Lego Project Data:
An Open Data Archive for Qualitative Video Research

Jacob Davidsen
0000-0002-5240-9452
Department of Communication & Psychology
Aalborg University

Mathias Thomsen
Aalborg University

Paul McIlvenny
0000-0003-2327-2124
Department of Culture & Learning
Aalborg University

Abstract

This article introduces and documents the collection and processing of raw video and audio recordings of an experimental Lego puzzle team game, which led to the archiving of the audio-visual data and ancillary materials in an open format suitable for sharing and reuse. The primary motivation was for the data to be included in demonstration packages for immersive qualitative analysis and transcription software tools that work natively with 360-degree video data. The data is made available in an open data archive with a Creative Commons license.

Keywords
Video data, Big Video, Archive, Open Data, Audio-visual

1. Introduction

This article gives a short introduction to the Lego Project open data archive that is shared with a Creative Commons (CC) license. It is intended both as an archive that others can use and learn from, as well as an example of a methodical way to document a data archive for qualitative video research. Besides the importance of creating and maintaining private video data archives for research projects (Cary 1982; Grimshaw 1982; Fitzgerald 2019), and concerns about how data technologies shape qualitative research in profound ways (Evers 2011) that need to be traced, there are increasing calls for archives that are open (Vos and Fernandes 2017) and transparent (Perkel 2018), in addition to being well documented. As yet, there are only a few academic journals in specific fields, often backed by commercial publishers, that support

publishing data archives (e.g. Mendeley Data), but they are infrequently used within the Social Sciences and Humanities. In parallel, there is some discussion about how open, non-proprietary publication formats for data (and software) archives generally should be designed, standardised and cited (Candela et al. 2015; Corti and Gregory 2011; Jeffries et al. 2019; Mons 2018; Nowogrodzki 2020). Since archiving video data is not yet a dominant practice in qualitative social science and humanities research, it is important to discuss and find the right formats. This article is our first attempt at publishing openly a complex video data archive.[1] To do so, however, can entail certain constraints on data collection.

2.   Data collection

In our video-based qualitative research, many of the audio-visual recordings we make are collected under ethical conditions (cf. GDPR in the European Union since May 2018) that make it difficult to share the recordings publicly and openly, even with informed consent and when anonymised. It was imperative for the larger *Big Video* endeavour (McIlvenny and Davidsen 2017) and the *BigSoftVideo* team that we also collect rich data that we could use as data examples when demonstrating the software we have developed and are developing in future for video-based qualitative research and immersive qualitative analytics (IQA). To gain informed consent for unrestricted public dissemination, under a Creative Commons (CC) license, an experimental setting had to be constructed.

In order to collect rich audio-visual data for unrestricted use in academic research, an experiment was devised by Jacob Davidsen in July 2019. An artificial situation was designed in order to illustrate the desirability of using multiple cameras and microphones to adequately record and document the talk and situated action of all of the participants no matter where they were or how mobile (Speer 2002). The approach to camerawork was informed by the methodological strictures of data collection in conversation analysis (Mondada 2013).

In a large room, tables and screens were laid out at opposite ends of the room in a symmetrical fashion. Subjects were enrolled in the experiment and they signed informed consent forms. Many 2D and 360-degree cameras, as well as mono, binaural and ambisonic microphones, were positioned to record the setting and the action from multiple perspectives during the short time-frame of the experiment (see the Figures and Table 1 below).

On the 2nd August 2019, six subjects were divided into two teams: P1-P3 in Group 1 and P4-P6 in Group 2. The two groups were assigned to two standing tables at one end of the room (Group 1 at table 1 and Group 2 at table 3). At the other end were two screens behind which another table for each group was to be found (table 2 for Group 1 and table 4 for Group 2). In Figure 1, the members of Group 1 are shown

---

[1] An alternative method to help visualise the spatiality of a complex recording setup across multiple spaces and temporalities is supported by the SQUIVE (Staging QUalitative Immersive Virtualisation Engine) Virtual Reality prototype developed by Paul McIlvenny. See McIlvenny (2020a) for a description and a link to a video demonstration.

assembled around table 1. The 360-degree video camera (G1-T1-YI360) was positioned on the table. Wireless microphone transmitters can be seen on the table and attached to a waist band.

Figure 1 - Group 1 working together on building the model at table 1.

One member of each group (P2 and P5) was assigned to go between their group and the respective hidden table. On the hidden tables (T2 and T4) were two identical Lego models. Each mobile group member could go to their hidden table and inspect and touch the model, but they were not allowed to remove it or take photos. In Figure 2, the member of Group 1 (P2) who runs between tables 1 and 2 is shown inspecting the Lego model hidden from the rest of her group. The 360-degree video camera (G1-T2-YI360) was positioned on table 2.

Figure 2 - P2 inspecting the Lego model at table 2.

The small embedded images in the main image in each Figure show different simultaneous views on the scene around the camera, i.e. from four 90-degree angles at the same moment as we see in the main image. This provides a little more context for a 2D viewport image when the 360-degree video footage is rendered in a flat rectangular frame.

The task was for the mobile member of each group (P2 and P5) to give verbal and embodied instructions to the other two members of their group how to reconstruct the model exactly from the many bricks on their group's table. The mobile group members were *not* to touch the bricks that the other members of their group were handling. The two groups were competing to complete the task before the other group. The competition lasted about fifteen minutes at which point one group completed the task. The language used in all social interactions was Danish (a translation has been provided using subtitles and a transcript - see below). The camera team comprised three persons (R1-R3), including two of the authors, Jacob Davidsen and Mathias Thomsen.

The diagram in Figure 3 illustrates, from an aerial perspective, the spatial configuration of the participants, the significant furniture and the primary audio-visual recording devices. Note that, with hindsight, even more video recording devices could have been used; for example, they could have been placed overhead above the tables to better capture the intricate, embodied work of the groups to inspect and assemble their lego models.
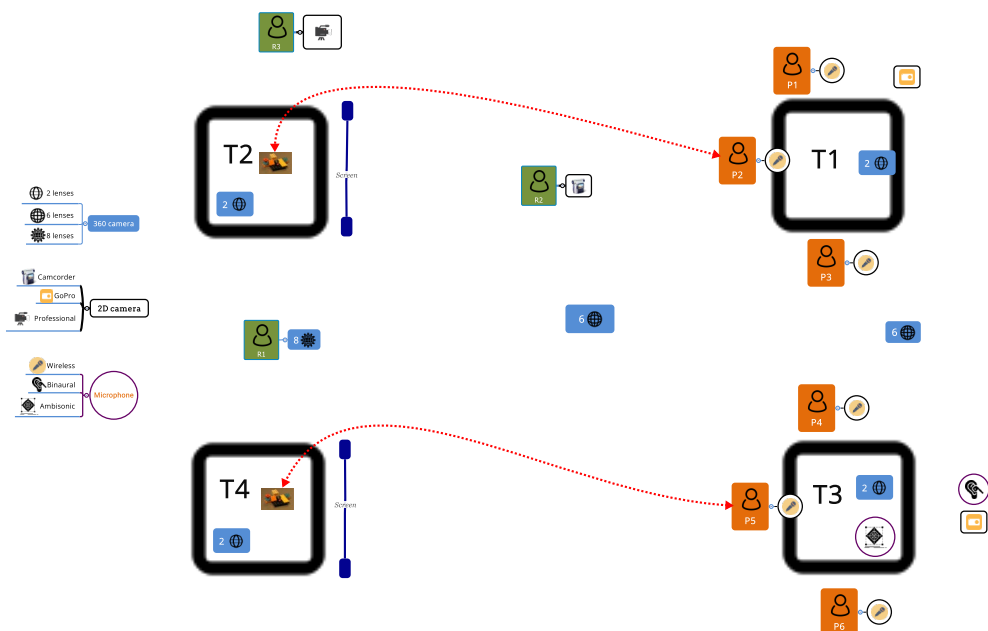

Figure 3 - Aerial map of room and participants -
DOI: 10.6084/m9.figshare.13292729.

Adding more detail, the diagram in Figure 4 shows the different items of recording equipment, and the attached lists make explicit the known settings of the recording devices and the data recording formats.

The figures are also available in *SVG* format on *FigShare*. This format can be rescaled without loss of resolution.

Note that not all of the recordings included in this open archive are perfect. Where possible, less than satisfactory recordings have been included to illustrate the problems that can occur when collecting audio-visual data (see Pink et al. 2018 for a discussion of the significance of the concept-metaphor of 'broken data'). For instance, some of the wireless microphones were not transmitting their signal to the receiver without some noise, interference and cross-talk (e.g. G1-P3). Nevertheless, one microphone recording (G2-P5) was not included in the archive because it was incomprehensible. In addition, because of physical placement, the ambisonic microphone recording (G2-Rode NT-SF1 ambiX) cannot be stitched together with any of the 360 cameras recording simultaneously in the scene. It should be replayed alone as spatial audio.

More detailed information about the recording formats and codecs of each audio-visual file in the archive can be revealed by using the free, multiplatform software *Media Info*.[2]

The total raw archive from all recording equipment was about 350GB for 15:00 minutes of recording time. The open archive indexed in this article is 3.4GB for 3:43 minutes of recording time.

All the participations have given their written consent for the recordings to be made public under a CC BY-NC-SA 4.0 license, which allows others to download, copy and use the data made available in the corpus, provided it is for non-commercial purposes and full attribution is always given as per the license.
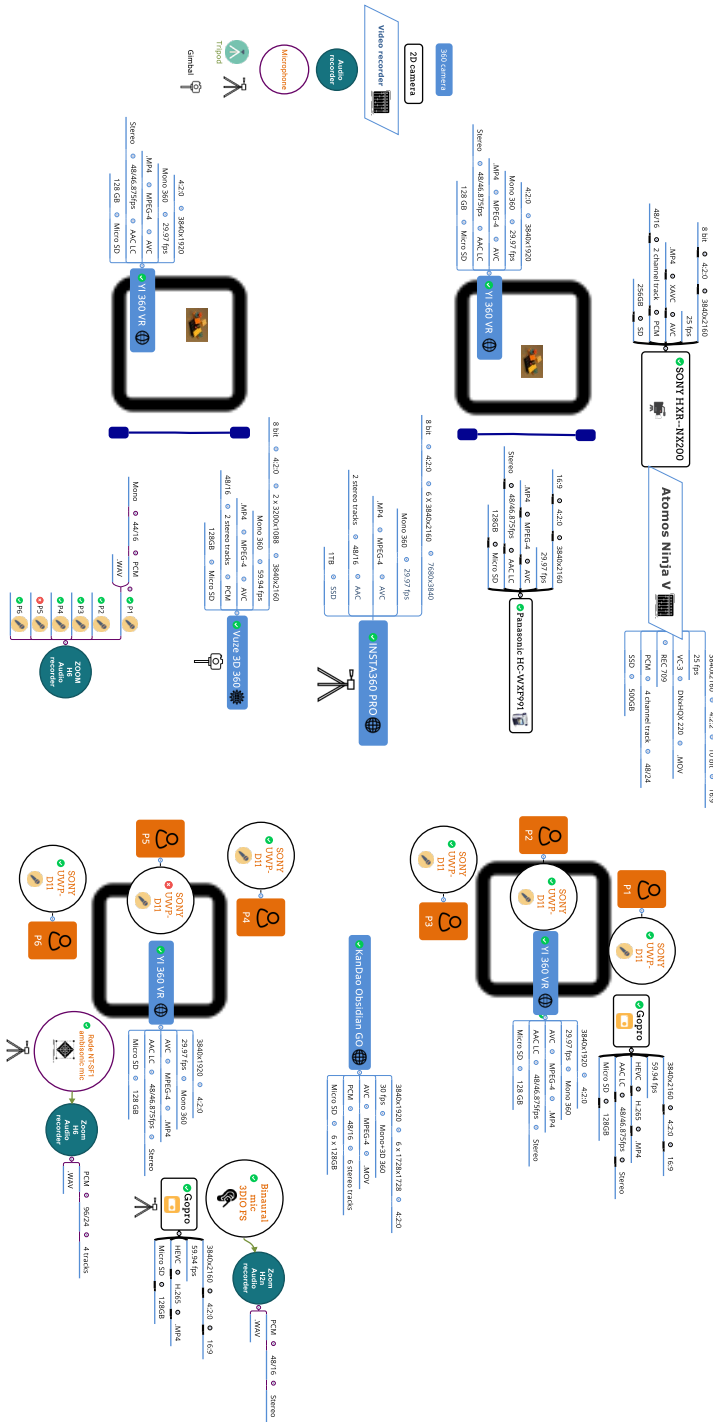
---

[2] mediaarea.net/en/MediaInfo

Figure 4 - A mapping of the equipment used to collect data -
DOI: 10.6084/m9.figshare.13292744

3.   Data post-processing

The video and audio recordings were post-processed using a non-linear editing suite, *Adobe Premiere CC*, and a digital audio workstation, *Reaper*, as well as proprietary software for 360-degree video and ambisonic audio that is designed to work with the raw data derived from the respective camera or microphone. Moreover, several filters and plugins were used to re-stitch and recalibrate the 360-degree footage, and improve the audio voice quality. The editing work was undertaken by Mathias Thomsen and Paul McIlvenny.

In a further step, the post-processed data was synchronised temporally to one master recording so that all recordings could be played simultaneously while in sync. In order to restrict the size of the archive, an excerpt of three minutes and forty-five seconds was chosen. Synchronised clips were exported that had the same IN and OUT timecodes. We have produced an online video that gives instructions how to synchronise multiple video files and export them as clips with identical IN and OUT timecodes (Thomsen 2020).

As a result of all the post-processing, the files contained in the archive are *not* the *raw* files that were recorded in the recording device. For many 360-degree cameras, the raw data are often recorded as separate videos for each lens, which must be stitched together in post-production, resulting in an equirectangular format video that can be played with an appropriate digital media player or in a VR headset or on YouTube (with the correct meta-data). The 360-degree video files in the archive are equirectangular. See Figure 5 for a frame grab from one of the 360-degree cameras (Insta360Pro). P2 is just about to reach the hidden table 2, while her team continue to work at table 1. Group 2 is working at table 3. R1 is filming table 1 and R2 is filming table 3. R3 is filming a wide shot of tables 1 and 3. The output from the Sony professional 2D video camera was redirected to the Atomos Ninja V recorder, which records in a high resolution 4K format called DNxHQX 220. This requires a large storage capacity on SSD, and the resulting file is too large to fit in the archive. Instead, it has been rendered in a more compressed, lower resolution H.264 MP4 format.

Figure 5 - A 360-degree camera frame grab in an equirectangular format
showing all the participants in the room.

For ambisonic microphones, the raw data are often recorded as four separate channels of mono audio that taken together are arranged in a format called A-Format, which is not particularly useful on its own. After conversion to B-Format, the four-channel 1st order ambisonic spatial audio file can be manipulated in a Digital Audio Workstation (DAW) and played back on an appropriate digital media player or on YouTube (ambiX ACN/SN3D with the correct meta-data) or, best of all, in a tracked VR headset with headphones (such as in *AVA360VR*). The file of the ambisonic recording in the archive (T2-Rode NT-SF1 ambiX) is of the latter four-channel format.

The free, open source software *Subtitle Edit* was used to create English subtitles (.srt) for two of the videos (G1-T1-GoPro and G1-T2-YI360; see below).[3] These two videos were chosen to best illustate the talk of Group 1 and the self-talk of the runner (P2) in Group 1 when she is hidden behind the screen at table 2. Her self-talk cannot be heard clearly on the audio recordings of other cameras. The subtitles will appear in overlay when the video is played with a desktop media player, such as *VLC* or *PotPlayer*.

4.   Data archiving and use

The archive was assembled from the post-processed data by Paul McIlvenny. The complete archive (compressed 3.2GB) is available for download from Zenodo, a non-commercial, open access online archive service.[4] It is a static archive that has a simple folder structure (see Figure 6).
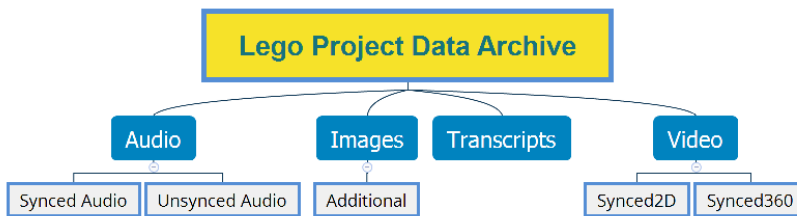

Figure 6 - Lego Project Data Archive folder structure.

---

[3] www.nikse.dk/subtitleedit/

[4] doi.org/10.5281/zenodo.4292236

In Table 1, all the audio-visual files contained in the folders are tabulated according to format, persons, position in the room and extra notes.

Filename	Format	Group/Person	Position	Notes
G1-T1-GoPro.mp4	2D Video	Group 1	table 1	Subtitles
G2-T3-GoPro.mp4	2D Video	Group 2	table 3	
Panasonic2D.mp4	2D Video		mobile	
Sony+Ninja2D.mp4	2D Video	Groups 1/2	static	
G1-T1-YI360.mp4	360 Video	Group 1	table 1	
G1-T2-YI360.mp4	360 Video	P2	table 2	Subtitles
G2-T3-YI360.mp4	360 Video	Group 2	table 3	
G2-T4-YI360.mp4	360 Video	P5	table 4	
Insta360Pro.mp4	360 Video	Groups 1/2	static	
Kandao-ObsidianGO.mp4	360 Video	Groups 1/2	static	
Vuze360.mp4	360 Video		mobile	
G1-P1.wav	Mono	P1	table 1	
G1-P2-runner.wav	Mono	P2	tables 1/2	
G1-P3.wav	Mono	P3	table 1	Cross-talk and dropouts
G2-3DIO.wav	Binaural	Group 2	table 3	cp. G2-T3-GoPro
G2-P4.wav	Mono	P4	table 3	
G2-P6.wav	Mono	P6	table 3	
G2-Rode NT-SF1 ambiX.wav	AmbiX	Group 2	table 3	cp. G2-T3-YI360
R2-Task-description.wav	Mono	R2		Instructions
LEGO blocks1.png	Image		table 3	Lego model
LEGO blocks2.png	Image		table 4	Lego model
Lego Transcript.txt	Transcript	Group 1	tables 1/2	

Table 1 - Meta-data for files in the archive.

In addition to the English subtitles for two of the videos, the data archive also includes a full transcript in plain text of the talk of Group 1 with an English translation.

A second, dynamic archive of the same data has been assembled as a demonstration project for the *AVA360VR* (Annotate, Visualise, Analyse 360-degree Video in VR) software tool. This second archive will also be made public when the *AVA360VR* software is released. It will be incorporated into the demonstration package for that software along with video and 'volcap' help guides.

Some files in the data archive has also been used as part of a demonstration in the *CAVA360VR* (Collaborate, Annotate, Visualise, Analyse 360-degree Video in VR)

software, and as the source files for an example project using a beta version of the 360-degree video transcription software *DOTE* (Distributed Open Transcription Environment) that we are developing.

5.   Conclusion

We do not intend this article to be a comprehensive overview and assessment of digital video archiving strategies. Instead, we have a much simpler goal, which is to document a specific solution that we developed to deal with complex recording settings with multi-cam and multi-mic setups. Researchers and/or technicians who collect data often forget over time the specific settings and properties of the recording technology they deployed. Moreover, if the data collected is to be shared with others not involved in the recordings, then it can be difficult to navigate and reuse an impenetrable archive. The innovation in our pilot project was to visually document some of the meta-data, both technical and spatial, that accompanies data collection, but is often lost or remaindered. In this way, some key aspects of the video and audio data are recoverable. In addition, a better understanding of the relationship between the recordings makes possible a more exploratory take on the archive, and this can be undertaken by others. For example, we have tended to prioritise the team comprising P1, P2 and P3; the other team has not been the focus of our attention. However, the archive contains the recordings of several cameras and microphones that capture the talk and activities of that other team. Moreover, the rich corpus of recordings enables new viewings (from different angles and framings) and listenings (spatially) if one learns how to navigate the archive from this documentation. From a pedagogical perspective, the technical details can prove helpful for other researchers and students who wish to collect data in a similar complex setting.

If we agree that it is crucial to the integrity of our research that others can inspect and re-analyse the same (raw) data, then more work is needed to create and maintain the shared formats and open archives that enable new, common practices to develop. This article, and the archive it describes, is a small step towards testing the waters with a less risky example. We appreciate that many qualitative research projects cannot for serious reasons grant the same sort of access to the original (or post-processed) recordings. Even if the recordings are anonymised, the worthy goal in open scholarship of giving others the opportunity to work with the originary data independently, and so to challenge or build on the results of the data owners, may not be achievable. More work is required to explore further a question we often neglect in qualitative research that is anchored in audio-visual archives, namely how to manage and publish our data archives in a more open and transparent fashion within the limits of GDPR and ethical constraints. In addition, an open discussion is needed about whether or not the goal of standardisation and interchangeability of data that is archived for complex cases is feasible or desirable given the specificity of each data collection setting and the disciplinary concerns (Pels et al. 2018).

Acknowledgements

References

Candela, Leonardo, Donatella Castelli, Paolo Manghi, and Alice Tani. 2015. 'Data Journals: A Survey'. *Journal of the Association for Information Science and Technology* 66 (9): 1747–62. https://doi.org/10.1002/asi.23358.

Cary, Mark S. 1982. 'Data Collection: Film and Videotape'. *Sociological Methods & Research* 11 (2): 167–74. https://doi.org/10.1177/0049124182011002004.

Corti, Louise, and Arofan Gregory. 2011. 'CAQDAS Comparability. What about CAQDAS Data Exchange?' *Forum: Qualitative Social Research* 12 (1). https://doi.org/10.17169/FQS-12.1.1634.

Evers, Jeanine C. 2011. 'From the Past into the Future. How Technological Developments Change Our Ways of Data Collection, Transcription and Analysis'. *Forum: Qualitative Social Research* 12 (1). https://doi.org/10.17169/FQS-12.1.1636.

Fitzgerald, Richard. 2019. 'The Data and Methodology of Harvey Sacks: Lessons from the Archive'. *Journal of Pragmatics* 143 (April): 205–14. https://doi.org/10.1016/j.pragma.2018.04.005.

Grimshaw, Allen D. 1982. 'Sound-Image Data Records for Research on Social Interaction: Some Questions and Answers'. *Sociological Methods & Research* 11 (2): 121–44. https://doi.org/10.1177/0049124182011002002.

Jefferies, Neil, Fiona Murphy, Anusha Ranganathan, and Hollydawn Murray. 2019. 'Data2Paper: Giving Researchers Credit for Their Data'. *Publications* 7 (2): 36. https://doi.org/10.3390/publications7020036.

McIlvenny, Paul, and Jacob Davidsen. 2017. 'A Big Video Manifesto: Re-Sensing Video and Audio'. *Nordicom Information* 39 (2): 15–21. https://doi.org/10.5281/zenodo.3972007.

McIlvenny, Paul. 2019. 'Inhabiting Spatial Video and Audio Data: Towards a Scenographic Turn in the Analysis of Social Interaction'. *Social Interaction: Video-Based Studies of Human Sociality* 2 (1). https://doi.org/10.7146/si.v2i1.110409.

———. 2020a. 'The Future of 'Video' in Video-Based Qualitative Research Is Not 'Dumb' Flat Pixels! Exploring Volumetric Performance Capture and Immersive Performative Replay'. *Qualitative Research* 20 (6): 800-818. https://doi.org/10.1177/1468794120905460.

———. 2020b. 'New Technology and Tools to Enhance Collaborative Video Analysis in Live 'Data Sessions''. *QuiViRR: Qualitative Video Research Reports* 1 (December), a0001. https://journals.aau.dk/index.php/QUIVIRR/article/view/a0001.

Mondada, Lorenza. 2013. 'The Conversation Analytic Approach to Data Collection'. In *The Handbook of Conversation Analysis*, edited by Jack Sidnell and Tanya Stivers, 32-56. Oxford: Wiley Blackwell.

Mons, Barend. 2018. *Data Stewardship for Open Science: Implementing FAIR Principles*. Boca Raton: CRC Press.

Nowogrodzki, Anna. 2020. 'Eleven Tips for Working with Large Data Sets'. *Nature* 577 (7790): 439–40. https://doi.org/10.1038/d41586-020-00062-z.

Pels, Peter, Igor Boog, J. Henrike Florusbosch, Zane Kripe, Tessa Minter, Metje Postma, Margaret Sleeboom-Faulkner, et al. 2018. 'Data Management in Anthropology: The Next Phase in Ethics Governance?' *Social Anthropology* 26 (3): 391–413. https://doi.org/10.1111/1469-8676.12526.

Perkel, Jeffrey M. 2018. 'A Toolkit for Data Transparency Takes Shape'. *Nature* 560 (7719): 513–15. https://doi.org/10.1038/d41586-018-05990-5.

Pink, Sarah, Minna Ruckenstein, Robert Willim, and Melisa Duque. 2018. 'Broken Data: Conceptualising Data in an Emerging World'. *Big Data & Society* 5 (1): https://doi.org/10.1177/2053951717753228.

Speer, Susan A. 2002. ''Natural' and 'Contrived' Data: A Sustainable Distinction?' *Discourse Studies* 4 (4): 511–25. https://doi.org/10.1177/14614456020040040601.

Thomsen, Mathias. 2020. 'How to synchronise and export multiple video clips recorded simultaneously using a Non-Linear Editor'. Zenodo. https://doi.org/10.5281/zenodo.4090664.

Vos, Rutger, and Pedro Fernandes. 2017. Open Science, Open Data, Open Source. Pfern/Osodos V1.0.0. https://doi.org/10.5281/ZENODO.1015288.