences are active on TikTok, Instagram Stories or Snapchat - social media platforms that offer ready-made storytelling formats. As a result, storytelling with videos is even more familiar.



Meanwhile, traditional media, such as television and news, are exploring new ways to attract younger audiences, of which UGV forms one particular strategy. In fact, for public service media (PSM), UGV contributes to their universal appeal as it allows them to connect to the community at large 'Van Dijck and Poell (2)'. At the BBC, for example, user-generated content projects facilitate an increased representation of its public, while building social capital amongst its participants 'Hutchinson (3), Bonini (4)'. Also at the Flemish PSM, VRT, TV shows that incorporate user generated video trigger a prolonged engagement with its participants 'Claes et al, (5)'. However, in contrast to the above mentioned social media platforms, traditional media are not yet sufficiently organised to process such UGV, both in terms of creative content production as in terms of quantity and image quality. Also, traditional media aims to create professional media content *with* UGV, which is different to those social media platforms that publish content *of* UGV.

In this paper, we present the design and development of Storymaker, i.e. a collaborative tool that aims to bridge the gap between incoming user generated videos and the workspaces of different editorial teams. Furthermore, Storymaker acts as an archiving tool for UGC. This tool is deployed by different editorial teams of VRT (Flemish public service media), which we discuss in the study section. We then present the results and reflect on our learnings. We conclude this paper with future work.

## **PROBLEM CONTEXT**

The context of our study is VRT, i.e. public service media of the Flanders region in Belgium that consists of 3 TV stations, 5 radio channels and 3 online channels. Several of these brands are increasingly combining professionally generated video content with those provided by its users. At the 2018 edition of 'De Warmste Week' for instance, i.e. a week-long charity event organised by VRT in general and aired live on one radio station, media producers received 1500 UGV's (and 15.000 photos) of users that wanted to share their fundraising activities. This video and photo content was approved manually after which they were broadcast in 'carousel' on the live stream of the event. It proved difficult to process this content in storytelling formats. In contrast to professional footage, the editorial team did not know what the UGV contained. Indeed, they did not have production sheets, scripts or storyboards to prepare for an actual story production. Here, we recognized the need to support producers to process that type of content in more automated ways, thereby creating time to focus on storytelling with that content. In particular, we recognized two needs:

1. Processing UGV and integrating it into the existing editorials workflows proved to be rather difficult, tedious, and repetitive. Before the final edit is finished, the uploaded content had to be transferred between workspaces, copied to several storages and downloaded by several people within the same editorial team. Several ad-hoc methods to annotate the content exist, yet there is no structural way to collaborate. In addition, it is difficult for editors to retrieve UGV.
2. In order to request UGV, the editorial team sets up a call-for-contributions. Here, they select a tool that fits best for their audience and their team, such as allowing an upload via Wetransfer, Dropbox, a Qualifio form, etc. Uploads are limited in size and the user experience for the contributor is limited to dropping a file in a folder or



hitting a submit button, after which he or she hopes their content is received well by the editorial team.

## **Prototyping process**

At VRT innovation, i.e. the research and development department of VRT, we build prototypes as a way to bridge problems and needs from our media production teams with novel technologies and theoretical insights. Developing such bridging concept as a tangible artifact allows us to reflect on the needs, stumble upon practical issues and, as a result, sharpen the research question we aimed to tackle and provide 'intermediate knowledge' 'Dalsgaard and Dindler (6)'.

Indeed, our prototypes are not meant to be implemented in production, due to the following reasons:

1. Constantly changing requirements. When we start building prototypes, both technical and functional requirements are not clear yet and are being investigated. Whilst we are working on a small part of the problem, new requirements emerge, change or become obsolete altogether.
2. Development cycle speed. When building software for a production environment, many elements need to be taken into account, including code testing, monitoring, manuals, etc. to name a few. During prototyping, we omit most of these elements for the sake of our development cycle speed. We need to be able to change course abruptly when the research requires it, and even start all over again from scratch when needed. Setting up a complete production environment would take a lot of time, while it is still uncertain whether we are building the 'right thing' 'Savoia (7)'.
3. Dependency on partner services. Most of the time, we develop our prototypes in the context of subsidized research projects, in which we translate novel technologies from project partners to tangible solutions for the media industry. During the duration of the project, we can use this technology to study our research questions. However, these terms of use end with the project. In addition, as public service media, we have to explore the market before closing any deal with a (technology) provider. This could possibly require us to adapt the prototype to another technology provider.
4. Support. We are focused on translating research questions into prototypes and quickly iterating on those prototypes. We provide support for the duration of our studies, which is a fixed period of time. After the testing period, we shut down our prototype and evaluate. Most of the time, we organise a couple of test cycles and iterate on the same prototype for a longer period of time, yet we always require a testing period to have an end date. In other words, we are not able to provide continuous support on all of our prototypes, which a production environment typically requires.

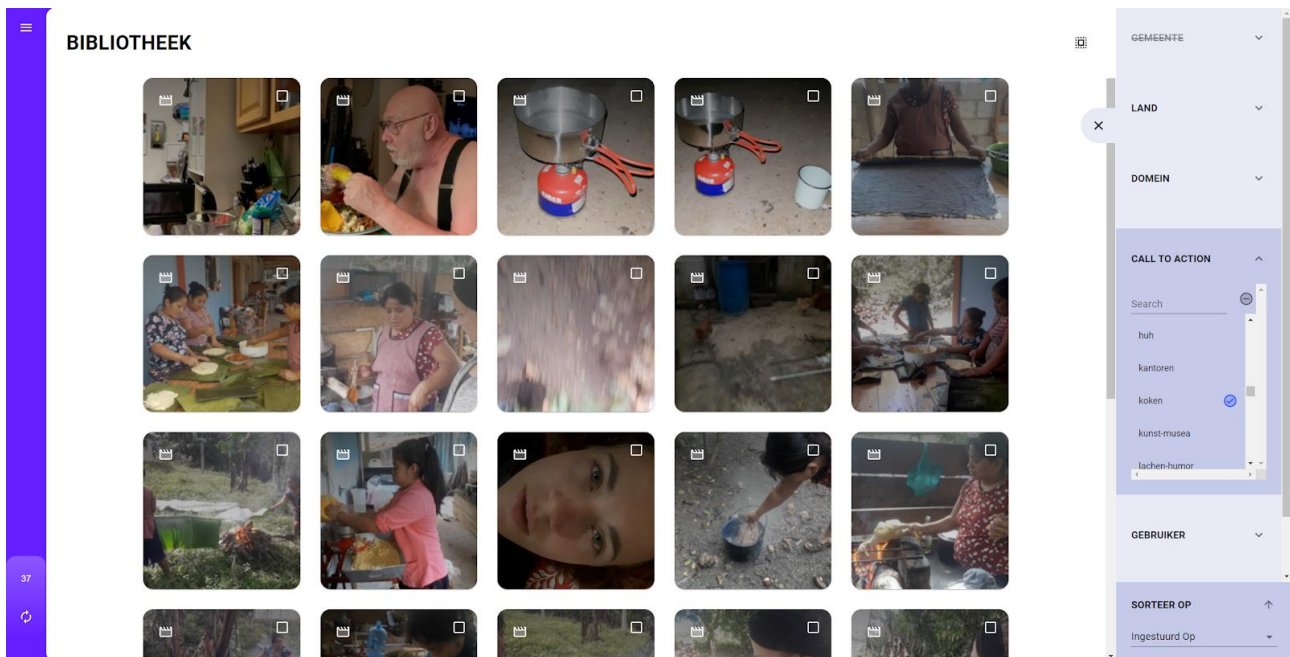


Figure 1 – Screenshot of the Storymaker’s main interface. On the right, the interface offers settings to filter on user, time, location and the call-to-action (i.e. the assignment).

## STORYMAKER

We present Storymaker; a prototype that allows us to easily filter and label UGV through the use of content metadata (e.g. camera type, timestamp, location) and user input (e.g. name). Moreover, metadata is enriched using services provided by Storymaker (i.e. object recognition). For each new uploaded media contribution (UGV), the asset is sent through a chain of analytical services to enrich the available metadata with object recognition data, data from sentiment detection, speech-to-text and topic extraction.

The aim of Storymaker is to provide an editor a clear overview, consisting of an interface that presents videostills of the UGV, and allows search and filter options (see Figure 1). This allows him/her to focus on the story s/he wants to bring to the audience, rather than spending time on organising incoming content. Secondly, UGV can be shared across different departments and teams of the same organisation, as it makes clear which content was used in which context and by whom.

Then, the editor can preview the UGV (see Figure 3), organise and select, and export to existing platforms for editing and publication, such as Adobe Premiere or Experience Manager. In addition, editors are enabled to create and save a project, which is the blueprint for a Story. Each project can have multiple timelines. A timeline contains a selection of assets to allow editorial team members to easily validate. Once the selection is approved by the team, only selected assets are exported - and not others. As a result, assets are included in the standard media flow to prepare for publication and broadcast.

## STUDY AND RESULTS

We take a user-centered approach in the evaluation of Storymaker, which means we start by identifying needs of the users, i.e. professional media producers, and work iteratively. During this process, we regularly request feedback of those users via semi-structured interviews and organised observations. Moreover, when producing and publishing media using user generated content, the tools that are being used today are not yet standard among editorial teams, which is in strong contrast to the production of traditional professional media, where the media flow has been optimized and standardized for years. A gradual approach, such as that of building and iteratively approving a prototype, allows the change in the way the actual production is tackled. In this section, we present how the first version of Storymaker was implemented and what we learned from this, i.e. the results. We then discuss how these results were fed back to the prototype.

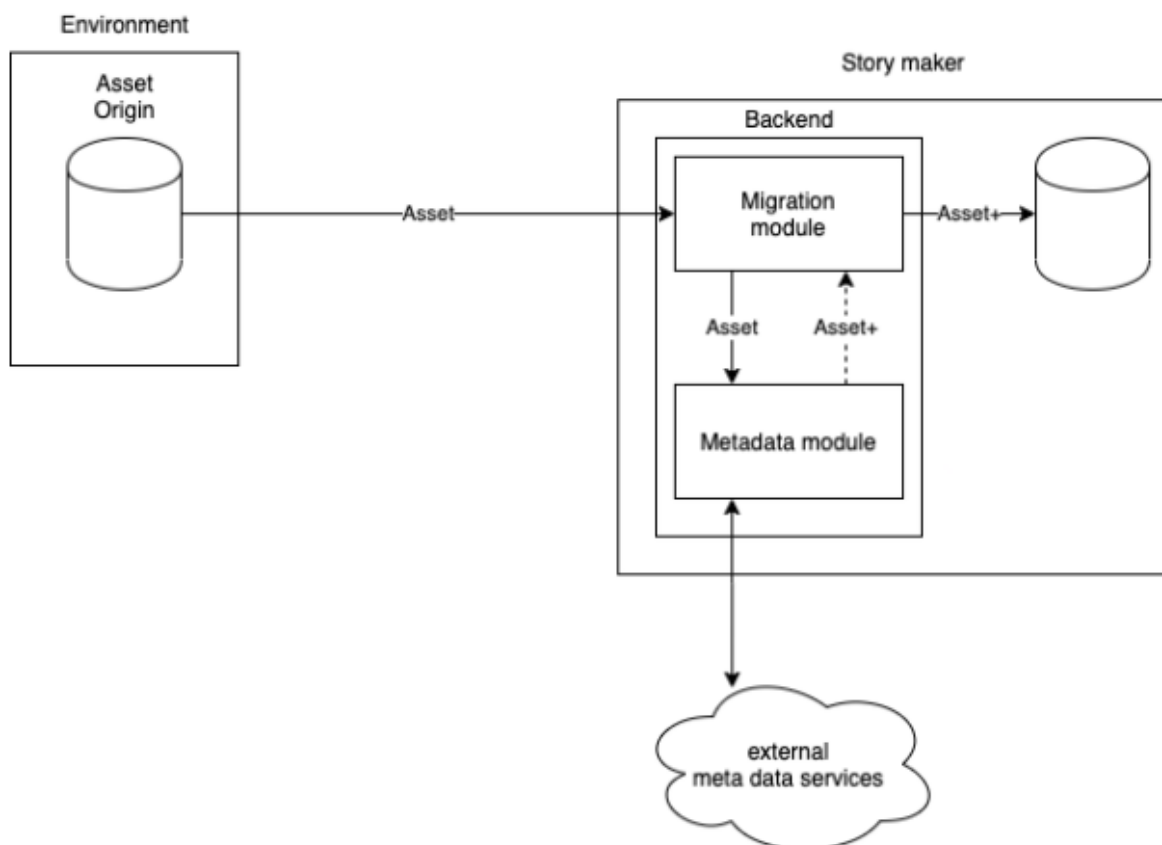


Figure 2 – Architectural diagram of the metadata chain.

### First iteration - use case of weekly TV item

A first, basic version of the Storymaker was deployed to produce the weekly TV item “From Alaska to Zimbabwe” in TV show “Iedereen Beroemd”, starting from August 2019 till May 2020. This version consisted mainly of an administrator tool to manage call-to-actions (i.e. assignments for the contributors). Each of these calls-to-action corresponded to a



unique link to an online upload form. This link was shared with a pre-selected group of users around the world. The uploaded videos were collected inside a Dropbox folder. The path name was predefined, containing the call-to-action, country and username for easy retrieval and could be considered as a primitive type of filtering that worked well for this specific use case.

This first version was a specific use case in which one media producer managed the complete production cycle from organising calls-to-action, contacting contributors, selecting contributions, revising, editing and finalising the TV item. We learned how producers want to scan the UGV in visual ways. In particular, Storymaker should give a structured visual overview (e.g. in a tiled format) of all incoming UGV and its metadata, and allow easy scrolling through the video's timeline. This producer also requested ways to search specific UGV she had seen earlier and remembered, and filter on, and receive an overview of geographic locations of the contributions.

We also learned that it is important for the producer to interact with the contributors as a way to request more contextual information or even additional video footage. However, the medium for communication depended on the contributor, ranging from e-mail conversations to WhatsApp chat messages. Also, already preparing for the implementation of our next iteration with other editorial teams, we learned this need was both specific for the use-case and the way the editorial team wants to operate. This situation would require too many tailor-made implementations and maintenance. As a result, we decided to search for integrations with existing tools.

### **Second iteration - use case of daily news item**

Then, the second iteration included a dynamic way of filtering UGV in a graphical user interface. We decided to discard the call-to-action management as a separate functionality, which opened the possibility to choose other ways of collecting contributions of users. In this version, an off-the-shelf customer engagement platform, i.e. Qualifio [8] was implemented in Storymaker.

When this version was almost finished, the Covid-19 crisis took place, causing a shift in the way broadcasting companies operate. The crisis created a demand for user-generated video content, as - similar to 'De Warmste Week' - the situation caused social initiatives to arise, ranging from a daily applause for people working in care environments, to highlighting the tasks of essential professions (such as bakers, mailmen, etc.). Those initiatives were set up immediately and organised country-wide; a scale that is difficult to capture without several professionals involved (which was prevented by the crisis as well). Here, UGV as crowd-sourced production formed the solution to capture those initiatives and connect them within (often daily) media productions. Besides the practical benefits of deploying UGV, also the qualitative, 'human feel' of this type of content connected to the sense of that time.

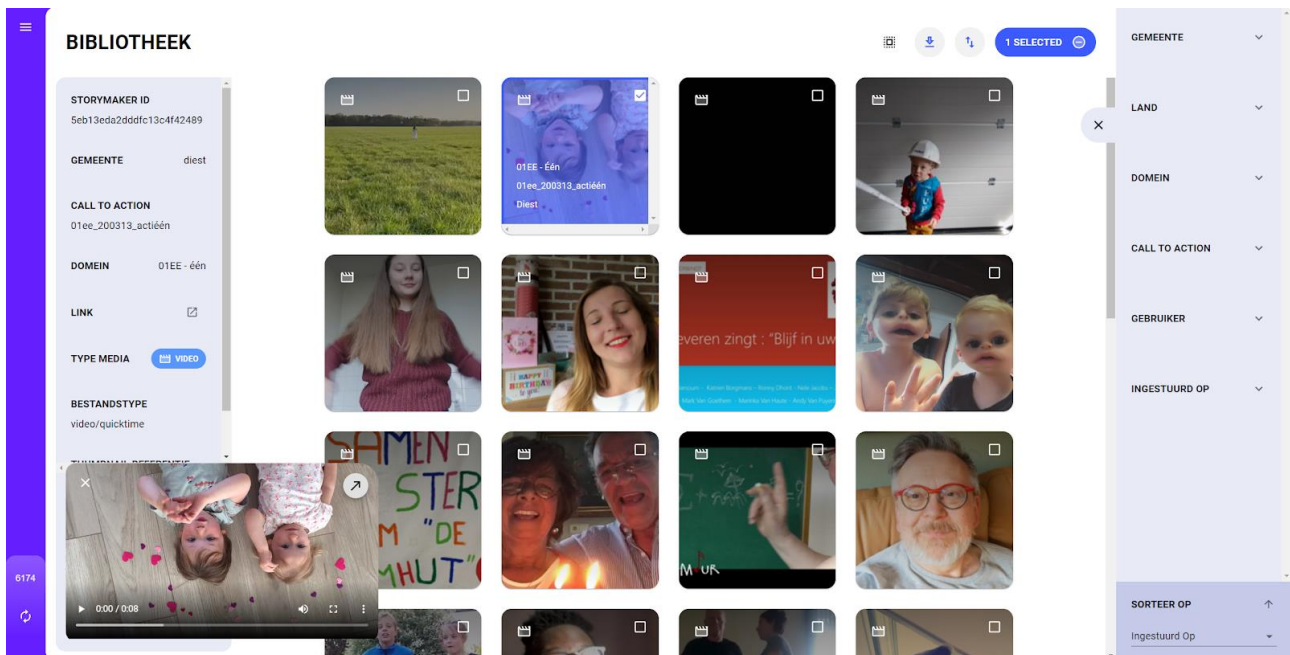


Figure 3 – In the Storymaker interface, UGV can be previewed in an overlay while selecting UGV via the checkbox in the corner on top right of each videostill.

The situation caused Storymaker to be implemented at a fast pace. In particular, Storymaker facilitated journalists that were responsible for the daily applause item (on TV) to edit UGV and prepare for broadcast in a short time span. More specifically, the daily applause was captured by users at 8pm, sent via a Qualifio call-to-action form, then processed and prepared for broadcast at 8.20pm. This tight schedule would have been impossible to keep without the help of a tool like Storymaker.

In particular, the editorial team was enabled to set up a Qualifio form themselves and integrate it with their website on which the call-to-action was launched (i.e. “*Send us your applause videos*”). By integrating the videos that were submitted through Qualifio into Storymaker, the editor had direct access to the videos, and was able to make an initial selection. Often, this selection was based on the location of the contributor as they found it important to have a daily geographical spread of the processed UGV; a filter that Storymaker could provide.

In addition, UGV proved valuable as a time document that reflects the specific situation of the crisis (which is a particular interest of a PSM). As a result, the Storymaker was integrated with VRT’s A/V pipelines by building an integration using Storymaker’s API. VRT uses Aspera Faspex [9] as a means to transport and process A/V content to and from its Arvato workspaces for A/V editing. Because of this integration, an editor can easily select qualifying user-generated content and get started with editing right away. Moreover, only relevant user-generated content is sent through the Faspex A/V pipeline, which reduces processing time and storage capacity.



## **DISCUSSION AND FUTURE WORK**

In this section, we discuss our main findings of collecting and processing UGV in a real production environment. We link these findings to our future work.

### **Enriching metadata**

One of the things Storymaker implies, is a well-structured metadata-set tied to the UGV. Preferably, Storymaker connects to the existing content chain in the beginning of the production process, at the moment when the user contributes videos to the professional media production, in order to avoid loss of metadata along the way. Secondly, Storymaker enriches metadata (i.e. through object recognition) where possible, which makes integration of this enriched data back into the media production content chain paramount to assist production flows later on.

### **Building a modular system**

Storymaker is built as a flexible, modular system. As it is intended to facilitate the organisation of UGV to create stories, it should fit within existing content submission and distribution systems. For example, it is not defined beforehand what type of metadata will be used for filtering, it is derived from the incoming UGV. As such, Storymaker is easy to deploy in any particular use case. Several of Storymaker's functionalities are also accessible via an API, which allows an easy integration in custom environments. Also, every use case has different and specific requirements in terms of metadata. Therefore, metadata is enriched by using separate metadata services, which can be enabled, disabled and customised as needed. The two iterations we discussed above, already proved how each TV item was in need of such a customisable solution.

### **Refining the production experience**

As a next step, we are focusing on adding a map-based overview of contributions, in order to facilitate the selection of UGV to balance the geographical spread. Then, Storymaker will be implemented to process UGV that are retrieved via VRT's radio smartphone applications, including Radio 2, MNM and Studio Brussels. These stations are currently setting up social initiatives that request UGV. Here, the distribution channels are not a linear TV broadcast, but content for amongst others, Instagram and Facebook Stories, requesting a different integration with the A/V pipeline. These radio items will thus serve specific use cases to further refine the Storymaker from production, as well as contributor's side.

### **Augmenting the contributor experience**

Until now, we largely focused on the experience of the media professional. As mentioned earlier, another need is to focus on the experience of the contributor, i.e. the audience. In contrast to social media platforms, in which you upload videos, add stickers or text yourself and publish immediately, the process of delivering UGV for professional production is rather 'black-boxed'; it is unclear whether and how their videos will be used until broadcast. We are currently exploring solutions to improve this experience, including





'instant gratification' effects (i.e. providing immediate feedback to the user that is preferably shareable). Such experiences range from a map that shows all the UGV contributions to a 'track and trace' overview for video processing.

### **Towards a smart system**

The ultimate ambition is to make Storymaker even smarter providing suggestions and recommendations. This could be achieved, for example, by artificial intelligence (AI) services such as object recognition, face detection and automatic A/V quality assessment. When, at a certain time, a lot of similar content is being submitted, Storymaker might alert the editor of this and provide him/her with a collection of this content.

### **CONCLUSION**

In this paper, we presented the design and development of Storymaker, i.e. a tool that allows media professionals to categorize and filter user generated videos, facilitates collaboration and supports archiving. Through an extensive iterative design process, ranging from a first, basic version that was deployed by one professional to produce one TV item to a second version that was used by several teams of media professionals to produce TV and online content, we learned how such systems should be modularised to fit the needs of each media item and allow for easy adoption within the organisation.

### **ACKNOWLEDGEMENTS**

This work was executed within the imec.icon project Essence, a research project bringing together academic researchers and industry partners. The Essence project was co-financed by imec and received project support from Flanders Innovation & Entrepreneurship.

### **REFERENCES**

1. Jenkins, H., S. Ford, and J. Green (2013). *Spreadable media: Creating value and meaning in a networked culture*. New York: NYU Press.
2. Van Dijck, J., & Poell, T. (2015). Making public television social? Public service broadcasting and the challenges of social media. *Television & New Media*, 16(2), 148-164.
3. Hutchinson, J. (2014). Extending the Public Service Remit through ABC Pool. In: G.F. Lowe and F. Martin (eds.), *The Value of Public Service Media*. Göteborg: Nordicom, 223–243.
4. Bonini, T. (2017). The Participatory Turn in public service media. *Public Service Media Renewal: Adaptation to Digital Network Challenges*, Peter Lang, 101-116.
5. Claes, S., Wijnants, M., Libot, C., & Bauwens, R. (2019, September). # TheDay: Triggering User Generated Videos in Participatory Media Productions. In *IFIP Conference on Human-Computer Interaction* (pp. 409-417). Springer, Cham.



6. Dalsgaard, P., & Dindler, C. (2014, April). Between theory and practice: bridging concepts in HCI research. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1635-1644).

7. Savoia, A. (2011). Pretotype it. *Pobrane z: [http://www.pretotyping.org/uploads/1/4/0/9/14099067/pretotype\\_it\\_2nd\\_pretotype\\_edition-2.pdf](http://www.pretotyping.org/uploads/1/4/0/9/14099067/pretotype_it_2nd_pretotype_edition-2.pdf)* (10.12. 2017).

[8] <https://qualifio.com/>

[9] <https://www.ibm.com/nl-en/products/aspera>