AV in DH 2014 workshop

Sound and (moving) images in focus

Workshop at DH2014, 8 July, Lausanne, Switzerland

Location: Swiss Tech Center, Room 5A

Table of Contents

Programme

Schedule

09.00-10.00 Keynote Andreas Fickers – "If content is king, context is its crown": Doing digital media history in the age of abundance

10.00-10.45 Session Data collection & Exploration

- Kleppe – Tracing the afterlife of iconic photographs using IPTC
- Ordelman – Using computer vision to facilitate exploration of television archives

10.45-11.00 Break

11.00-12.00 Keynote Arjan van Hessen – Infrastructures: will they be used?

12.00-13.00 Lunch

13.00-14.30 Session Exploration & Analysis

- Baaren & Van Gorp – 'Disclosed' readings of transmedia content: a demonstration of TROVe
- Nyhan & Flinn – Oral History, audio-visual materials and Digital Humanities: a new 'grand challenge'?
- Huang & Lawaetz – Radio Sound and the Measuring of Sensuous Voice Qualities
- Clement – The Hermeneutics of Distant Listening to Spoken Word Texts with High Performance Sound Technologies for Access and Scholarship

14.30-14.45 Coffee break

14.45-16.00 Session Analysis & Presentation

- Van Gorp, Olesen, Fossati, & Noordegraaf – Emerging Cinema/Emerging Methods: Developing a Tool for EYE's Jean Desmet Collection
- Henderson – The EVIA Digital Archive Project: A Time-Based Media Annotation and Online Access System for Digital Humanities Research, Teaching, and Collaboration
- Sanders & Hagedoorn – How to publish AV research online

16.00-17.00 Synthesis & Concluding remarks

Keynote Andreas Fickers - "If content is king, context is its crown": Doing digital media history in the age of abundance

The recent trends towards digitization and online dissemination of historical sources force the community of professional historians to rethink their traditional research methodologies and practices. Basically, this lecture aims at reflecting on a number of epistemological and methodological challenges of doing digital media history:

- How does the digitization of sources and its online accessibility affect the classical concept of "archive"? (redefinition of archival evidence)

- What new search strategies are necessary in order to cope with the masses of information at hand? (new heuristics of search)

- How to critically interpret digital sources in terms of their authenticity and reliability? (questioning the concept of "original" in digital era)

- How to make use of new digital tools and techniques for the analysis and / or visualization of large data sets? (appropriating digital literacy)

- What new forms of storytelling do digital environments offer for the historian? (multimedia and transmedia storytelling)

Special attention will be paid to the observation that there seems to be a growing gap between the fast development of new techniques and tools for doing digital history and the rather slow appropriation of such tools and techniques by the historical community. Digital literacy - especially in the field of media studies and history - it seems, can't keep pace with the rhythm of innovation in digital technologies.

About Andreas Fickers

Andreas Fickers is Professor for Contemporary and Digital History at the Faculty of Language, Literature, Humanities, Arts, and Education / Luxembourg University . After finishing high school in Belgium, he studied history, philosophy and sociology at the universities of Aachen (G) and Reims (F), specializing on the history of science and technology and the sociology and philosophy of knowledge. Fickers rejoined Aachen University as research and teaching assistant in contemporary history after practical courses at the German Museum for science and technology in Munich and Bonn. His interest in media technologies resulted in a deeper concern for media history (especially radio and television). In 2003 Fickers was appointed Assistant Professor for television history at the department for Media and Culture at Utrecht University, and set up a broader research agenda for a comparative European history of television. In October 2007, he was appointed Associate Professor for Comparative Media History at Maastricht University and became a member of the department of Arts and Culture and the research group Science, Technology and Society. Since September 2013 Fickers works as Professor for Contemporary History at Luxemburg University with a special focus on epistemological and methodological issues concerning digital historiography. His scholarly ambition is to bridge theoretical and methodological approaches from various disciplines (especially from history of technology, media and cultural studies) in order to further develop my ideas of a cultural history of media technologies in the digital age. Fickers is engaged in several European research networks, such as the Tensions of Europe network and the European Television History Network. He is editor-in-chief of the open access e-journal on European television history and culture VIEW, a European multi-media platform for television research. His complete CV can be found here.

Keynote Arjan van Hessen – Infrastructures: will they be used?

The need to handle massive amounts of digitized or digital born audiovisual data that can be used for academic research, not only calls for reconsidering traditional methodological approaches, but also for building solid and sustainable infrastructures. This means creating or converting content according to international accepted standards, facilitating continuous and persistent access to tools and data, and developing a long term preservation policy. Meeting these goals is only possible if institutions join forces across disciplinary boundaries and collectively invest in high cost services.

These requirements have to be met for any data, but the nature of this type of sources – unstructured pixels, color scales and sound waves, asks for interventions at many levels. With regard to online access and providing easy traffic of this type of data there is the need for device- and bandwidth dependent streaming facilities. To be able to analyze the content, tools are needed that enable quantitative measurements of use of color, speed of movements, but foremost tools that support the annotation and interpretation of the verbal and nonverbal layers of the document. We need to know the who, when, where and why at document and fragment level. At present the speedily development of speech and image recognition is helping to counter the backlog in metadata attribution.

The hardest challenge though, lies in finding a compromise with the owners of the data who perceive the charging of copyright fees as rightful in all circumstances (money), and in creating commitment from a broader community of scholars then the circle of media-studies mindset). Especially in the realm of contemporary history, the use of a photo, film, interview recording or documentary, should be as ubiquitous as written documents are for any scholar. This lecture will address the issues described above and offer examples of various successful and non-successful attempts of applying Digital Humanities principles to research based on audiovisual and oral history sources.

About Arjan van Hessen

Arjan van Hessen has been a researcher of speech and language technology since 1986 when he started his PhD in Phonetics. After his PhD he got a 3 year position as Post-Doc at the Universities of Dusseldorf and Ulm. Thereafter he moved to Lernout and Hauspie, a software company in Brussels where he was working on speech recognition. Once back in the Netherlands, he decided to work partially in the industry (ComSys, 1999-2002 and Telecats, since 20002) and partially at the Human Media Interaction group at the University of Twente. His main interest is speech recognition and the use of HLT in the industrial, the cultural heritage and the academic world, as can be seen in the OH-project BalkanVoices. Since 2009 he is member of the executive board of CLARIN-NL, a 6-year NWO infrastructure programme targeted to increase the use of HLT in the humanities. Since 2012 he is working for CLARIN-ERIC, a European Research Infrastructure Consortium stated to make digital language resources available to scholars and researchers of all disciplines, in particular humanities and social sciences. Other activities:

- Member of the board of NOTaS, an organisation of Dutch HLT-SMEs and HLT-Universities, targeted to boost the use of HLT in the Netherlands (and abroad).

- Member of board of Stichting Levende herinneringen, an oral history project focussed on the preservation of the recorded memories of people who lived in Dutch Indonesia

- Member of the executive board of CLARIAH, the successor of CLARIN.

His CV (in Dutch) can be found here.

Abstracts

Kleppe – Tracing the afterlife of iconic photographs using IPTC

In recent popular and scientific products of historiography ever more similar looking photographs of (mostly symbolic) happenings are used as an instrument to tell a story about important events or structural trends. These types of images can be called iconic photographs. We consider iconic photos as photographs that have been reproduced more than once and have a special composition. They refer to archetypes and have the potential to be an archetype itself, and thus represent more than what is being displayed. Although this symbolic meaning is immediately obvious, it may change over time. Since a group of people know these photographs and attribute the same meaning to it, they are part of a 'collective memory' (Kleppe 2013a). Well-known examples of iconic photographs are the photo of a girl running naked and screaming after a napalm bombardment in Vietnam, 1972 or the man in front of a line of tanks at Tiananmen Square, Beijing, 1989. These examples can be considered to be 'global super icons' . However, iconic photographs also exist within a national context (Paul 2008). This paper describes how we determined which Dutch photographs can be called iconic by using IPTC technology. This technique is being used in the media industry to transfer information on photographs in unified standards. We applied this technology to efficiently research two elements of the above formulated definition of an iconic photograph: 1) which photo is published most often in a dataset and therefore functions as an iconic image and 2) how can thechanging symbolic meaning of iconic photos be studied?

The International Press Telecommunication Council (IPTC) develops technical standards for news organisations. By applying the same standards, a photographer can embed the information of a photo inside the file and send this to an editor at a newspaper who can download the metadata in their local ICT-infrastructure. This technique not only facilitates the exchange of files between journalists, but also academics, digital libraries and cultural heritage institutions can use IPTC to include information about their objects in the digital files (Grijsen 2012, Reser 2012).

We used this technology to determine which Dutch photos are published most often by analysing 5.000 photographs in 400 Dutch history textbooks, published in the period 1970 – 2000. All photos were digitized and analyzed by assigning 41 variables (such as topic, caption, person and year). However, finding similar images in a large dataset can be a challenge given the high level of subjectivity when interpreting photographs (Finnegan 2006, Rose 2007) and the lack of standardized thesauri to describe photographs (Kleppe 2012). To overcome this 'semantic gap' (Smeulders 2000) we formulated a list of historic events based on a literature review. Together with all factual information about the photo we added this data in IPTC fields that are embedded within the digital file of the photograph by using the commercial software program Fotostation Pro. This program not only allowed us to do full text searches through all assigned metadata but we were also able to share our research data with other researchers who could import the information in the IPTC fields of their photo–, editing- and viewing software. Moreover, we were able to export all values to csv-files that were importable into statistical software packages such as SPSS.

By making frequency tables of the list of historic events we calculated which topics were most present in the set of photographs. We then manually went over the images that illustrate these topics to find the images that were used most often, answering our first research question. Results show that a photograph of 1912 of socialist politician Pieter Jelles Troelstra is used most often in the analyzed textbooks. On the photo, Troelstra gives a speech in which he pleads for universal suffrage.

To study our second research question on the changing symbolic meaning of iconic photographs we could return to our database since each photo is not only described based on the list of historic events

but we also noted down factual information such as the chapter title in which the photo was published and the accompanying caption. By going over this information we could examine in which context this photo of Troelstra was used. By simply typing in the name 'Troelstra' in Fotostation Pro, we found all books in which the photo was used and could go over the information in the accompanying IPTC field. By following this approach we could answer our second research question on the changing symbolic meaning of iconic photographs. In the case of Troelstra, we found that the photo is incorrectly dated in one-third of all history textbooks. Nowadays, in Dutch historiography Troelstra is not so much known for his plea for universal suffrage but mainly for his failed attempt to start a revolution to overthrow the queen in 1918. Our research shows that the photo of 1912 is used to illustrate the events of 1918 instead of the demonstration of 1912 in one-third of all Dutch history textbooks, clearly illustrating the changing symbolic meaning of this iconic photographs. This outcome gives ground for several follow-up historiographical research questions focusing on both the afterlife of photographs (Kroes 2007) as well as the selection processes of historical gatekeepers (Kleppe 2013a).

Even though our database is relatively small, the case-study of the photo of Troelstra shows that by adding metadata in the IPTC-fields, we were able to quickly track-down all the textbooks in which the photo is used and determine the context in which the photo is used. Studying this afterlife can even be taken a step further when databases with the same approach can be linked, e.g. collections that are described with the ICONCLASS System (Brandhorst 2012) or the GTAA (Oomen 2010). Therefore we made our database available for future researchers (Kleppe 2013b) in order to be reused to answer Humanities research questions or further exploration on the use of IPTC within the Digital Humanities. However, going over the frequency tables and subsequently looking up the information in the database and IPTC fields remains a manual process. Even though the software allowed us to quickly search through all data, we still had to rely on our own judgement and scrutiny. We envision future developments in image retrieval by using image recognition will be of assistance for this type of research (Wu 2009). At this moment several search engines offer Reversed Image Lookup (RIL) such as [Google Images](#) and [Tineye](#). Users can upload an image and retrieve similar like images located at websites. This type of technology is already being used by commercial parties to [track-down copyright protected images across the web](#) 3 , to assess the impact of scholarly images online (Kousha 2010), to find patterns in large image databases (Manovich 2009) or to analyse the reuse of digital images of cultural and heritage material (Terras 2013). The next step of applying this type of Reversed Image Lookup would be its application to recognize similar images in a dataset, allowing researchers to automatically retrieve the most often published images and analyze its afterlife.

References

Brandhorst, H.: "The Iconography of the Pleasures and Problems of Drink: Thoughts on the Opportunities and Challenges for Access and Collaboration in the Digital Age". In:*Visual Resources*, 28(4), 384-390 (2012).

Finnegan, C.A.: "What is this a picture of? Some Thoughts on Images and Archives", In:*Rethoric & Public Affairs* 9, 116 – 123 (2006).

Grijsen, C.: "In perspectief: behoud en beheer van born-digital fotoarchieven" (In perspective: conservation and management of born-digital photo archives). In: *Fotografisch Geheugen 75*, 24 – 26. (2012).

Kleppe, M.: *Canonieke Icoonfoto's*. De rol van (pers)foto's in de Nederlandse geschiedschrijving (Canonical Iconic Photographs: The role of (press) photos in Dutch Historiography). Eburon, Delft (2013a).

Kleppe, M.: *Foto's in Nederlandse Geschiedenisschoolboeken (FiNGS)* (Photos in Dutch History textbooks) http://www.persistent-identifier.nl/?identifier=urn:nbn:nl:ui:13-l37n-bi(2013b).

Kleppe, M.: "Wat is het onderwerp op een foto? De kansen en problemen bij het opzetten van een eigen fotodatabase" (What is the subject of a picture? The opportunities and difficulties in setting up their own photo database). In: *Tijdschrift voor Mediageschiedenis 2*93 – 107 (2012).

Kroes, R.: P*hotographic memories: Private pictures, public images, and American history.*Dartmouth College, Hanover (2007).

Kousha, Kayvan, Mike Thelwall, and Somayeh Rezaie. "Can the impact of scholarly images be assessed online? An exploratory study using image identificationtechnology." *Journal of the American Society for Information Science and Technology* 61.9 (2010): 1734-1744.

Manovich, Lev. "Cultural analytics: Visualing cultural patterns in the era of more media."*Domus* (923) (2009).

Oomen, Johan. & Brugman, Hennie: "Thesauri gekoppeld" (Thesauri Linked), In: *Digitale Bibliotheek 2* (5) 18- 21 (2010)

Paul, Gerhard, *Das Jahrhundert der Bilder. 1949 bis heute* (Göttingen 2008).

Reser, G., & Bauman, J.: "The Past, Present, and Future of Embedded Metadata for the Long-Term Maintenance of and Access to Digital Image Files". In: I*nternational Journal of Digital Library Systems (IJDLS), 3*(1), 53-64 (2012).

Rose, G.: *Visual Methodologies – An Introduction to the interpretation of Visual Materials.*Sage Publications, London (2007).

Smeulders, Arnold W. M. e.a., "Content-Based Image Retrieval at the End of the Early Years", *IEEE Transactions on Pattern Analysis and Machine Intelligence 22*, 1349 – 1380 (2000).

Terras, M. M., and I. Kirton. "Where do images of art go once they go online? A Reverse Image Lookup study to assess the dissemination of digitized cultural heritage." S*elected papers from Museums and the Web North America,* 237 – 248 (2013) .

Wu, Zhong, et al. "Bundling features for large scale partial-duplicate web image search."*Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on.* IEEE, (2009).

Baaren & Van Gorp – 'Disclosed' readings of transmedia content: a demonstration of TROVe

Media play an important role in the (re-)articulation of discourses and the construction of cultural meaning. With digitization, convergence and the emergence of social media, the construction and diffusion of news and opinion increasingly operate within a transmedia sphere, in which different media interact (e.g. Henkins, 2006). The newly developed research tool TROVe (Transmedia Observatory) monitors the spread of news and public opinion across Dutch transmedia content: television, radio, blogs, online newspapers and Twitter. In this paper, we demonstrate the role of TROVe in the research process of media scholars. More specifically, we will look at the dynamics of 'distant' and 'close' readings at play in the context of the exploration and analysis of transmedia content.

The research phases of media researchers, and humanities scholars alike, are dynamic processes (Bron et al, 2012). While exploring the archive, media researchers browse, jump from one to the other topic, and create meaningful links between nodes. Media researchers draw links between different documents while also keeping the larger picture in mind (Van Gorp, 2013). Franco Moretti coined these processes of getting a helicopter-view, alongside with a look at individual documents, as 'distant' reading versus 'close' reading (Moretti, 2000). 'Distant' reading is defined as 'the crunching of large quantities of information across a corpus of textual data or its metadata' (Burdick et al, 2012, 18). Close reading concern an in-depth inspection of the characteristics and interpretation of one or a small selection of documents.

TROVe offers both a distant view on 'trends' in the different data collections and a 'close' reading of individual documents. The tool enables to explore and identify important moments and actors in a transmedia debate, for example about political subject, at a glance. When entering a keyword, a time line indicates frequencies, starting points, ending points and peaks/bursts. These visualizations prompt researchers to further analyse how documents and/or collections are related to each other. These lead to rephrasing and adjustment of existing research questions and related queries. Second, TROVe also offers filters and facets with wordclouds of related topics (contextual words found in the same document as the original keyword) and related actors (Twitter accounts, websites and names of organizations). These wordclouds, again, give an indication of the characteristics of the debate, the topics, the persons, broadcasters and web publishers. The results feed further questions about what is said by whom in relation to time and context.

When the user notices an interesting or unexpected evolution or pattern, s/he can zoom in and inspect individual documents. By doing so, dots appear, which offer a pop-up window with metadata (e.g. title, viewing rates), and (links to) the primary source. Tweets, blogposts and online news are embedded within the pop-up. Audiovisual sources are provided by an external link, of which a selection also leads to streamed audiovisual content (via the portal Academia.nl). These metadata and original content of individual documents provide material for close reading. Audiovisual and textual content (such as tweets) can be systematically analysed by means of existing methods in humanities and media studies such as discourse analysis and/or qualitative content analysis.

In sum, TROVe enables both close and distant reading and supports the existing research dynamics of media studies. However, for media scholars, switching between individual (audiovisual) documents and larger contexts is an iterative process that remains largely internalized (see e.g. Van Gorp, 2013). The interface plays an important role here: collecting (audiovisual) documents for close reading requires a continuous process of returning to graphs and contextual wordclouds. In turn, graphs and wordclouds can only be understood by understanding the primary documents that co-create them. Using TROVe thus requires researchers to switch dynamically between the graphical 'helicopter' view to the individual document by actively clicking, zooming, changing views and interpreting visualizations. This process can make them more aware of the patterns and evolutions in which their individual documents of interest are embedded. The continuous and seemingless switching between distant and close reading in TROVe may create a form of 'disclosed' readings with more and/or different opportunities for the media researcher. Future research and discussions may tackle the implications of (new) research questions and (mixed) research methods that these disclosed readings imply.

References

Bron, M., Van Gorp, J. , Nack, F. F., de Rijke, M., Vishneuski, A. & de Leeuw, J.S. (2012) A Subjunctive Exploratory Search Interface to Support Media Studies Researchers. *SIGIR 2012: 35th international ACM SIGIR conference on research and development in information retrieval.* Portland: ACM. Burdick, A., Drucker, J., Lunefeld, P., Presner, T. Schnapp, J. (2012) *Digital_Humanities*. Cambridge, MA: MIT Press. Jenkins, H. (2006). C*onvergence Culture: Where Old and New Media Collide.* New York: New York University Press. Moretti, F. (2000) "Conjectures on World Literature," *New Left Review,* 1 (Jan–Feb.), pp. 54–68. Van Gorp, J. (2013) Looking for what you are looking for: a media researcher's first search in a television archive. *VIEW: Journal of European Television History and Culture, 2*(1).

Nyhan & Flinn – Oral History, audio-visual materials and Digital Humanities: a new 'grand challenge'?

Oral historians have long recognised that voice and gesture can communicate information, knowledge, emotion and interpretation in ways that text cannot. Indeed, oral history artefacts can be studied not only for the words they contain but also for features like interjections, gestures and silences that can, among other things, contain clues about an interviewee's emotional state (see, for example, Good 2006). Tebeau has written that "the digital age has changed the field in dramatic ways, in particular by extending the possibilities for using voices both in research and public interpretive projects" (2013) and his Cleveland Historical project is exemplary of this. Nevertheless, it can be argued that this process has not gone as far as is necessary, as Frisch has put it "Everyone knows that the core audio-visual dimension of oral history is notoriously underutilized" (Frisch 2006 p.102)

Technological developments—often based on advances made in the Digital Humanities community involving structured and semantic markup—have opened up a plethora of new ways to process audio-visual materials. As Frisch has written

Oral history audio and video can now be placed in an environment in which rich annotation, cross-referencing codes, and other descriptive or analytic 'meta-data' can be linked to specific passages of audio-visual content … This will, of course, allow the audio-visual files themselves to be 'searched, browsed, accessed, studied, and selected for use at a high level of specificity' (p.103) … On this software frontier, audio and video documentation becomes as richly and easily accessible as a well-organized text-based-reference book, and far more easily accessible (op cit, p.104).

This is indeed true, and these methods are and will surely remain important way of making such materials tractable. However, it is notable that these methods continue to privilege largely text-based approaches to Oral History; after all, what is meta-data but natural language codes inserted into a text in order to make explicit its meaning or constituent parts? Methods being developed in other fields that have, as yet, seen relatively little take up in Digital Humanities, for example, image and facial recognition, acoustic approaches to sentiment analysis, 3D imaging and modelling, digital narratology and storytelling etc offer methodologies that could be fruitfully brought to bear on the capture and especially the analysis of such sources. Not only might such approaches offer new interpretative strategies—that are neither founded upon nor predominately focused upon text—for engaging with audio-visual materials, but they could contribute to a more thorough and sustained reassessment of the dominance of the 'written' word in fields like Digital Humanities and Oral history.

This paper will be based on two case studies drawn from the 'Hidden Histories' project, which is undertaking oral history research on the history of Digital Humanities (see, for example, Nyhan, Flinn et al. 2013). The first case study will identify how oral history recordings currently tend to be made available in Digital Humanities fora. Special emphasis will be placed on a critical evaluation of a recent *Digital Humanities Quarterly* special edition of oral history materials that we edited as part of our Hidden Histories research (Nyhan ed. 2012). The second part of this paper will discuss ongoing research into the analysis of the c. 50 oral history recordings that are being made during the course of our project. In this section a 'blue sky' approach will be taken as we reflect on the new kinds of research on and with the recordings that technologies such as those discussed above could facilitate.

The technologies and methodologies discussed will include:

1. Advances in the area of visual computing where, for example, it has been recognised that emotional changes are often signalled by changes in the colour of a person's skin. Jimenez et al (2010) have published a 'practical appearance model for dynamic facial colour' and have

shown that it can easily be integrated with existing and planned animation and rendering projects. Leaving the details of the technical implementation aside, how might techniques such as these be applied to videos of oral history interviews in order to analyse them in new ways? What new kinds of searching of visual records might they support and what kinds of new knowledge about oral history could be created as a result of such applications?

2. Once a rather recherché research area, since post 9/11 in particular numerous advances in the area of image mining and facial recognition of actors in photographs and videos can be noticed, though the investigation of the application of these techniques to the Digital Humanities and Cultural Heritage sectors has so far been relatively fragmented.   What new kinds of questions might such techniques be able to support? For example, how might advances in this area allow oral history photos and videos to be studied in terms of the gestures, body language and the spatial location of participants?

3. Sentiment analysis and opinion mining combines techniques from NLP, machine learning and statistics and is usually applied to texts of various kinds in order to detect words that give an insight into the attitude of the actor under study. What approaches exist for performing such an analysis on the spoken word and how might they be applied to oral history research?

As stated above this section will be very much 'blue-sky'; to ensure that the possibilities discussed remain grounded in the research questions that they might be applied to reference will be made throughout to the Hidden Histories project and the issues we are pursuing there.

In summary we will aim to identify a number of desiderata for the treatment of digital oral history recordings. It will also be argued that by harnessing methods that are, as yet, underutilised in Digital Humanities, that research on audio-visual materials can open new and as yet little known vistas for Digital Humanities researchers. Indeed, when one considers the wealth of audio-visual information available online one might say that such research has the potential to open a new "grand challenge" for the field. Perhaps this is one of the most fundamental ways that we can set about remedying the problem posed in the call for papers for this workshop, namely "how to overcome the contrast between audiovisual material being a steadily increasing body of data and the fact that it is relatively poorly represented in the field of the Digital Humanities"?

References

Frisch, M., 2006. Oral history and the digital revolution: toward a post-documentary sensibility. In R. Perks & A. Thomson, eds. *The Oral History Reader*. New York: Routledge, pp. 102–114.

Good, F., 2006. Voice, ear and text: words, meaning and transcription. In R. Perks & A. Thomson, eds. *The Oral History Reader*. New York: Routledge, pp. 362–373.

Jorge Jimenez, Timothy Scully, Nuno Barbosa, Craig Donner, Xenxo Alvarez, Teresa Vieira, Paul Matts, Veronica Orvalho, Diego Gutierrez, Tim Weyrich. 2010. In *ACM Transactions on Graphics (Proc. SIGGRAPH Asia), Vol. 29*, No. 5, pp. 141:1–141:8, Seoul.

Nyhan, J., Flinn, A. & Welsh, A., 2013. Oral History and the Hidden Histories project: towards histories of computing in the humanities. *Literary and Linguistic Computing*, p.fqt044.

Nyhan, J., 2012 Ed. D*igital Humanities Quarterly [a special edition on Hidden Histories: Computing and the Humanities c.1965-1985]*. 6:3http://www.digitalhumanities.org/dhq/vol/6/3/index.html

Tebeau, M., 2013. Listening to the City: Oral History and Place in the Digital Era. *Oral History Review, 40*(1), pp.25–35.

Huang & Lawaetz – Radio Sound and the Measuring of Sensuous Voice Qualities

Until the digital revolution (Lauer 2013), accessing audio archives and resources has been very difficult. As a consequence, radio as an oral culture has rarely been studied systematically, on the basis of a large amount of data and clean datasets. The recent opening of several broadcast radio archives (BBC, DR) calls for new methods for investigating radio recordings as sound and oral sources.

Through a demonstration of tools, visualisations and findings, we will present and discuss the tools for analysis of radio voices which have developed and used for a recent study. The study focuses on capturing and measuring the sensuous voice qualities of the continuity announcers of one key Danish radio channel (P1). By sensuous qualities, we understand qualities in vocal performance that go beyond the textual and semantic qualities, which may be analysed in transcripts. Radio recordings in the present study span a period from 1940 till 2012. Through an interdisciplinary approach borrowing methodological tools and insights from linguistics, rhetoric and acoustics we introduce and demonstrate a new method for handling voice recordings within the field of cultural studies.

The analysis of recorded voices can be divided into three layers: (1) the acoustic signal, (2) the perception of the signal and (3) the contextualizing knowledge of the performative situation, space and role that affect the interpretation of the results of the acoustic and perceptive analysis. (Lawaetz Submitted). The *acoustic signal of the voice recording* is analysed as raw data. It is possible to measure the pitch and create modulation profiles which, in the future, may be extracted into searchable data criteria. It is also possible to measure speech and articulation rate. For this the following free-software are used: Audacity, Praat, Wavesurfer, and a basic tool, Excel. The *perception of sound* is individual. Nevertheless we need terminology to describe the quality of voice. Audiologopaedics already have a vocabulary and a schemata, though for dysfunctional voices. A schemata for media-voices will be presented with this paper. The contextualising k*nowledge of the performative situation* is important for the interpretation of the acoustic analysis. The central questions are: Which 'role' does the speaker perform and how does s/he perform in the actual space (interaction with technology such as distance to microphone affects not only the recorded signal, but also the voice performance (Brixen 2014)). Since the archives are only presenting parts of what was actually broadcasted, the analysis must be combined with other historical documentation and researched material.

Through a combination of software tools (*Praat, Wavesurfer, Audacity, Excel*) the non-verbal voice qualities are measured, revealing a non-conscious, institutional preference regarding female continuity announcers' voices. Fundamental frequency histograms show that certain female modulating is preferred. This adds a new layer to the history of the Danish broadcast cooperation and opens up for wider discussions of gender and voice in public sphere.

Although the method for analysis was created on a sample of only 98 voice recordings we would like to discuss the possibilities of applying the method on other oral sources. With modifications the method might be useful for researchers working with oral recordings in general and part of it may be turned into an algorithm.

The LARM-project: The recent study is performed with the Danish radio RI-project LARM (2010-2014). The overall focus of the LARM-project has been the accessibility of radio archives for research, through the interface larm.fm. More than 40 persons from universities, the broad cast cooperation and the national library have been collaborating to further develop the field of radio studies in general and more specifically create an virtual archival interface with research tools for humanities scholars.

References

Brixen, E.B., 2014. A Study of F0 as a function of vocal effort. I AES 54th International Conference. London, UK. Lauer, G., 2013. Die digitale Vermessung der Kultur Geisteswissenschaften als Digital Humanities. I Big Data. Das neue Versprechen der Allwissenheit. Berlin: Suhrkamp Verlag Berlin, s. 99–116. Lawaetz, A., Submitted. Danmarks Radios stemmer – undersøgt gennem et studie af kanalspeakeren på P1 1925-2012. Phd. København: København Universitet.

Clement – The Hermeneutics of Distant Listening to Spoken Word Texts with High Performance Sound Technologies for Access and Scholarship

Developing computing systems that facilitate critical "close listening" alongside "distant listening" (such as data mining and visualization) practices seems to hold the promise of creating critically productive tools for analyzing digital audio collections of poetry performances, field recordings, oral histories, and other spoken word artifacts. Since most of these collections lack transcripts, however, it is important that we consider the viability of a hermeneutics of listening that relies on paralinguistic sonic patterns for insight. Michael Chion, for example, emphasizes *causal* and *semantic* listening over what he calls *reduced* listening (or listening to sonic traits independent of a sound's cause or a comprehension of linguistic meanings) (Chion, "Three Modes," 50; emphasis added); he claims that our lack of a descriptive system for such sounds precludes our ability to interpret them.

I would like to consider the nature of non-verbal and paralinguistic sounds within the context of descriptive systems in ethnography and transcription studies as well as within projects conceived for computational analyis such as Alan Lomax's Cantometrics (Lomax 2009), the Text Encoding Initiative's *Guidelines for Electronic Text Encoding and Interchange* (2014), and the author's ongoing digital humanities project, High Performance Sound Technologies for Access and Scholarship (HiPSTAS). HiPSTAS participants include twenty humanities junior and senior faculty and advanced graduate students, as well as librarians and archivists from across the U.S. interested in creating access to and scholarship with large collections of spoken word collections. A significant part of the HiPSTAS project includes introducing participants, all of whom had never used advanced machine learning technologies and visualizations for accessing and analyzing audio, to the ARLO (Adaptive Recognition with Layered Optimization) software designed by HiPSTAS collaborator David Tcheng. Originally developed for acoustic studies in the fields of animal behavior and ecology to begin exploring the use of machine learning for data analysis, ARLO uses spectrograms to extract sonic features for matching, discovery (clustering), and automated classification (prediction or supervised learning).

Over the course of the last year HiPSTAS participants project have pursued a variety of scholarly interests using ARLO with the PennSound project, which includes over 5,000 hours of poetry readings (these are all described at https://sites.google.com/site/nehhipstas/project-pages). Eric Rettberg argues that the nature of a poem can change dramatically when performed in front of an audience as the poem becomes a dialogue between audience and poet. Finding moments of applause or laughter can indicate significant moments of dynamic exchange. "In close listenings of four versions of William Carlos Williams's 'This is Just to Say,'" Rettberg argues that a version in which Williams seemed to try to get the audience to laugh and failed showcases the power of laughter at poetry readings, transforming the poem from a whimsical delight to a highly serious, academic reflection on the nature of art and poetry" (Rettberg). Chris Mustazza's research focuses on the materiality of poetry recordings and to what extent provenance can be determined based on recording artifacts from a particular device. Marit MacArthur is interested in what she calls "monotonous incantation," which is in part signified by the repetition of a falling cadence within a narrow range of pitch. MacArthur posits that "ARLO can provide superior analysis of pitch patterns in the human voices, compared to other

programs commonly used for pitch tracking, in part because it uses spectral analysis as output" (MacArthur). She seeks to explore the influence of "monotonous incantation" across collections and movements, including the Beats, the Black Arts movement, and the black church.

In particular, I am thinking through the historical context of descriptive systems and case studies from HiPSTAS participants to suggest that reconceiving the hermeneutics of sound studies within the framework of an algorithmic criticism (Ramsay 2012) that embraces intervention, defamiliarization, and deformation as inherent aspects of the interpretive process reframes our understanding of how close and distant listening to sound in the absence of transcripts can be enabled.

References

Chion, Michael. "The Three Listening Modes" in *The Sound Studies Reader*. Ed. Jonathan Sterne. New York: Routledge, 2012. 520-532. Print. Ramsay, Stephen. *Reading Machines: Toward an Algorithmic Criticism*. 1st Edition. University of Illinois Press, 2011. Print. TEI Consortium, eds. *TEI P5: Guidelines for Electronic Text Encoding and Interchange*. P5. TEI Consortium. 20 Jan. 2014. Web. 15 May 2014. http://www.tei-c.org/Guidelines/P5/.

Van Gorp, Olesen, Fossati, & Noordegraaf – Emerging Cinema/Emerging Methods: Developing a Tool for EYE's Jean Desmet Collection

The digitization of film heritage collections and the emergence of digital film-related datasets comes with the promise of finding new pathways to knowledge on cinema's history. Although a number of initiatives exist that make use of digitized film collections for studying film history (e.g. Tsivian 2009, Pafort-Overduin 2009, Maltby, Biltereyst and Meers 2011, Manovich 2013), film scholars generally have been slow in seizing the opportunities of digital film collections for e-humanities research. Within the context of the research project Creative Amsterdam: An E-humanities Perspective (CREATE) at the University of Amsterdam's Centre for Cultural Heritage and Identity, the authors currently prepare research on the emergence of cinema as a new cultural industry in Amsterdam, in the early 20th century. Based on existing data and collections, such as the Desmet collection at EYE Film Institute Netherlands, Cinema Context, IMDB, the Dutch Biography Portal and the Netherlands Theatre Institute collection, we intend to map the social and economic networks in which cinema entrepreneurs and consumers operated. Using existing and new computational tools for analyzing moving image content, we will connect the outcome of the network analysis to an interpretation of the films that were produced, distributed and consumed. The challenge we take up is to find out which new questions about the history of early cinema can be answered on the basis of which datasets, thus contributing to unlocking the potential of audiovisual and AV-related collections.

As pilot data for our theoretical, technological and methodological exploration, we selected EYE's collection of film distributor and cinema owner Jean Desmet (1875-1956). Inscribed on UNESCO's Memory of the World Register in 2011, Jean Desmet's unique collection of films, posters, photographs and business documents from the period 1907 to 1916 offers an unparalleled insight into cinema's formal and industrial transition in these years; from early cinema's intermedial attractions to its institutionalization as a narrative art form distributed in purpose-built cinemas. Archivists and scholars already studied the Jean Desmet collection. The discoveries have cleared up a number of misconceptions, and the film-historical appreciation for historical genres such as Italian diva films, German melodramas and French comedies has been changed. Also the wealth of information about film distribution and cinema operation in the period 1907-16 has been studied extensively (e.g. Blom 2003). The question, now, is what the use of computational techniques, and the development of a demonstrator-tool in particular, can add to existing scholarship. What do we learn that has not been known before?

Taking into consideration existing scholarship on digital methods in film and media studies, the paper considers the opportunities and challenges related to the development of a demonstrator-tool to explore, analyze and curate the Jean Desmet collection. The Jean Desmet collection consists of four different (audio)visual datasets: 900 films, 2000 posters, 700 vintage photographs, and 120.000 scans of the company archive. All datasets are related to the film distribution activities of Jean Desmet, thus chiefly centered around the same set of films and their screenings in movie theatres. The company archive, then, sheds light on the business activities of Desmet and all films that did not make it to the big screen. By developing a demonstrator-tool, we aim to tackle the heterogeneity and diversity of the collection by a range of computational techniques and methods. First, we provide the user with overviews of the composition of the collection, by means of time-based visualizations (e.g. topic streams), network-visualizations and/or geographical visualizations. Second, we explore how we can enrich the metadata of the four collections, for instance by linking them to other databases such as IMDb, Wikipedia and the historical film distribution database Cinema Context (cf. Dibbets 2011). Third, we look at ways in which we can visualize the gaps in the archive, and provide the user with the necessary information needed to analyze the collection. In addition, we try to automatically analyze the content of the cultural objects in the collections (films, posters and photographs), in the same vein as Manovich' Soft Cinema approach (e.g. Manovich & Kratky 2005), in order to draw links between formal (e.g. colour, style) and textual aspects of the cultural objects and the available metadata. Importantly, the demonstrator is developed by a multidisciplinary team of archivists, software developers and film scholars. Together, they explore the theoretical, technological, methodological and film historical possibilities of the Jean Desmet collection.

The research questions of film scholars do not only instigate but continuously feed the development of the tool, highlighting the importance of a user-centred approach in the development of research tools for the humanities.

References

Blom, I. (2003) J*ean Desmet and the Early Dutch Film Trade.* Amsterdam: Amsterdam University Press. Dibbets, K., "Cinema Context and the genes- of film history", in N*ew Review of Film and Television Studies*, 8.3 (2011) pp. 331-42. Maltby, R., D. Biltereyst, P. Meers. (2011) E*xplorations in New Cinema History: Case Studies and Approaches.*Chichester, Wiley-Blackwell. Manovich, L. & Kratky, A. (2005) *Soft cinema: navigating the database.* Cambridge MA: MIT press. Manovich, L. (2013) Virtualizing Vertov. Unpublished paper, available online: Http://softwarestudies.com/cultural_analytics/Manovich.Visualizing_Vertov.2013.pdf. Pafort-Overduin, C. (2009) Compartmentalisation and its Influence on Film Distribution and Exhibition in The Netherlands, 1934-1936. In D*igital Tools in Media Studies. Analysis and Research. An Overview*, M. Ross & M. Grauer (eds.), 55-68. Bielefeld: Transcript Verlag. Tsivian, Y. (2009) Cinemetrics, Part of the Humanities' Cyberinfrastructure. In D*igital Tools in Media Studies. Analysis and Research. An Overview*, M. Ross & M. Grauer (eds.), 93-100. Bielefeld: Transcript Verlag.

Henderson – The EVIA Digital Archive Project: A Time-Based Media Annotation and Online Access System for Digital Humanities Research, Teaching, and Collaboration

This presentation will discuss how the EVIA Digital Archive Project (www.eviada.org) has addressed the needs of digital humanities scholars for annotating, analyzing, and disseminating time-based media. I will focus on the ways scholars engaged in ethnographic fieldwork and sound and moving image performance analysis have used the project's tools and online access system for research, teaching, and collaboration. I will also touch on the ongoing challenges of publishing and providing

long-term online access to annotated video collections, and will provide a brief demonstration of the project's desktop digital video annotation tool and online search and browse delivery system.

Scholars such as ethnomusicologists, folklorists, anthropologists, and dance ethnologists, make a large number of audio and video recordings when conducting ethnographic fieldwork. Finding tools and systems for the preservation, analysis, and dissemination of these recordings has been a constant challenge. To meet this need EVIA Project staff and contributors have created a workflow for preserving the ethnographic field video recordings created by these scholars, and have developed a suite of software tools for digital video annotation, online collection searching, technical metadata collection, peer review, and controlled vocabulary and thesaurus maintenance as well as an online delivery system to provide free website access to these richly annotated recordings.

The presentation will illustrate how digital humanities scholars are using the project's video annotation tool in their research, from a choreographer, dance ethnographer, and digital archivist who are collaborating on a project to document, preserve, and archive each component of a substantial choreographic work, to an ethnomusicologist collaborating with interlocutors in the field when annotating her video recordings of musical performances. I will also discuss how scholars have used the EVIA online delivery system to analyze and compare text-based ethnography to media-based ethnography, and hypermedia ethnography; to study fieldwork methods and the documentation of ethnographic field sites; to analyze cultural performances; and to link their published EVIA collections with their scholarly monographs.

Since the launch of the EVIA website in 2009, and the end of external funding support the same year, the project has been dealing with a number of ongoing challenges involving funding; technological obsolescence; project maintenance; upper level administrative support; depositors' time needs for annotating ten hours of digital video; and online publication delays caused by the peer-review process. I will briefly comment on these challenges, the various solutions we have employed to address them, and the remaining challenges for which we are still seeking solutions.

EVIA Project summer institutes have been the principal channel through which individuals deposit ethnographic materials in the archive. After an application and vetting process, selected institute scholars sent ten hours of their fieldwork video recordings to us for preservation transfer and the preparation of digital access versions of their media for annotation. These scholars then spent an intensive two-weeks at Indiana University Bloomington, working with the EVIA Project's video annotation tool, the "Annotators Workbench" (AWB), an open source software platform developed for assembling a corpus of digital video files, segmenting them using a three-level hierarchical scheme (event, scene, action), and annotating each segment.

Using the AWB scholars create complex and nuanced analyses for time-based media not possible with other multimedia platforms or text publications. Its annotation capabilities are ideally suited to the analysis of time-based media because depositors can cue scholarly annotations to specific events in the recordings as they play, and, up to the second, can very precisely tie written descriptions and analyses to sound and moving images. Completed collections are peer-reviewed and published to the project's online search and browse website.

In other online digital humanities publications which link descriptive text with audio and video files, media examples are secondary to the descriptive and analytic text. By contrast, the EVIA Project search and browse website is a standalone digital media publication that foregrounds video recordings in conjunction with description and analysis. End users may view a sound and moving image collection's descriptive and analytical annotations time-coded to each segment of video content as it plays and can search published collections using keywords and controlled vocabulary faceted searches; access every

annotated action, scene, or event in the recordings; create a playlist of various examples for class instruction; and email links to specific scenes of interest. Intellectual property issues are handled through an end user license agreement, by permission documents, and through an AWB function that enables depositors to block end users from viewing culturally or politically sensitive material.

With this time-based media annotation tool and delivery system, the EVIA Project has endeavored to facilitate the engagement of digital humanities scholarship with sound and moving images, uniquely combining preservation, annotation, and online scholarly publication to meet the needs of ethnographic scholars working with digital video recordings.

Sanders & Hagedoorn – How to publish AV research online

Although various tools and functionalities are currently being developed to search, explore, and analyse digital audio-visual sources, this paper will focus on the next step: the question how to publish such Digital Humanities research online. How can Digital Humanities research with audiovisual materials benefit best from the affordances of online publication? This paper will discuss various scenarios for the online publication of Digital Humanities research. It will draw from both existing online publications, by scholars and by media professionals, and scenarios newly developed as part of a university course on doing television history online. The results will then be applied to a selection of proposals within the workshop.

The use of digital material – either digitised or born digital – within Humanities research is increasing, and has initiated a wide range of initiatives to support research and analysis of digital materials, including audio and/or video materials. Contemporary projects include the aggregation of digital collections (e.g. EUscreen, Europeana Newspapers); the development of tools and functionalities for searching and exploring such collections (e.g. AVResearcher, TROVe); and tools to analyse digital materials (e.g. AXES for image recognition). Furthermore, Digital Humanities scholars increasingly use the internet as a platform for dissemination, for instance, through the open access multi-media journal*VIEW*, the online scholarship platform In *Media Res*, and *Vectors*, which operates at the intersection of culture, creativity, and technology. With this in mind, we continue the process and consider the next step: the publication of DH research with AV materials. We address the question how the publication of such research can benefit best from the affordances of the internet. What kind of tools, functionalities, and designs best serve its online publication? To answer such questions, this paper will analyse and discuss various scenarios for the online publication of Digital Humanities research.

The paper will first discuss two existing examples of online publications as best practice examples: the online journals *VIEW* (http://journal.euscreen.eu) and *Vectors*(http://vectors.usc.edu). It will then discuss online media projects, as the presentation of interactive narratives is increasingly being developed by filmmakers, media artists, and journalists. Like scholars, these makers publish interactive narratives generally presenting an argument and using a variety of sources and source materials, including videos, documents, and photographs. Examples which are analysed and will be discussed as best practice include the Japanese interactive documentary *Soul-Patron* (http://www.soul-patron.com), the Dutch television series *Na de Bevrijding* [After the Liberation] (http://www.nadebevrijding.nl), and The New York Times interactive long read *Snow*(http://www.nytimes.com/projects/2012/snow-fall/#/?part=tunnel-creek). While such narratives may not seem the most obvious examples for the online publication of scholarly work, analysing them can be useful to consider how to narrate DH research with AV material, how to navigate various sources of information, and how contextualize such research narratives.

The EUscreen project aggregates archive materials, mostly in the form of videos, of European broadcasters, and makes these available for the general audience as well as for teachers and researchers. One of the challenges the project is currently facing is the development of tools and functionalities to provide context for these materials. Within the framework of the EUscreenXL project (2013-2016), ways to publish research including EUscreen materials are being investigated. The second part of the paper will discuss scenarios newly developed by students within a course using EUscreen. In Spring 2014, a group of 27 bachelor students at Utrecht University participated in a course on Dutch television culture conducted by Sonja de Leeuw, Eggo Müller, Berber Hagedoorn and Willemien Sanders. Students were required to research Dutch television history in an international context using the portal EUscreen and other online archives and collections, such as Academia (provided by the Netherlands Institute for Sound and Vision), Open Images, as well as YouTube. Moreover, they were assigned to develop scenarios for the online publication of their research and to write a research paper accounting for the development of their scenarios and designs. For example, some students developed a timeline as the structuring device for the narrative, some used geographical mapping as the main point of entry, whilst others used the idea of an online museum with various exhibition rooms as a narrative structure. A number of the scenarios developed in the course will be discussed as best practice examples.

The various examples of online publications including AV materials discussed above will inform the distillation of a number of prototypical formats for publishing Digital Humanities research online.

In the last part of paper the findings of the analysis of various publications discussed will be applied to a selection of the proposals within the workshop, focusing on the ways to narrate research, ways to navigate sources, and ways to contextualize the narrative. The results will be discussed and will subsequently inform the development of publication formats and related tools and functionalities, for publishing DH research including AV materials online within the framework of the EUscreen project (http://www.euscreen.eu).

Stępińska & Wyszyński – Content Analysis System for Television (CAST): more than an archive, less than an assistant (retired)

Several decades ago Berelson (1952) suggested five main purposes of content analysis, that is: describing substance characteristics of message content, describing form characteristics of message content, making inferences to producers of content, making inferences to audiences of content, and predicting the effects of content on audiences. Both, social science and humanities use content analysis for some of these purposes: while the former is primarily concerned with the effects that media messages may have on audiences, the latter tries to analyze how the media content reveal 'truths' about a society and culture. Despite of the methods used during content analysis, either quantitative or qualitative, scholars still face the same challenge, that is how to record and store audiovisual data. In particular, for longitudinal studies monitoring media content, collecting a significant amount of digital material seems to be a stumbling block since it requires not only an access to the sources of the data (TV stations), but also a digital repository.

The aim of this paper is to identify obstacles for the integration of AV sources in the Digital Humanities. In particular, our goal is to present Content Analysis System for Television (CAST), owned by the Faculty of Political Science and Journalism at the Adam Mickiewicz University in Poznan (Poland), and discuss challenges and problems that we face while employing this system to our research.

The system was designed to serve not only as an archive of a TV content, but also as a digital tool that might be applied in a process of coding and analyzing audiovisual data. The system allows us to collect and record a content of approx. 20 TV channels (Polish and foreign) and store it for a period of 5 years.

Speech found in videos is transcribed into text. One of the especially interesting features is an option of creating a codebook and collections of clips that can be coded within the system. In practice, a coder is able to watch and code at the same time, thanks to a extremely flexible user interface. The software also provides an opportunity to a frame by frame video preview, finding repeating video sequences (e.g. commercials), as well as an extensive search of all recorded material. The features mentioned above make the system an exquisite archive. However, when we plan to convert the system into a more 'assistant' mode, we see not only advantages, but also obstacles. The system is powerful and let us reach amounts of data that we could not process otherwise, but human analysis can be more precise and necessary in different stages of the research process, including the selection of data, cleaning up of the sample, and analysis of the content. However, the risk of using big data is that insights might be limited in the absence of theoretical frameworks that produces strong research questions. Previous studies also highlighted the weaknesses of big data analysis in terms of accuracy and interpretation. Nevertheless, we would like to emphasize an emerging productive relationship between research based on "small" and "big" data.

Pagenstecher – Mapping Testimonies. The Interview Archive "Forced Labor 1939-1945" (retired)

Digital technology offers new opportunities for Oral History: Instead of reading transcripts, the historian may listen to the original audio or watch the video file. But the task of preparing large, multilingual interview collections for the creation of searchable online archives poses a number of technical and methodical challenges, that will be discussed using the example of the online archive "Forced Labor 1939-1945. Memory and History".

The "Forced Labor" archive commemorates more than twenty million people who were forced to work for Nazi Germany. Bringing together almost 600 personal narratives from 26 different countries, it is one of the most sophisticated multilingual interview archives. As a digital memorial it aims at a transnational culture of remembrance.

The online platform at http://www.zwangsarbeit-archiv.de supports the combination of biographical and geographical approaches in research and in education. 590 former forced laborers, prominent survivors as well as neglected victim groups, tell their life stories in detailed audio and video interviews. Since 2009, the audiovisual media can be consulted online, together with time-coded and indexed transcripts, German translations, short bio-graphies, photographs and other documents.

A detailed mapping function locates the survivors' experiences of camps and factories. Together with various search categories and a full-text search through the media files, it enables a targeted, direct retrieval of individual interview passages. On the other hand, the online environment is designed to "combat" the use of the interviews as a mere quarry for quotations: A table of contents, a biography and the user interface design aim to support an integral understanding of the whole testimony in its narrative structure and its biographical meaning.

Based on the archive, an educational DVD, regional projects with memorial museums and smartphone apps have aimed at situating survivors' testimonies within localized remembrance cultures and class rooms. The creation of multilingual approaches for mobile devices is underway.

Apart from presenting the public online archive, the presentation will demonstrate the internal, collaborative indexing platform. It will reflect the project's strategies in preserving, analyzing and presenting individual memories in a transnational digital context, and discuss future perspectives of interview archives within Digital Humanities.