

11. Interstitial Data

Tracing Metadata in Archival Search Systems

Jasmijn Van Gorp

Abstract

Metadata do not merely give explicit information about records in the archive but can also be considered a source of information about the (historical) context in which they are created. This chapter combines the insights of critical data studies and archival studies to formulate a hands-on approach to tracing metadata in archival search systems. The approach, which builds further on Loukissas's local reading strategies, consists of two distinct phases: an exploration phase to trace and select and an analysis phase to trace and compare. The author concludes that a lot of data necessary to understanding metadata in search systems is hidden—different forms of what can be considered “interstitial data.”

Keywords: archive, metadata, television, search system, tracing, local reading

Metadata, or data about data, provide the essential context of a record in an archival collection (Kitchin 2014, 4). Similarly, Pomerantz (2015) refers to metadata as a map that represents the complexity of an object in a simpler form. Metadata can be automatically generated and/or created by humans, such as archivists or users (cf. Noordegraaf 2015). Without metadata, it would be virtually impossible to find a record in large-scale archival collections. First, they support the identification of any given record “at a glance.” Second, metadata are based on a logic and classification system that provides information of a given record within a collection. Metadata also make records retrievable through descriptions of the records.

However, metadata do not merely give explicit information about records. Within critical data studies and archive studies, it is common to consider metadata also a source of information about the (historical) context in which

they are created. In their discussion of new media art archives, Rinehart and Ippolito (2014) show how metadata can shape the historical record. Each metadata standard “frames”: it provides a point of view and determines “what we choose to remember and what we choose to forget” (Rinehart and Ippolito 2014, 60–62). Loukissas (2019) elaborates on how (meta)data can also serve as cultural markers of past collection practices. Data, he contends, can be “locally inscribed.” He proposes an active understanding of context and considers it to be assembled through a combination of social, technological, and spatial practices. Within archive studies, a similar perspective has been advocated by, among others, Eric Ketelaar (2001), who shows how archives reveal the context in which they are created through “tacit narratives.” With “tacit narratives,” Ketelaar means all practices and technologies that leave traces. Traces can be found not only in metadata that is available but also, or even more so, in what is not available, since “archiving also entails what should and what should not be kept” (Ketelaar 2001, 136).

It is this intriguing interplay between availability and unavailability that I will further investigate in this chapter. Using a case study that enforces what is at the margins, the “forgettable” broadcast genre of interstitials, I investigate the different traces of metadata.¹ To this end, I test and extend a proposed reading by Loukissas (2019, 62–69) by discerning two phases of “tracing metadata” as comprising a hands-on approach for studying metadata in archival search systems. While Loukissas focuses on all kinds of collections data in data infrastructures, I focus specifically on the function of metadata in the scholarly practice of searching in archival search systems. The two goals of this chapter are to investigate the extent to which traces of archival practices can be revealed through an analysis of metadata and, at a practical level, to help train students in critical reflection on the role of metadata when making use of archival search systems. This chapter concludes that talking about “interstitial data” is fruitful for encapsulating metadata’s various forms and related practices.

Trace and Select

Most archives have online search systems that are based on semantic searches with words matching the metadata of the records. The user defines

1 I borrow the concept of “forgettable television” from Polan (2013) who conceptualized it as “programming designed to be forgotten at virtually the very moment of its original viewing—as virtually most TV was” (347).

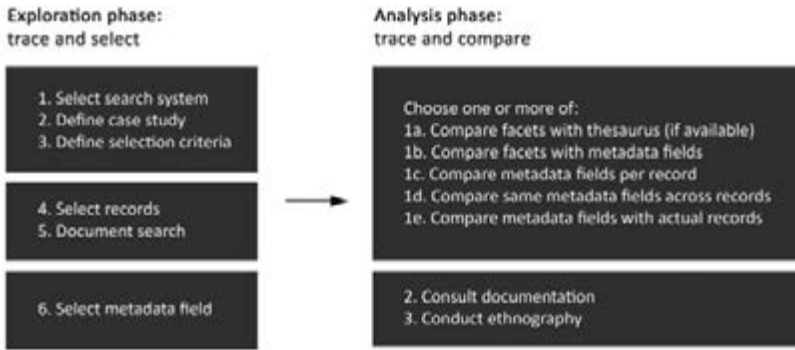


Fig. 10. The two phases of tracing metadata in archival search systems (image created by the author).

a query or conducts a “keyword search” (cf. Althaus and Phalen 2010), after which the system retrieves matching records in the search system based on their metadata and ranks the records in a result list. The metadata, as present in these search systems, are the object of my analysis. For the analysis of metadata, I outline a tracing method that consists of two distinct phases: an exploration (or “preparatory”) phase and an analysis phase. In the first phase, I explore a search system using keyword searches and trace my journey in the search system. The goal of this exploration phase is to select records for the analysis phase and to come into grips with the search system and with the relevant metadata fields.

The first step of the exploration phase is to select a search system. I selected a search system that I know very well, as I co-developed it: the CLARIAH Media Suite. The Media Suite is a new research infrastructure for audiovisual data in the Netherlands, aimed specifically at researchers (Melgar et al. 2018). It provides access to an extensive collection of the Netherlands Institute of Sound and Vision, containing almost two million television and radio broadcasts from public broadcasters. A second step is to define a case study with which to build a corpus. For my case study, I focus on what is seen as a “forgotten part” of television history: interstitials. Interstitials are the broadcasts that air in between the actual programs, such as trailers, commercials, and idents (Ellis 2012). As Johnson (2012) explains, archival access to interstitials is limited because they are often not saved, except in the case of full-day recordings, which are comparatively rare.

Once a search system and a case study are defined, an important step is defining a selection strategy for records and making this explicit (cf.

Creeber 2004). To achieve a balanced selection, I applied four search criteria. I wanted to select (1) various types of interstitials, (2) across a broad historical spectrum, which preferably have (3) access to the video and (4) are found by applying various search strategies, i.e., using different queries, filters, and settings for each record. I ended up with a selection of nine records: one commercial for STER for Duyvis tiger nuts; one introductory announcement of an old news program for the hearing impaired; one film announcement for a Saturday night film as part of the series *Jiskefet*; five television announcements of the fairy tale *Klaas Vaak*, with one original clip and four rebroadcasts of the same clip within a children's television program; and one full-day of recording of a Sunday on the channel NPO 1.

Tracing the journey in this exploration phase can be understood in the reflective sense or as “critical tracing” (van der Tuin and Verhoeff 2021, 197): walking back your own footsteps and reconsidering what you have found and why, by means of a critical reflection on your search and selection process. Systematic documentation or “logging” in research journals should be a standard phase of any research project (Borg 2001). Therefore, it is important to keep track of this exploratory phase by writing down one's keywords, filters, and settings. Many search systems have automated tools for this, such as history tracks and bookmarks. The Media Suite has bookmarks and URLs that reveal the tools used and the search terms.² In the method of tracing metadata, documenting helps the user not only to grasp semantic searches in search systems but also to find relevant metadata fields to analyze in the second phase. I consulted my research journal and noticed four metadata fields that were used extensively, triggering some questions: the descriptive metadata fields labeled “genre,” “title,” and “description” and the administrative metadata field labeled “date.”³ Based on my research journal, I also reconstructed the search paths for each item to have it as reference for my metadata analysis (see box 1 for an example).

2 In item 1, for instance, it is “searchTerm”: “\”aankondiging\””: <https://mediasuite.clariah.nl/tool/resource-viewer?id=2101608040033514831andcid=daan-catalogue-aggrandst=aankondiging#>. The URL not only contains the search term (st=aankondiging), but also the Media Suite Item ID (2101608040033514831), the collection (daan-catalogue-aggr), and the tool used (resource-viewer). The URL helps both to trace back your steps and to analyze them.

3 While descriptive metadata describe the content and provide context (title, author, publisher, subject, description, etc.), administrative metadata inform us when and how the dataset was created, on technical aspects, and who owns and can use the data. Pomerantz (2015) also identifies three other types: structural metadata, preservation metadata, and use metadata.

CLARIAH Media Suite > Tool "Search" > Clear search > Add new Facet Broadcast type > select facet broadcast type: "Collection band" (<verzamelband) > 48 items in result list > select facet Genre (series): reclame > 6 items in result list of which 5 items 'commercials database' and 1 item 'ster reclame' > select Item *Ster reclame (Ster commercial)*

Main criteria: type (commercial), search strategy (add facet broadcast type)

URL selected item: <https://mediasuite.clariah.nl/tool/resource-viewer?id=2101608040033945131andcid=daan-catalogue-aggr>

Box 1. Example of a search path, criteria and URL as taken from my research journal.

Trace and Compare

In the second phase of the tracing method, the analysis phase, I looked at traces in the material sense. These concern materializations of elements that might be invisible or "traces that mark what has been 'there'" (van der Tuin and Verhoeff 2021, 196). Both Loukissas (2019) and Ketelaar (2001) refer to all kinds of processes that leave traces. First, the data handlers, the archivists, and their collection practices leave traces. Second and relatedly, traces are also left by technological processes that transfer data from the physical sources (paper, reel, etc.) to the interface, such as digitization and indexing. I also add here the technical practices of standardization and normalization that are characteristic of data infrastructures.⁴ While "standardization" refers to setting up one standard to which all other practices have to adhere, "normalization" means having data conform to a certain format so that it becomes machine legible.⁵ As these definitions reveal, both practices are difficult to discern for humanities scholars with no computer science background (like myself) but are in essence aimed at increasing technical interoperability when collections of different archives are aggregated. Third, and often forgotten, is the user (you) who searches the search system. The user also leaves traces, which loop back to tracing in the exploration phase to show how the user explored the archive in the first phase.

4 Rinehart and Ippolito (2014, 60) explain that metadata standards indeed contain traces of its historical context, but also discuss the benefits of these standards for data infrastructures: platform independence, portability, accessibility, extensibility, and longevity.

5 I compiled my own working definition of "normalization" by reading Hoekstra et al. (2019) and Loukissas (2019) and of "standardization" by reading Van Zundert (2013).

For my analysis phase, I follow one of the local reading strategies proposed by Loukissas (2019, 62–69), who points to six features that shape data.⁶ A first element he discerns is *classification*. In classification systems, the world is divided into segments. These segments are used to help in administrative or knowledge production (Bowker and Star 1999). In my case study, I translate local classification to the object at hand, the metadata of television broadcasts, and view genre as a classification system. Second and third, Loukissas looks at *schemata*, the ways of recording metadata, and *constraints*, the conditions that apply for inscribing data such as technical limitations. The last feature type he offers is *errors*. Data are often cleaned or filtered, but ideally the mistakes are kept, as they are relevant as traces (Loukissas 2019, 67). OCR mistakes are commonly known, but the question for my research would be whether I can discern mistakes in audiovisual data collections. Another element, which I consider to be key, is *absences*. As archives define what is kept, shown, or put on display, they also leave out data and metadata. With the last element, *rituals*, Loukissas refers to cultural practices that can be seen, in my case, as the archivists' practices.

These material traces can be revealed by applying “a comparative lens,” a useful framework again provided by Loukissas (2019). Comparison is at the heart of most research, Berger (2016, 21) contends. While comparison is usually meant in either a diachronic or synchronic sense, I compare items within the same search system. Hoekstra and Koolen (2019) refer in this context to *datascope*s, different representations of the same item within archives and infrastructures. My aim, then, is to look for patterns across the six elements as I compare different representations of records in metadata.

To help identify and understand the six elements, I use the documentation of the Netherlands Institute for Sound and Vision: a metadata handbook as used by archivists, the metadata translations for the Media Suite, and the Media Suite's user manual. In addition, I send emails to archivists and developers of the Media Suite when I have questions about some of the metadata fields. I choose to stay as close as possible to students' situations and investigate the limitations of doing this kind of text-driven research with only a limited amount of help from practitioners.

6 Two decades earlier, Ketelaar (2001) discusses similar elements but in more general terms. He also discusses absences, classification, and rituals/practices but does not specify constraints, schemata, and errors.

Metadata Field Genre

In my first comparison, I compared the facets with the controlled vocabulary list and with the key words of the selected items. I focused on the main category for television collections, which is the television genre. Genre is characteristic for both film and television studies and refers—to use a simple definition—to a combination of style and content elements.⁷

First, I wanted to generate an overview of genres as shown in the facets, the filter mechanisms to narrow down the search list, and an aggregation of the key words of all records matching the query. When searching on the full collection with an empty query, the facets show only three types: promos (315), promotions (162), and commercials (24,589). The more general category of interstitials is missing, as well as announcements, idents, logos, trailers, and teasers. This implies that users are not steered toward most interstitial types when using facets.

A similar picture is painted by the thesaurus *Common Thesaurus for Audiovisual Archives* (GTAA), which provides the controlled vocabulary for the audiovisual archives in the Netherlands. Archivists use a thesaurus for disambiguation and for tagging records with key words that are used to optimize retrieval. The GTAA, which is part of the Media Suite's resource viewer, mentions three different terms for "commercials": the Dutch term with capitalization, the Dutch term without capitalization, and the English term as it is also used colloquially in Dutch. I find similar synonyms for the English word "promos." In other words, the GTAA—in its core function—does not normalize the data: it has multiple categories for the same genre. And while "leaders" and "trailers" are not categories in the genre facet, they are mentioned in the GTAA.

The GTAA also uncovers rituals, as it provides for short definitions of the different terms as seen by archivists that are given especially for disambiguation purposes. The definitions show that the common denominator for commercials, promos, and leaders are "(very) short" and "small," by means of the diminutive Dutch suffix "pje" (e.g., filmpje). Promotions are defined as the opposite of commercials, which indicates informal schemata to define interstitials against the "main" genre, commercials. The definitions in the GTAA point to the importance of commercials and promotions for the archive, which is enforced by the fact that the GTAA contains all variations of commercials and promos but not of leaders and trailers. Lacking in facets and GTAA are interstitials and idents. This may be due to the use

7 For advanced theories on television genre, I refer to Mittell (2004).

of English language and the lack of an apt Dutch translation for interstitials and idents, expanding the “semantic gap” to a language gap. Absence, then, is also related to language constraints.

If I then look at the records of my nine items, I notice that only two items have a genre that refers to the interstitial: the commercial is tagged with “Reclame,” and the original leader for the television broadcast is classified as a “Promo” (and not as leader). The other six are classified with the genre of the broadcasts that they encompass, such as “comedy,” “entertainment,” or “news.” The genre metadata field shows that interstitials are chiefly not classified as interstitials, which might be explained by the constraint that each broadcast can have only three genre tags, a requirement that may push interstitial genres into the absent place. In addition, it also shows that the longer an interstitial’s duration is, the more likely it is to be considered a broadcast type that is worth mentioning and archiving.

Leaders:	/
Reclame:	/
commercials:	<i>radio and television commercials</i>
reclame:	<i>advertisements, mostly (ultra)short production, of which the main goal is to sell a specific product or service</i>
Trailers:	<i>short promo videos [filmpjes] for television programs and feature films, not for trailers as vehicles</i>
Promo:	<i>mostly (ultra)short production, of which the main goal is to get attention for a radio or television program</i>
Promos:	<i>promotion spots for radio and television programs</i>
Promotion:	<i>production that is aimed at convincing the viewer or listener, without a commercial intent</i>

Box 2: GTAA entries with definitions (if available) in italics. Translated from Dutch by the author except for “reclame,” which is the Dutch word for commercial.

Metadata Field Title

Titles are the first relevant check of records for users of search systems. The full metadata in the Media Suite shows 27 different title fields, which offers me a plethora of options for comparing title fields with each other.

The interstitial type is mentioned in title fields of only one out of nine records. Not surprisingly, it is the commercial that also has “commercial” in

its title. In the generic title aimed at media professionals, the interstitial type is even mentioned three times: “STER COMMERCIAL; Ster commercial (1993) Duyvis: Tiger nuts – Gerard Cox in tiger outfit is advertising Duyvis tiger nuts.” The genre is mentioned on different occasions in the title and the first time in all capitals so that it could be easily detected. It also shows a technical ritual of the title fields by the Media Suite title field: all title fields are squeezed into this generic metadata field for title. It is only through comparison with the other title fields that I realized that this main title is a composite.

The schemata and annotation rituals differ per title field. Since the archive also has a museum, some items also have a “museum” title field. The museum title of the commercial is “Gerard Cox in tiger costume is advertising Duyvis tiger nuts,” starting with the name of a celebrity rather than the genre (commercial) or the section name (STER), which is mentioned in the general title. The interstitial type “commercial” gets a lot of attention in the program’s title when aimed at media professionals, but it is played down when aimed at the museum audience, which is engaged by using the celebrity’s name. It shows an implicit assumption about the kinds of information museum audiences are looking for.

While title is an important metadata field, it does not provide many traces in the case of the other eight records. This absence of interstitial types in the title field might be indicative of the ritual that interstitials should not be put “on display” in the title field. In its title, the announcement of *Klaas Vaak* is just named “clip,” a generic term that could equally be referring to a regular clip. The title of the weekly recording, which is a full overview of all programs in the order of broadcast, only mentions the programs and leaves out all the interstitials. In the oldest item, the news broadcast, the title field does not mention the leaders, but the tape fields do. The reels contain the news items and have separate titles, of which two are named “leader” followed by a code: FHD. An archivist told me that FHD means it is digitized from film. The video’s materiality and the way it is saved are shown in the reel title, a hidden datum resulting from a ritual. Absence can be taken quite literally, as only six out of twelve reels play out. The titles of the tapes, even when the play-out or video is missing or broken, show what was there and make the invisible visible again.

Metadata Field Description

Like the metadata field title, the “abstract” or description of a record is frequently used to understand the content of a document (Althaus and

Phalen 2010). In the case of the Media Suite, I compare the metadata field description across the nine items and the description of each item with the actual record, the video content.

The first question is whether the records are described (and thus classified) as interstitials. In line with my previous findings, the commercial is described extensively. While the title of the *Klaas Vaak* clip labels it with the generic label “clip,” the description starts with “Promoclip,” making it easier to identify it as an interstitial. None of the description fields of the other items mentions the interstitial type. The documentation manual shows that it is the ritual not to mention “leaders” and “credits” in descriptions.⁸ Other metadata fields show information on the provenance of descriptions, such as the names of the annotators as well as the source of the description. In five records, the annotation field shows that the description was not made by watching the video but by consulting “information,” such as a broadcast magazine. In the case of the weekly recording, the broadcast times are set crudely to 19:15–21:15, followed by the name of the reel (‘\nDs783/V7483’). The ritual leaves out all the interstitials, as the time schedule is taken from a broadcast magazine and is not based on the video itself.

It is particularly instructive to compare descriptions of duplicates that Loukissas (2019, 60) considers to be “key to learning about heterogeneity of data infrastructures.” The *Klaas Vaak* items are very clear examples of duplicates. The four times in which the *Klaas Vaak* announcement was mentioned in the description field were standardized: the same sentence appeared in all four descriptions. In addition, in all four occurrences in the video, the same voice-over announces the clip with the exact same phrase. This shows that standardization is related to schemata—that is, a standard—and most likely also relates to automatically generated metadata. Interestingly, the announcement of the film is also mentioned in the description and follows the same order of words as the *Klaas Vaak* items. However, this time, it coincides with an error. In the video, a voice-over announces a film in a 1970s timbre, which is ambiguous, as the program is broadcast in 2006. When I recognized a famous, contemporary Dutch actor in the video, I realized this was a parody of an announcement. This shows that the description obscured the fact that it is a parody: the archivist did not use the word “parody” but described the item in the very same manner that a “real” announcement would be described. Therefore, the schemata of archival descriptions have transformed the fictional announcement into

8 It specifies that if credits are interesting for re-use, these may be mentioned in the annotation field.

a real announcement, which can be considered an error but equally an indicator for the ritual of description practices. It shows that Loukissas's elements of constraints, errors, rituals, and schemata can also reinforce each other.

Absences are an element that comes explicitly to the fore in the descriptions. In the case of the weekly recording, a two-hour recording of all programs on a Sunday, I tried to find the duplicates and the separate broadcasts, but to no avail. This full-week recording collection that is saved for media history gives insight into missing programs, which is very valuable for media history. In addition, the video shows that the recording contains twelve idents, five announcements, and twelve commercials. All twenty-nine interstitials were left out of the description. However, they are mentioned in the metadata field summary, the description field to provide information on the series to which a broadcast belongs: "The programs are interrupted by commercial breaks (STER), promos, trailers and *Postbus51*-commercials." Again, commercials are mentioned first, and leaders and idents are not mentioned at all.

Metadata Field Date

While I did not expect there to be interesting traces in the pure administrative metadata, I found the date field in the CLARIAH Media Suite to be a treasure trove of traces. In total, there are 120 different date fields. I wanted to check whether the date fields might enable me to reconstruct the entire trajectory of physical objects in the archive to items in the catalog and the Media Suite, which Ketelaar (2001, 138) calls the "semantic genealogy of the record." To this end, I lined up a selection of date fields per record. I discuss the fields in depth for one example, namely, the commercial (box 3).

The date metadata fields are very technical and almost impossible to understand without the direct help of practitioners. At first, I surmised that the broadcast had been put on a VHS tape and digibeta in 1982, but then the editors of this book asked me whether that is possible. I enquired with the archivists of Sound and Vision, who were also inclined to think that this is an error or, more precisely, a randomly chosen date, as the commercial was put on a "collection band" (*verzamelband*). At that point, I decided to show my other interpretations of the date fields of the commercial to the archivists of Sound and Vision. I came to realize that I challenged myself to pick the commercial as example, as it was put on a collection band, and it stems from the early 1990s. In general, the more recent a broadcast is, the

fewer data transformations are inflicted on its metadata and the easier it is to reconstruct the path.

The commercial was broadcast in 1993. It has a specific publication date (January 1, 1993), which is a trace of a constraint that this date field cannot be left blank and should be set on the first of January of the year. In the metadata field annotation, it is mentioned that the exact date is unknown, confirming the constraint. All dates are incorporated in the field “sort date,” which is algorithmically calculated as the most likely publication date when fields are left empty. There are various fields called “Date created.” One of these was close to Museum Genre, so I interpreted it as the date the metadata field was created specifically for the museum, which was confirmed by the archivists as the most likely interpretation.

The dates in the metadata also show when the digital file was entered into the system by means of the field “Asset Date.” “Asset Date” shows that the commercial’s metadata record was ingested by the system in 2008, at about 9:30 in the evening. All other records were ingested in 2008 or 2009. The “Last Updated” date refers to all kinds of updates but most likely to the conversion of the old iMMix catalog to the DAAN catalog; it ranges between 2017 and 2021. The first moment of conversion is most likely “Date created [logtrack]” in 2017. “Asset items created” and “Date last updated” have the same annotation system, namely, the unix timestamp dates that are in schemata, which is incomprehensible for humans. An archivist pointed me to a converter that translates this into a humanly readable date. In this sense, it is an invisible date that only becomes visible through an archivist’s expert knowledge.⁹

Carrier date of the vhs:	1982—01—01
Carrier date of the digibeta:	1982—01—01
Broadcast date:	1993—01—01
Sort date:	1993—01—01
Date created [museum]:	116216280000 (<i>2006—10—30</i>)
Asset item date:	2008—04—11T21:30:09Z (<i>2008—04—11 at 09:30:09pm</i>)
Date created [asset item]:	1207949409000 (<i>2008—04—11</i>)
Date last updated [logtrack]:	1488757329517 (<i>06—03—2017 at 00:42:09am</i>)
Date last updated [conversion]:	1606473164285 (<i>2020—11—27</i>)

Box 3: Date metadata fields for the selected commercial (converted dates in Italics).

9 Converter: <https://www.epochconverter.com/>

Conclusion

In this chapter, I have proposed a metadata tracing method that expands on one of Loukissas's local reading methods. The genres of the interstitial are quite invisible in the facets used in search systems to come to selection and therefore proved themselves to constitute an interesting case study for investigating the interplay between available and unavailable data. First and foremost, it is thanks to the efforts of the archive that most of my interstitials were saved in the first place—most of the records were saved on collection bands. These records, then, were found because I articulated more advanced search strategies. This implies that it is important to spend quite some time on the exploration phase but also that my fifth selection criterion—using different search strategies—is a good one to follow, especially when one's chosen case study is a bit off the beaten path. In terms of (training) critical reflection on search systems, I realized how much my search and my selection are related to oft-hidden decisions that I tried to unravel.

Yet I came to realize that the “interstitiality” of interstitials reaches beyond the case study itself and points at different forms of what can be considered “interstitial data,” data that is hidden at first glance. First, the concept refers to all data that comes to the fore through comparison. The videos shed light on rituals while revealing what was missing in the archive. The duplicates of records were the most illuminating records for the investigation of rituals, while they are also very difficult to locate in large collections. That is why it is important to spend considerable time on the exploration phase, as it allows one to find striking and interesting examples for the analysis phase. It is in the combination of the different representations that even more data and tacit narratives arise.

Second, interstitial data also points at the data that are considered less important for re-use and which therefore disappear from the displays of search systems. The commercial is “on display” in all possible ways: in the facets, the GTAA, the title, and the description. This shows how the archive, at least when it comes to interstitials, is built up for commercials. Commercials are clearly regarded as the most reusable category of interstitials. The “title” field can be considered the “chief” metadata field and is left almost empty in case of interstitials. The GTAA sheds light on the rituals of archivists by providing other types of spellings and definitions. The “description” field is the place where all Loukissas's elements come together and reinforce each other. Tacit narratives were also found in an unexpected metadata field, the date field. This field is

key to the investigation of the provenance of an item and its tacit narratives, also in cases in which the title and description fields are missing or incomplete. My analysis shows that classification, absences, and rituals run through all metadata fields together with constraints and schemata. The six elements helped me to get a better grasp of the mechanisms of local data. Probably due to the limited number of records I analyzed (only nine), I could barely identify errors.

My small-scale analysis enabled me to see some hidden data, but I also came to realize that some invisible data remains invisible. The older the record in the archive, the more difficult it becomes to gather all the information necessary to understand the (different transformations of) metadata. It also costs the precious time of multiple practitioners to unravel past practices. Knowledge about archiving practices has been increasingly preserved in documentation, but this documentation is inherently partial. Interstitial data therefore also refers to invisible data that remains invisible. It is also related to my own standpoint or skills, as a television scholar who knows the search system very well but is neither a trained information scientist nor a trained computer scientist—I was not able to look under the hood of the search system myself.

This raises a question, which is particularly relevant to students: To what extent it is possible to conduct this research without the help of practitioners? I could quite easily understand the title fields, the GTAA, and the description fields, as these are lengthy texts that can be analyzed through comparison. The description fields are an especially interesting metadata field. However, thesauri such as the GTAA are often not available within search systems, while it really helped me to understand the rituals. My advice is to consult the documentation of search systems and see whether they made use of e.g., Europeana or Unesco thesauri. The field that is impossible to understand without help of practitioners is the date field. It was the date field that triggered me the most, and my next plan is therefore to dig into date fields together with practitioners and data scientists to investigate whether we can come a step closer to reconstructing the genealogy from inception and broadcast, through digitization and ingestion, to searchable entry in the Media Suite. This research shows that there is still so much more to investigate for the sake of bringing to the fore historical knowledge about the collections and their operating practices. As Loukissas writes, local reading of obscure data “stimulates curiosity” (67), and that is exactly what it did to me.



Fig. 11. Follow this QR code to the playlist “Finding Interstitials in the Television Archive,” curated by the author.

Acknowledgements

The author would like to thank the archivists and developers of Netherlands Institute for Sound and Vision for their help in understanding metadata and the editors for their input and feedback. This research has been made possible by the NWO infrastructure project CLARIAH-PLUS and the CLICK-NL project Re-Frame.

References

- Althaus, Scott L., Jill A. Edy, and Patricia F. Phalen. 2002. “Using the Vanderbilt Television Abstracts to Track Broadcast News Content: Possibilities and Pitfalls.” *Journal of Broadcasting & Electronic Media* 46, no. 3: 473–92. https://doi.org/10.1207/s15506878jobem4603_9.
- Berger, Arthur A. 2016. *Media and Communication Research Methods: An Introduction to Qualitative and Quantitative Approaches*. 4th edition. Los Angeles: SAGE Publications.
- Borg, Simon. 2001. “The Research Journal: A Tool for Promoting and Understanding Researcher Development.” *Language Teaching Research* 5, no. 2: 156–77.
- Bowker, Geoffrey C., and Susan Leigh Star. 1999. *Sorting Things Out: Classification and Its Consequences*. Inside Technology. Cambridge, MA: MIT Press.
- Creeber, Glen. 2004. “Introduction.” In *Fifty Key Television Programmes*, xiii–xvi. London: Arnold.
- de Leeuw, Sonja. 2012. “European Television History Online: History and Challenges.” *VIEW Journal of European Television History and Culture* 1, no. 1: 3–11. <https://doi.org/10.18146/2213-0969.2012.jethc002>.
- Ellis, John. 2011. “Interstitials: How the ‘Bits in Between’ Define the Programmes.” In *Ephemeral Media: Transitory Screen Culture from Television to YouTube*, edited by Paul Grainge, 59–69. London: Palgrave Macmillan.

- Hoekstra, Rik, and Marijn Koolen. 2019. "Data Scopes for Digital History Research." *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 52, no. 2: 79–94. <https://doi.org/10.1080/01615440.2018.1484676>.
- Johnson, Catherine. 2012. *Branding Television*. London: Routledge.
- Ketelaar, Eric. 2001. "Tacit Narratives: The Meanings of Archives." *Archival Science* 1, no. 2 (June): 131–41. <https://doi.org/10.1007/BF02435644>.
- Kitchin, Rob. 2014. *The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences*. London: SAGE Publications.
- Loukissas, Yanni Alexander. 2019. *All Data Are Local: Thinking Critically in a Data-Driven Society*. Cambridge, MA: MIT Press.
- Melgar Estrada, Liliana, Roeland Ordelman, Jasmijn Van Gorp, and Julia Noordegraaf. 2018. "Media Suite: Unlocking Audiovisual Archives for Mixed Media Scholarly Research." *Selected papers from the CLARIN Annual Conference 2018, Pisa, 8–10 October 2018*. <http://www.ep.liu.se/ecp/159/014/ecp18159014.pdf>
- Mittel, Jason. 2004. *Genre and Television: From Cop Shows to Cartoons in American Culture*. New York: Routledge.
- Noordegraaf, Julia. 2015. "Crowdsourcing Television's Past: The State of Knowledge in Digital Archives." *Tijdschrift voor Mediageschiedenis* 14, no. 2: 108–20. <http://doi.org/10.18146/tmg.139>.
- Polan, Dana. 2013. "It's Fun to Eat: Forgotten Television." In *How to Watch Television*, edited by Ethan Thompson and Jason Mittel, 347–54. New York: NYU Press.
- Pomerantz, Jeffrey. 2015. *Metadata*. Cambridge, MA: MIT Press.
- Rinehart, Richard, and Jon Ippolito. 2014. *Re-collection: Art, New media, and Social Memory*. Cambridge, MA: MIT Press.
- van der Tuin, Iris, and Nanna Verhoeff. 2021. *Critical Concepts for the Creative Humanities*. Lanham, MD: Rowman and Littlefield.
- van Zundert, Joris. 2012. "If You Build It, Will We Come? Large Scale Digital Infrastructures as a Dead End for Digital Humanities." *Historical Social Research / Historische Sozialforschung* 37, no.3: 165–86. <https://doi.org/10.12759/hsr.37.2012.3.165-186>.

About the Author

Jasmijn Van Gorp is Assistant Professor in the Department of Media and Culture Studies at Utrecht University. Her research is situated at the intersection of audiovisual data and the digital humanities. Since 2019, she has co-led the "Audiovisual Data and Media Studies" work package of the Dutch digital research infrastructure CLARIAH, which develops the Media Suite. > j.vangorp@uu.nl