

LOOKING FOR WHAT YOU ARE LOOKING FOR

A MEDIA RESEARCHER'S FIRST SEARCH IN A TELEVISION ARCHIVE

Jasmijn Van Gorp

Department of Media and Culture Studies
Faculty of Humanities
Utrecht University
Muntstraat 2A , 3512 EV Utrecht
The Netherlands
j.vangorp@uu.nl

Abstract: In this essay, the author reflects on her first search with the online search system of the Netherlands Institute for Sound and Vision. This was part of a pilot study on how media researchers use the audiovisual archive. Her search was being logged, video taped and sound recorded, she had to 'think aloud', and all of this in the presence of a fellow researcher from computer sciences who observed her search behaviour. By showing how she found some relevant programmes among more than 1.2 million items, this article illustrates how archival *finding* by media researchers can be understood as archival *looking* or 'exploratory search'.

Keywords: search system, exploratory search, archive, Russia, Dutch television

Three years ago I worked for the first time with [iMMix](#), the online search system of [the Netherlands Institute for Sound and Vision](#). Given the clumsiness and inefficiency of that first search, I would have preferred to forget about it. Let alone write an article about it. However, it was part of a pilot study on how media researchers use the audiovisual archive. For that reason, the search was being logged, videotaped and sound recorded, I had to 'think aloud', and - to make the surveillance complete - all of this in the presence of my fellow researcher of computer sciences who observed my search behaviour. Now, three years later, I am ready to take the recordings off the shelves and turn them into objects of analysis. I apply a user-centred approach to the digital audiovisual archive by showing how a media researcher interacts with it and ultimately, discovers relevant programmes among more than 1.2 million items. This article illustrates how archival *finding* by media researchers can be understood as the more processual archival *looking* or 'exploratory search'. Media researchers, in their first stage of research, have usually no specific title in mind, but mix searching and undirected browsing, jumping to related items so as to explore an entity, theme or event.¹

¹ See Lunn, Brian K., 'User needs in television archive access: Acquiring knowledge necessary for system design', *Journal of Digital Information*, 10(6), 2009, and Bron, Marc, Jasmijn Van Gorp, Frank F. Nack, Maarten de Rijke, Andrei Vishneuski, and Sonja de Leeuw, '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 Oregon, ACM, 2012a

They depend heavily on serendipity – stumbling upon ‘lucky accidents’ they were not necessarily looking for. Personal memory or prior knowledge plays an important role in this search process.² The user deploys his/her memory to be able to formulate key words, but also to contextualise the programmes retrieved and ultimately, to make a selection of programmes for further scrutiny.³ Moreover, prior knowledge and the process of remembering is also related to the search system itself. Users learn how a system works by using it, mediating the way one interacts with it.⁴ In this practical case study, I will show what role personal memory plays in a search that was set up as a user study for a larger project. Although the research setting felt more like a public performance on national television including accompanying nerves, it turned out that my search was quite prototypical for the way media researchers search the digital audiovisual archive.

1 Setting The Scene

It is the morning of February 3 2010. The online search system of the Netherlands Institute of Sound and Vision, [iMMix](#), consists of metadata descriptions of around 1.2 million television and radio programmes of Dutch public broadcasters, provided by professional annotators.⁵ It has a simple and an advanced search option that allows users to search on specific fields. After the search results have been displayed, they can be refined by a filter system. However, the search system is developed for an experienced group of users: the broadcast professionals. Broadcast professionals, in their need to reuse material, often know what they are looking for, using a directed search, for instance, by providing a title or specific content.⁶ Consequently, iMMix mainly supports ‘known item’ search and can be considered as a ‘known item search system’. Media researchers, however, would benefit from an ‘exploratory search system’, a system that supports them in their browsing and jumping from one topic to the other.⁷

The aim of our [BRIDGE](#) project is to develop and test new ‘exploratory search systems’ that support media researchers in their exploration of the audiovisual archive.⁸ BRIDGE is a collaboration between [Intelligent Systems Lab Amsterdam](#) (University of Amsterdam), [Centre for Television in Transition](#) (Utrecht University) and the Netherlands Institute for Sound and Vision. For the first study of BRIDGE, we set out to investigate how ‘the known item search system’ iMMix can be used for an exploratory search task in order to compile a list of features that would be desirable for the new ‘exploratory search system’. As test person, we were looking for a media researcher with no or just a little experience with the search system and someone with a ‘limited’ prior knowledge of the archival content in the archive, so s/he would have to rely heavily on the search system for support.

2 See for instance Wendy M. Duff, Emily Monks-Leeson, and Alan Galey, ‘[Contexts built and found: a pilot study on the process of archival meaning-making](#)’, *Archival Science*(12), 2012, p. 69-92, 2012. For the role of personal memory in archival search of photographs, I refer to the work of Martijn Kleppe, e.g. M. Kleppe, ‘Vind foto’s met behulp van foto’s’, *Fotografisch Geheugen* (72), 2011, p. 15-17

3 See for the specificities of media and humanities researchers’ search in the archive: Lunn 2009, and Franciska M.G. de Jong, Roeland Ordelman, Roeland J.F. and Stef Scagliola, ‘[Audio-visual Collections and the User Needs of Scholars in the Humanities: a Case for Co-Development](#)’, Proceedings of the 2nd Conference on Supporting Digital Humanities (SDH 2011), Copenhagen, Denmark, Centre for Language Technology

4 See for instance Christopher Hölscher and Gerhard Strube, ‘Web search behavior of Internet experts and newbies’, *Computer networks*, 33(1), 2000, p. 337–346. The study found that search experience can influence search behaviour, as more experienced users plan more and as such, are more goal-directed.

5 For more information on iMMix, the Netherlands Institute for Sound and Vision and its collection, see Sabine Lenk, ‘Images for the Future: Will They Make TV Scholars Happy?’, *Critical Studies in Television: The International Journal of Television Studies*, 5(2), 2010, p. 80-85

6 Bouke Huurink, Laura Hollink, Wietske van den Heuvel and Maarten de Rijke, ‘Search behavior of Media Professionals at an Audiovisual Archive: A Transaction Log Analysis’, *Journal of the American Society for Information Science and Technology*, 6(16), 2010, p. 1180-1197

7 For a technical discussion on exploratory search systems, see Ryan White, Gheorghe Muresan and Gary Marchionini, eds., ‘[Evaluating Exploratory Search Systems](#)’, SIGIR 2006: Workshop of 29th International ACM SIGIR Conference on Research and Development in Information Retrieval, Seattle, 2006. For a study on how non-professional searchers use a professional search tool, see Marc Bron, Jasmijn Van Gorp, Frank F. Nack, Maarten de Rijke, ‘[Exploratory Search in an Audio-Visual Archive: Evaluating a Professional Search Tool for Non-Professional Users](#)’, EuroHCIR 2011: 1st European Workshop on Human-Computer Interaction and Information Retrieval, Newcastle, 2011. Also interesting in this respect is the comparison between the search system of the National Film and Sound Archive of Australia and YouTube as analysed by Alan McKee, ‘YouTube vs the National Film and Sound Archive: which is the more useful resource for historians of Australian television?’, *Television and New Media*, 12(2), 2010, p. 154-173

8 See the BRIDGE-papers: Bron et al, 2011; Bron et al, 2012a; Marc Bron, Frank F. Nack, Maarten de Rijke and Jasmijn Van Gorp ‘[Ingredients for a User Interface to Support Media Studies Researchers in Data Collection](#)’, EuroHCIR 2012: 2nd European Workshop on Human-Computer Interaction and Information Retrieval, Nijmegen, 2012b; Marc Bron, Jasmijn Van Gorp, Frank F. Nack, Maarten de Rijke, and Lotte B. Baltussen, ‘Aggregated search interfaces in multi-session tasks’, SIGIR 2013: 36th international ACM SIGIR Conference on Research and Development in Information Retrieval, Dublin, ACM, 2013, forthcoming

Surprisingly, I was that one person who met both requirements. I am a postdoc in television studies who only partly shares a Dutch cultural background: I am a Dutch speaking Belgian who has recently moved to the Netherlands. If a Belgian has to search a Dutch television archive, one would expect that her memories of programmes are only sparse and that her way of defining new keywords is impeded by a limited knowledge of Dutch popular culture and history. Combined with the fact that I never used the search system before, this made me an ideal subject for a pilot study on the search behaviour of media researchers. In this article, I look back at my struggles interacting with the system iMMix.

For the following account, I use all materialized memories of that sunny day in February 2010 to reconstruct the experience: the notes I scribbled down during the search, the data log, the subsequent interviews conducted by my fellow researcher and the video of the computer screen with the voice-over of my spoken thoughts.

2 Running Out Of Keywords, Or How I Ran Into A Russian

In order to *start* a search, I have to define a topic. I decide to focus on my academic geographical area of interest: Russia and, more specifically, on the 'representation of Russians on Dutch television'. Once arrived in the building of the Netherlands Institute for Sound and Vision, I log into the computer with an account that allows streaming content, while a camera behind me records my actions and a microphone captures my comments.⁹ My fellow researcher from computer sciences installs a log program. Let's read the recordings.

00:00:00 ([see video 1](#)) 'The first thing I do is explore the archive', I comment, still confidently about my search. I start with the simple search option of the interface, uncheck the box of sound recordings so as to only retrieve television programmes. My first keyword 'Russian' ('Rus') results in 1593 items, which is 'way too much' to browse. As I want to specifically look for fictional programmes about Russians, I try to find ways to filter out news broadcasts, or alternatively, cluster all other fictional genres. I select all 18 fictional genres in the filter bar, push on the 'refine' button and no results are found. What I do not know is that the filter does an AND-, and not an OR-search. Naturally, there are no programmes belonging to *all* of the 18 genres I selected, so I fail.

00:03:15 I try to find an alternative method to filter items, using the advanced search version of the search system. I type 'Russian' in box 1 and type genre 'comedy' ('komedie') in the genre field of box 2, not knowing that, in contrast to box 1, the additional three boxes of the advanced search option contain a drop-down menu in which you have to *select* a term instead of *typing* it. Consequently, none of my attempts to filter by using an additional search box succeed.

00:06:21 Accidentally, I automatically select 'television drama' in the genre field, and iMMix retrieves two television programmes. I do not notice the selection, possibly because I am too preoccupied with the research topic. One of the two programmes is *Televisierechtbank* (*Court on television*). I expect that 'one of the characters in court will be a Russian, in line with their stereotypical representation as gangsters in films'.¹⁰ However, while opening the programme description, it turns out that one of the witnesses in the court is called mister K. Rus. I fail again.

00:08:21 Desperately looking for alternative ways to filter, I decide to enlarge the search string with the article 'a' (i.e. 'a Russian'), this time with success: only 21 programmes retrieved, a number small enough to browse. The majority of the items are news broadcasts however. I paginate and on page three I find a youth

9 iMMix for external users only allows searching on the metadata of programmes. I used iMMix for internal users (Extranet) that requires a login and allows streaming content and provides key frames of the programmes.

10 See for instance Barbara Korte, Eva Ulrike Pirker and Sissy Helff, eds., *Facing the East in the West. Images of Eastern Europe in British Literature, Film and Culture*, Editions Rodopi, 2010

programme *Waskracht* (2001) that is 'entirely devoted to Russia'. It contains very stereotypical topics: Russian dating services and vodka drinking men. I write it down as part of my selection, but do not watch it until I have selected all programmes. 'Searching and selecting is something else than watching', I explain to my colleague, 'watching is analysing'. I also select a programme called *Future Express* ('Rusland houdt u wel van ons?', transl. 'Russia, do you really love us?' episode 11, 2009), which is a travel programme about railroads around the world. In this episode, the reporter travels by train from Saint Petersburg via Novosibirsk to Chabarovsk in Russia. It matches one of my long existing ideas for a new research article: a study on the representation of Russia as a country in Dutch travel programmes. However, I forget to write down the name of the programme.

00:16:50 I have already found two programmes but 'have the feeling that there is yet more to discover'. I decide to look for items containing 'Russia', but am afraid that too many results will come up.

In the meanwhile, my fellow researcher observes that I continuously use the same keywords, 'Russia' and 'Russian', not thinking of any variations. In the post-test interview, surprised by his observation, I explain my limited imagination by saying that 'I was determined to find a Russian character in a fictional programme, and thought it was the fastest route to success. I deliberately did not coin more specific keywords, such as Moscow, to prevent me from missing something.' I wanted to have an overview about the topic that was as wide as possible. This obviously contradicted my need to find programmes in a small result list that is easy to browse. I continue apologising by blaming the setting: 'my mind did not work as usual due to the time limit and pressure of the test'. The observational setting is indeed stressful: after our first large-scale user study with iMMix the 22 freshmen students testified that they were very nervous due to the setting and the invigilation.¹¹ The question then is whether I would have searched differently if there was no surveillance and if I would have mastered the search system sooner with the help of a tutorial video, for instance.

3 Finding what I already know, or how I finally understood how the system works

00:18:23 Still thinking about ways to reduce the result list when searching on 'Russia', I decide to use some of my prior knowledge. I remember some non-fiction and entertainment programmes about Russia by the Dutch public broadcaster VPRO, so I automatically select 'VPRO' in box 2. The system renders the large amount of 504 results. I want to clear all fields and search again until I notice a programme called *Express VPRO Ruslandgangers* ('Express VPRO Russia-visitors', 1983). Excited about the title that sounds very much like a travel programme, I read the programme description. To no avail: it is a radio programme. Apparently, I had forgotten to uncheck the sound-recording box and thus, have to start over. 266 items are found.

00:20:43 ([see video 2](#)) I try to figure out, again, how to exclude news programmes, I'm still not mastering the filter menu. 'I am stuck, unless I take the time to browse through the 266 programmes one by one,' I think aloud. I again try to filter a genre by using the additional box 'entertainment' ('amusement') this time, with no results since I did not select it. My next attempt to decrease the number of items is checking the box for 'full broadcasts', which I apparently do automatically. A lot of other actions remain unnoticed by myself, unless my fellow researcher makes me aware of them ('Why are you doing this?'). The 'full broadcasts'-manoeuvre reduces the amount to 170.

00:21:30. I decide to browse the 170 items on 'Russia' and 'VPRO' and notice a programme that I already know: *Van Moskou tot Magadan* (*From Moscow to Magadan*, 2009), a series of 7 episodes with reports from the Dutch journalist Jelle Brandt Corstius' journey across Russia ([see clip 1](#)). It is, admittedly, one of my favourite programmes. I also notice another programme that I had seen when it was broadcast in 2008 but forgot about: an episode of the historical series *Tegenlicht* ('Under the Spotlight') about the Russian chess

¹¹ See Bron 2011

champion Yuri Kasparov. Using 'VPRO' as keyword, then leads to a bias towards my own preference, or finding what I already know. It is tempting to conduct research on programmes you already know. 'There should be more', I express my thoughts, hoping that the search system directs me to new treasures.

00:22:30 I browse the 170 items, while trying to remember the result list of my previous searches. I start to observe a trend in the data: there are many recent items, with a peak in the late 1980s and early 1990s. I check whether it is caused by a chronological ranking of the results, but that is not the case. 'It is very interesting', I explain, 'that the majority of the programmes are broadcast in the period around the collapse of the Soviet Union. Apparently, the end of the Cold War affects the frequency of representation of Russians and Russia on Dutch television'.

00:27:34 It has taken a while before I mastered the search system. 27 minutes to be precise. Now that I finally do understand the system, I start a new search on 'Russia' and 'VPRO', and at page seven I find a programme called *Nieuwe Maatjes* ('New Friends', 1987), about the life of a Russian migrant girl that lives in the Netherlands. Until now, I was thinking of conducting research on Dutch travel programmes about Russia, but my project takes a u-turn. Why not focus on Dutch programmes about Russian migrant children? 'Find some similar programmes', I instruct myself. For the first time, I manage to use the filter on the left side in which all tags are mentioned with their frequency in between brackets. I see 'children' with frequency nine, check that box only, click on the search button and one of the first results retrieved is *Circuskinderen* ('Circus children', 1992) about the Russian girl Angela.

4 Tracking Your Search, Or How To get Back On Track

Besides struggling with the search system, I also have difficulties memorising which keywords I have already used and which programmes I have already selected during the search. I use a pencil and paper to scribble down my thoughts and selected programmes for further scrutiny. The iMMix system provides a 'favourites' option, but because of either nerves or enthusiasm I do not spot the button. And, to make matters worse, while writing this article and checking out iMMix again, I notice a 'history' button as well. One advantage of the observational method is that my fellow researcher of computer sciences is sitting next to me. While I should follow the protocol to simply think aloud, I use him as a helpdesk ('Oh, how can I open the advanced search option again?') and as external flash memory ('Oh, which keyword did I use?').

00:35:21 Now that I finally master the use of the filter menu by just selecting one instead of many filters, I start looking for 'Russia' and 'satire'. And, yes, again a discovery: this time four episodes of *Keek op de Week* ('View on the Week', 1990-1991) in which comedians [Kees van Kooten](#) and [Wim de Bie](#) play different stereotypical characters. In the first episode (25-11-1990) an expert on Eastern Europe, dr. R. Clavan, comments on the food aid for Russia. One week later (02-12-2009), Gerrit Braks, a fictional anchor, comments about the content of food parcels for Russia, while in the next episode (09-12-1990), a former German teacher complains about the food aid to Russia. One year later, the food aid story line is dropped in favour of a sketch on a Dutch trader and his exhibition of 'tablecloth art' in Russia.

00:45:32 ([see video 3](#)) I am convinced I will not find anything else. I decide to use a different strategy and narrow down my search to the combination 'Russian' and 'crime' just to see what it will retrieve. I am immediately triggered by the programme *Future Express*, but do not remember that I have already found it 40 minutes before. I try to justify this carelessness by saying that 'I probably found it not relevant then, but now I do', which does not sound very convincing to myself either (in the interview later I admitted that 'I was too enthusiastic to write it down properly'). I get stuck again, commenting that 'I am running in circles'. The combination 'Russia' and 'crime' paves the way to *Factor*, with an episode named *Moordende Meiden*

(Murderous Girls, 2003), about a re-education camp in Belgorod, Russia, and female prisoners talking about the crimes they committed. I see a link with *Waskracht*. 'It is a reversed cliché of the items of *Waskracht*, about Russian dating services and vodka drinking men', I explain to my colleague, 'One would expect that crime is mostly linked to men, not to women'.

01:17:00 Well, what to do now? 'It is a dead end', I conclude, while overlooking my harvest. I am exhausted and sigh 'It is so important that BRIDGE will develop another search system because with iMMix you have to be so conscious about which boxes are checked and unchecked that you have no mental space left to think about your research question'. My analysis, however, shows the opposite: I manage to process what I find and to draw links between programmes that distract me from my actual search actions and from learning the particularities of the search system. Consequently and ironically, the fact that I am 'too much' focused on the content prevents me from finding more programmes. Time to watch the programmes.

5 Towards A New Search System

01:55:23 I am finally finished watching. Some programmes can not be linked to others, like *Tegenlicht*, so I leave them out of the selection. My mental linking process throughout the search has resulted in three clusters of programmes for further scrutiny: two travel reports (*Future Express*, *Van Moskou tot Magadan*), two programmes about migrant children (*Nieuwe Maatjes*, *Circuskinderen*), two current affairs programmes about stereotypes or counter-stereotypes (*Waskracht*, *Factor: Moordende Meiden*). And then, a separate category: satire (four episodes of *Keek op de Week*). Only while writing this article, I notice that 9 out of 11 programmes are broadcast by VPRO. It may indicate that VPRO is the broadcaster with the largest amount of Russia(n) programmes, yet it is equally possible that my prior knowledge has led me towards these items. What would have happened if I had not taken VPRO as a keyword, but used, for instance, the time period 1980s-1990s? Would I have ended up with an entirely different selection, if any? Although during the search I made apologetic comments about my slow understanding of the search system, my chaotic mindset and nerves, I am satisfied with the selected programmes. In addition, I discovered new programmes, such as *Waskracht* and *Keek op de Week*. I also managed to gain an overview, spot trends and a peak in television broadcasts situated around the time of the collapse of the Soviet Union.

This reflection on my first search shows that the road to discoveries is not wholly coincidental but rather an interplay between the search topic, the user's memory, the particularities of the search system, and the research setting. The user's personal memory not only plays a role in defining keywords, but also in the mental linking process during the search. In this case study, it is the activity of the user that highlights the dynamics at play in a digital archive: the exploratory search encapsulates the dynamics of browsing, creating meaningful links between nodes that make up what Wolfgang Ernst calls the 'dynarchive'.¹² 'We are looking for what we are looking for', one interviewee explains, referring to the process of exploring the archive on the basis of which media researchers compile their research corpus.¹³ It is a process of drawing links between programmes while also keeping the larger picture in mind.¹⁴ On a technical level, this pilot study shows that a search system aimed at supporting exploratory search should offer an overview at a single glance, provide inspiration for keywords and support remembering the search and selection process. What would happen if researchers would use such a search system designed for exploratory search when exploring the audiovisual archive?

Based on this pilot study and a large-scale user study with 22 students, the BRIDGE-project developed an

¹² See chapter 'Underway to the Dual System: Classical Archives and Digital Memories,' in: Wolfgang Ernst, *Digital Memory and the Archive* (edited and with an introduction by Jussi Parikka), University of Minnesota Press, 2013, p. 81

¹³ See Bron et al. 2012a.

¹⁴ de Jong et al. define this larger picture as a search 'directed towards a collection as a whole, in which case an entire dataset is the focus of attention'. See de Jong 2011

exploratory search system for the Netherlands Institute of Sound and Vision, called MeRDES (Media Researchers' Data Exploration Suite). MeRDES was launched in 2012 and incorporates two side-by-side versions of a standard exploratory interface and has three special features: word clouds to give inspiration, comparative visualisations to indicate trends, and a bookmark option and search history to keep track.¹⁵ MeRDES forms the baseline for other search systems such as CoMeRDA, AVResearcher, TROVe and QuaMeRDES, currently under development at the Netherlands Institute for Sound and Vision. Please visit the [blog of their Research and Development department](#) for the latest news. A demo version of MeRDES is available [here](#).

Acknowledgements

The author would like to thank Marc Bron for co-conducting this study and for his continuous support, and Lotte Baltussen, Johan Oomen, Bouke Huurnink and Evelien Wolda for their feedback on earlier versions of this article. This research, part of BRIDGE, has been made possible by the [CATCH](#) programme of the Dutch Scientific Organisation NWO.

Biography

Jasmijn Van Gorp is a postdoc in television culture at Utrecht University, the Netherlands. Her research is situated in the fields of media and cultural studies with a focus on Eastern Europe. She received a PhD in Social Sciences from Antwerp University (2008) and has been a visiting scholar at the Russian Film Institute in Moscow and at the Comparative Media Studies program at MIT.

¹⁵ For more information on MeRDES and the user studies we've conducted, see Bron et al. 2012a.