



AALBORG UNIVERSITY
DENMARK

Aalborg Universitet

Video Life Cycle Data Management (VDM)

The management of video data from educational research – Final Report (extended version)

Otrell-Cass, Kathrin; Andersen, Bjarne; Offersgaard, Lene; Eckardt, Max Roald

Creative Commons License
CC BY-NC-SA 4.0

Publication date:
2018

Document Version
Version created as part of publication process; publisher's layout; not normally made publicly available

[Link to publication from Aalborg University](#)

Citation for published version (APA):

Otrell-Cass, K., Andersen, B., Offersgaard, L., & Eckardt, M. R. (2018). Video Life Cycle Data Management (VDM): The management of video data from educational research – Final Report (extended version). Aalborg University.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal -

Take down policy

If you believe that this document breaches copyright please contact us at vbn@aub.aau.dk providing details, and we will remove access to the work immediately and investigate your claim.

Video Life Cycle Data Management (VDM)

The management of video data
from educational research –
Final Report (extended version)

October 2018

Kathrin Otrell-Cass
Bjarne Andersen
Lene Offersgaard
Max Eckardt



1. The project and the infrastructure

The **overall intention of this pilot project** was to explore the data management processes and develop and test infrastructures and resources that support research with video technology (from more to less sophisticated levels of analysis), particularly relevant to educational research, because:

- Video has become a standard research tool amongst educational researchers and is frequently shared;
- Video data captures highly sensitive data of and about people, including vulnerable groups like children;
- Video data requires that researchers safely collect, store, analyse, discard, and share video data (see recent updates to European data protection law, GDPR <https://tinyurl.com/yb8eda6d>).

What comprises video data: video, audio files, and field notes (ie associated word documents), metadata, ethical procedures (data management plans, informed consent). It includes raw, annotated and coded data.

The aim was to *explore, develop, pilot, and evaluate* how *existing* research infrastructure could be used and extended to support educational research using video. In particular, the project explored existing infrastructure, including: Kaltura, Edumedia2, CLARIN.dk, DMPonline, and data.deic.dk (now sciencedata.dk) to identify extensions and develop & test prototype modifications.

Project organisation and execution: The project involved team members from the following institutions: Aalborg University (AAU), The Royal Danish Library (KB), Copenhagen University (KU), and The University of Southern Denmark (SDU). In addition, experts from the following institutions were involved for piloting purposes: Waikato University, New Zealand, University of Oslo, Norway, and Aarhus University, Denmark. The project involved researchers and technical experts. The project was managed by a central team with expertise in different aspects to do with video data research/storage, headed by Kathrin Otrrel-Cass (AAU), Bjarne Andersen (KB), Lene Offersgaard (KU), and Max Roald Eckardt (SDU). Each institutional team leader was supported by an institutional team consisting of either researcher/s, technical staff, or both. The project worked alongside a project timeline, held regular management board meetings, whole team meetings, and consulted with members of the Danish e-Infrastructure (DeiC) to ensure the successful execution of the project. Since work on the Kaltura Mediaspace was a key aspect in the project, the team was grateful for the support provided by Kaltura who sponsored the licenses, organised the installation of the platform as well as providing on-call support. In order to execute some technical work to do with the implementation and set-up of a web-interface, the project involved also experts from the Reload, an IT solutions company.

The project aims were executed by drawing on real test cases (video data), provided by the research teams at AAU, KU, and SDU to ensure relevance for developing the prototype or any other resources. This means that examinations of existing infrastructure (for example: <https://DMPonline.deic.dk>) were always conducted in context. The project process involved the different steps Analyse-Define-Ideate-Prototype-Test as illustrated in Figure 1.

Design process in the VDM project

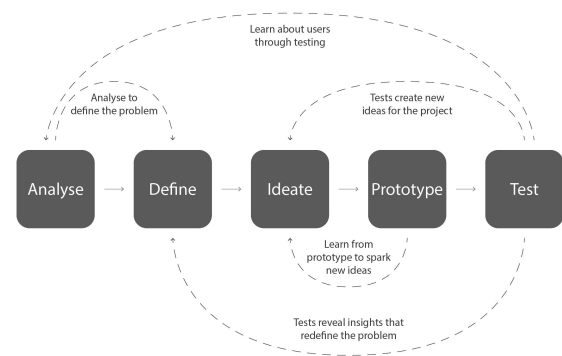


Figure 1: Design process in the VDM project

Academic rooting

Since this project was concerned with supporting the Danish research community it was important to ground the project in academic expertise. This was achieved by building on a national team of experts representing nested and connected research teams and initiatives situated in digital humanities research. The academic research teams involved in the project were connected to VILA - The Video Research Lab at Aalborg University, the video lab of the University of Southern Denmark at campus Kolding; and the section Center for Language Technology at Institute for Nordic Research at Copenhagen University.

2. Achievements of the pilot project

A key aim in the project was to develop new infrastructure for sharing video for research purposes based on existing platforms. For that reason, we explored: Kaltura, Edumedia2, CLARIN.dk, DMPonline, and data.deic.dk (now sciencedata.dk).

After an initial examination by the team, it was determined that Edumedia2 was unsuitable as a test platform so the project concentrated to work with the functionalities of a new and clean instance of Kaltura Mediaspace sponsored by Kaltura. data.deic.dk was also explored; however, it was also identified as unsuitable to support the sharing of data, and shared working with video files, although it represents a secure platform for the sharing of project data at a national level since international collaborators experience difficulties accessing the site. Efforts were therefore invested into exploring and developing the prototype for this project with infrastructures Kaltura Mediaspace, Clarin.dk and DMPonline. The results of the pilot project will be presented by the following topics that represent the nature and focus of the different work packages:

1. Needs Analysis
2. Design of Prototype
3. Implementation of Prototype
4. Pilot Testing of Prototype

2.1. Needs Analysis

Step one in this project was a literature review to identify what aspects academic research had been identified as necessary for video based educational research. The specific focus on this kind of research had to do with it being; a) a highly popular research method in the learning sciences, and b) being research that often dealt with highly sensitive data, thus demanding high levels of care in regards to the collection, processing, sharing, and storing of data. The literature review formed the basis for the needs analysis. The literature review identified the following key steps in video based research: Data collection; Documentation and Metadata; Ethics and Legal Compliance; Storage and Backup; Selection and Preservation; Data Sharing; and Responsibilities and Resources.

Based on three contextual cases supplied by the academic partners AAU, KU, and SDU, the project team developed a video life cycle management flowchart (see Figure 2 below or visit <https://bit.ly/2yDI2fZ>) to identify the different steps video based researchers would need to typically consider, from planning to long term archiving. As part of this process the researchers went through a dialogical process supported by card-sorting (identifying the different steps) to provide the input for the design of the flow chart.

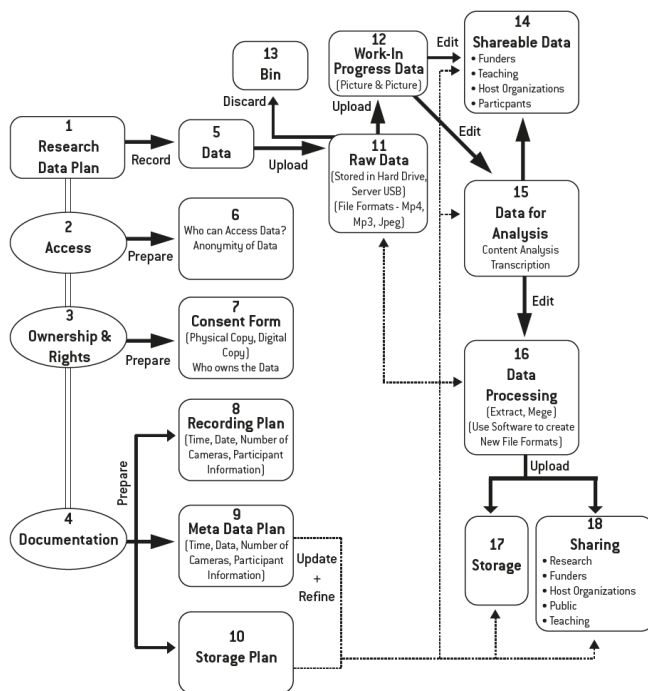


Figure 2: Video Life Cycle Flow Chart

The needs analysis was used for the Design of the Prototype (2.2) to identify key tools researchers require and functionalities that a sharing platform would need. The needs analysis formed the basis for further developmental activities in this project.

Activities	Deliverables
Stakeholder analysis (stakeholders: individual, systems, roles)	State of the Art on (video) data management planning within humanistic research (see appendix 1)
User cases / Process description of scenarios (maps stakeholders to services)	Stakeholder Analysis, Process Description (software, scenarios), this was achieved by using the three case studies
Specification of infrastructure and data management plan	Input into video life cycle data management flow chart
Software Catalog including the integration of the video data management with the national DIGHUMLAB network	Needs analysis based on test cases identifying software needs to then select those with the potential for development (see appendix 2)

Table 1: Activities and Deliverables Needs Analysis

2.2. Design of Prototype

The design of the prototype of infrastructure services had the aim to set up infrastructure that supports researchers' work through the video data lifecycle phases utilising existing infrastructure, including data.DeiC.dk, as part of the storage practices.

The design work was shaped by the development of a prototype mock-up interface (wireframes) that facilitated the discussions around what was viable and useful and qualified the prioritisation when it came to the actual implementation around the prototype.

The mock-up interface played an important role since it served as inspiration including possible future developments as kind of a “front-door” leading researchers working with video in the relevant directions. It resulted also in prompting the preparation of more static guidance that can be implemented at a “front-door” and the prototype “manage video research projects”-interface (developed in WP4) that could be implemented within this frame as well instead of being a stand-alone interface.

Based on the workflow analysis done in WP2 an initial mock-up user interface was designed and internally evaluated. This mock-up facilitated the discussions about the design of the actual working prototype.

Our mock-up interface design aimed at maximising usability with regard to operational management of data. The feedback from scientists in our field research had indicated that they would not endorse any overhead in administering data usage rights. Therefore, our user interface design broke down user rights management into three categories:

1. denial or permission to use data
2. purpose of use is for either research, teaching, or both
3. the domain limitation: (user whitelist, institution, country)

To manage usage rights for a file or a group of files, a user of our prototype would relate to each of the categories which would produce a new rule. These rules for sharing files stack as a cascade of inclusions and exclusions of users across institutions and countries/areas. The example below illustrates a wide permission of a file being used in Denmark, Norway, and Sweden with a supplement of Aarhus University. This cascade of permissions allows federated logins from Aarhus University to use the data even when they log in from another country. Also, in the case of reviewed permissions, users from different domains can be discriminated and respected with individual rules.

Mockup User Interface design for managing file sharing rules

The implementation displayed below is based on NodeJS, ExpressJS, and VueJS. It was planned to be connected to the Kaltura API to administer sharing permissions on a file and channel basis. This proved to be more complicated than initial probes into Kaltura’s development environment made it seem. In our initial research we anticipated that Kaltura’s nodeJS API would comply with their comprehensive documentation. However, the ported nodeJS version was substantially different and incongruous with its documentation and thus it made it very hard to actually implement working functionality.

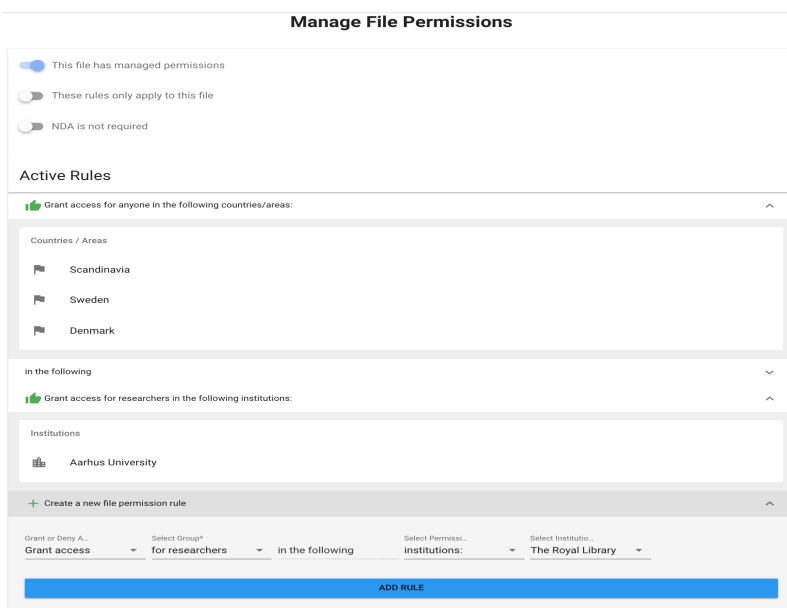


Figure 3: Manage File Permissions

Based on the evaluation of the initial mock-up interface and the original needs analysis, a second generation working prototype interface was needed for managing the concept “research projects” as well as “project members”. This was a necessary step since the datamodel in Kaltura with the frontend of Mediaspace does not embody this. The evaluation showed that using Kalturas “channels” functionality and sharing possibilities as a technical solution to implement Project- and member concepts was not obvious and operative for the users.

- Kalturas Mediaspace should serve as a front end for
 - uploading content — both video as well as other additional material (documents)
 - streaming access to content — based on secure sharing in 4 different “levels”
 - downloading of material in original format
 - adding of extra metadata defined by a project

Following the needs analysis, two further resources were developed. Since they were not identified as crucial for fulfilling the aims of the project, these resources were not tested with users. The two resources include dataflowtoolkit.dk and an interactive Informed Consent Form for Video Research.

DataFlowToolkit.dk

In response to the examination of the DMPonline template, the project identified the need to provide guidance that would support the data management process for researchers working with video. While the project had access to the DMPonline sandbox through DeiC, we decided not to add new guidance text. This is mainly due to the fact that we realised that this guidance would stay hidden for researchers not actively working with a data management plan within DMPonline.

DataFlowToolkit was simply produced through the content management system (CMS) on WordPress. A specific theme dedicated to ‘Qualitative video data’ provides specific information about topics such as ethics, metadata, and more.

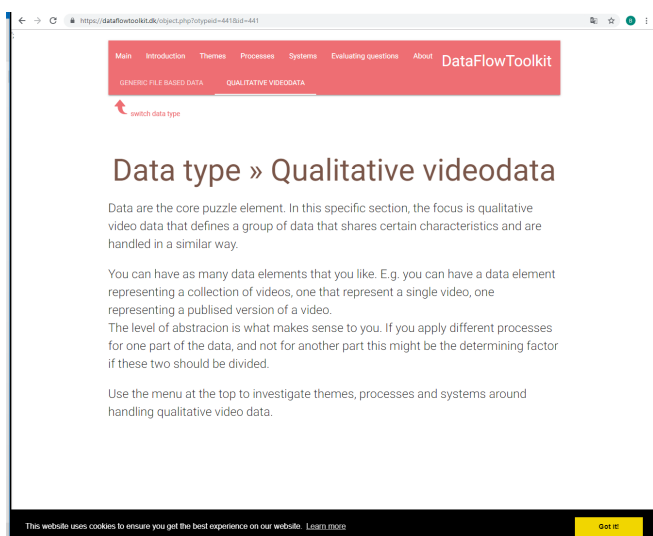


Figure 4: Screenshot DataFlowToolKit (<https://dataflowtoolkit.dk/object.php?otypeid=441&id=441>)

This is a first generation resource and can be developed further. Further details and the background for the first version of dataflowtoolkit.dk is available here; https://vidensportal.deic.dk/da/News/Data_management_i_praksis (doi: 10.7146/aui.243.174)

Informed Consent Form for Video Research

An interactive Informed Consent Form was developed to specifically support researchers working with video. This template was simply produced utilising Google Forms (available in Danish <http://bit.ly/2K6SqA9> and English

<https://tinyurl.com/y76p2qpw>) to scaffold the different steps and considerations researchers should be taking into account when consent is sought from participants, including from vulnerable participants (e.g. children). This provides a relatively simple process prompting a researcher to consider various aspects (includes also links to important background documents to do with ethical conduct), so that by the end a letter can be produced that is tailored to suit the requirements of a particular project. This template was also a first generation working prototype and was finished before the EU's EU General Data Protection Regulation (GDPR) was implemented.

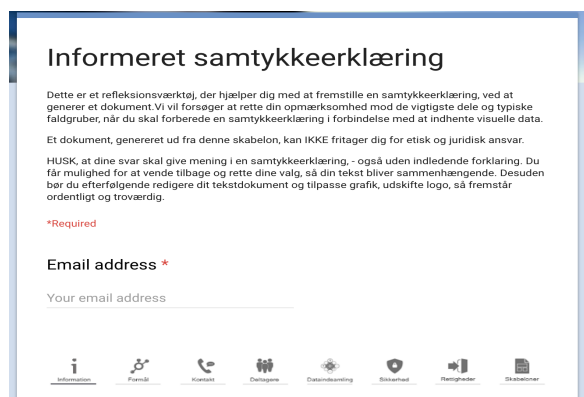


Figure 5: Screenshot of landing page of interactive Informed Consent Form for Video Research (Danish Version)

Activities	Deliverables
IT architecture / Design of prototype	Interactive mock-up for setting file and channel permissions. The prototype informs about a possible way in with minimal administrative overhead. The mock-up is kept in a code repository at the Royal Danish Library. Actual working prototype designed based on discussions around the mock-up as well as going through the requirements comparing what was desired with what was doable within the technical framework of Kaltura as well as the time available for the prototype implementation.
Finalizing the requirements for the upgrade of the existing Edumedia.dk platform	An upgrade of edumedia.dk would require: <ul style="list-style-type: none"> - reduction of administrative overhead to comply with GDPR laws - possibility to share video data and other relevant files through the same interface - reduction of administrative overhead for setting sharing permissions - possibility to generate shareable versions from any uploaded video files - possibility to supply public urls to individual video files for embedding in online publications - possibility to add and edit VDM-metadata
Design templates for data management plans	A review of the templates offered by https://DMPonline.deic.dk/ , which resulted a commented version of the DEFF DM-template. These ideas were incorporated into dataflowtoolkit.dk
Additional research support tool	interactive Informed Consent Form for Video Research (Danish and English Version)

Table 2: Activities and Deliverables Design of Prototype

2.3. Implementation of prototype

The delivered prototype interface for managing video research projects was developed by an external company; Reload, in close cooperation with KB. Reload and KB have worked together on Kaltura-based projects before, hence why Reload knew the Kaltura API up front and could step in and help the implementation when a critical resource had to leave this project for another position at SDU.

Reload implemented a new web application for the basic functionality of “starting a new video research project” as well as managing members of the research project. The new prototype application was developed in pure Javascript using the React.js framework as well as the standard Kaltura APIs (http://www.kaltura.com/api_v3/testmeDoc/).

The decision to include an external group (Reload) for the development of the web application halfway through the project timeline had initially not been anticipated and meant that we needed to focus on key functionality because of budget limitations. For this reason the decision was made to skip the implementation of WAYF-login, instead we used a dedicated login service from Kaltura. This meant that all users and testers within the project had to be added manually to the Kaltura platform before they could login. Since WAYF-login is working well in other known frontends to Kaltura, we did not need a proof of concept for it. Furthermore the dedicated login method made it easy to include researchers without a WAYF-login in the pilot test.

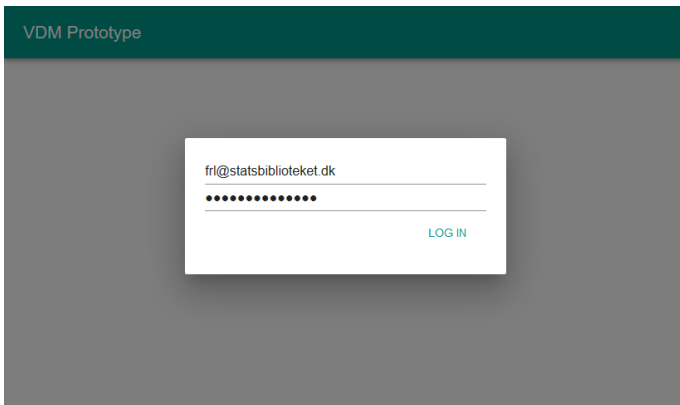


Figure 6: Login in box for the prototype

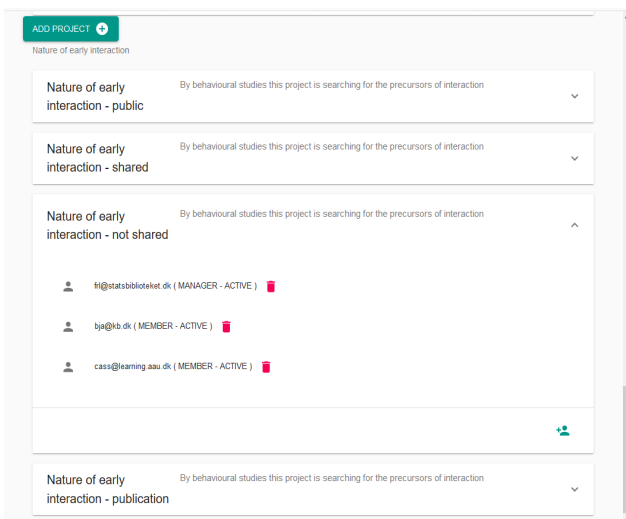


Figure 7: Working with a project in the prototype interface

In the prototype web interface four new Kaltura channels were added when a new project was started. These channels were setup with different rights settings as well as the correct members added to each of the necessary channels to have the four levels of sharing of content working right.

The work with Kaltura was originally planned to be done at the existing edumedia.dk (also called Edumedia2) platform that is also based on Kaltura. Budget was allocated in the project for additional licenses as well as hardware to be able to add another transcoding server to the platform and additional storage to cope with the pilot testing. The dialogue with Kaltura resulted in the company's interest in the VDM project and they agreed to sponsor the project by donating an extra license for Kaltura Mediaspace, including the required back end services. Thus, we did not have to work directly on the edumedia.dk platform but could work on a completely separated platform which could have presented problems for existing users of edumedia.dk.

The Kaltura Mediaspace interface was extended and re-worked with a focus in the prototype on:

- Additional Metadata capabilities
- Skinning to illustrate the VDM-project (<http://vdm.statsbiblioteket.dk>)
- Removing standard features not relevant to the VDM-project

The prototype went through two iterations, and the following table presents the issues identified in the prototype 1 and how they were addressed in the second generation prototype.

Issues identified in Pilot Test 1	Issues addressed / fixed for second pilot prototype
Difficulties finding relevant information on how to use the system.	Links to guides and information was added to the header menu in MediaSpace.
Workflow was based on four channels in MediaSpace. It was possible to deselect some of the channels in the VDM-admin-interface, which resulted in problems later.	The VDM-admin-interface was changed to make all channels mandatory. This also mitigated a request to redirect a "create channel"-function in MediaSpace to the VDM-admin-interface.
Users were unsure about how to delete a video.	This was made explicit in the pilot user guide.
Conflicting terminology between the VDM-admin-interface and MediaSpace regarding security settings on videos and channels.	The terminology in VDM-admin was changed.
Download of a video was disabled in the first pilot.	This was discussed and enabled only to the owner a video.
Uncertainty on how to upload a video.	Better descriptions were made and unused and confusing features in MediaSpace were removed.
More intuitive access to a user's own video.	Moving a link to the main menu header.
Public channel were not public.	Security settings for all four channels were revised.

Table 3: issues identified and addressed in Prototype 1

Implementation of integration with third party system (Clarín)

Integrating Clarín.dk was planned from the beginning to use an example to test how video-handling could be connected to other relevant platforms that hold video and/or metadata about video.

The integration used the Kaltura API to share a link to the video data stored in Kaltura. This link was used in CLARIN.dk to expose the sharable video data with resource metadata available in CLARIN.dk. It is important to note that only videos in Kaltura, that are in public channels and marked for publishing, can be exposed through CLARIN.dk.

Metadata that are exposed for a video are created in two steps. First, some of the metadata for the video inside Kaltura is fetched and marked for publishing. Second, a user gets the option to extend these metadata with more information. The

metadata that are added in the second step focuses on describing the videos in a way that allow external users, not being familiar with the data beforehand, to understand the provenance and content of the videos.

After the metadata has been curated for publishing in the CLARIN.dk curation workflow, the video metadata is assigned a PID which can then be used to cite the data in a persistent way.

CLARIN.dk then exposes the metadata via OAI-PMH allowing others to harvest the metadata. The metadata will be harvested by the European CLARIN metadata aggregator “CLARIN Virtual Language Observatory” (vlo.clarin.eu).

To work at production level the following points should be observed:

- Implement an administration interface where a system administrator can configure when to harvest the videos marked for publishing and which part of the metadata to harvest.
- Write a user guide that explains the different steps in the publishing process.
- It should also be documented very clearly in the VDM pilot interface how to publish the sharable videos and that completing the metadata as part of the CLARIN.dk curation workflow is needed to make the data citeable with a persistent identifier (PID).

The integration with CLARIN.dk can be seen as a proof-of-concept for the suggested architecture for publishing shareable videos. It is important to stress that only very selected data in the Kaltura backend can be expected to be shared as the videos are likely to be sensitive. It should also be checked that the CLARIN.dk and Kaltura integration does not allow for access to video collections that have to be protected as sensitive data.

Activities	Deliverables
Design of prototype	Documented in the project WIKI at the Royal Danish Library
Implement prototype interface	Working interface hosted at a cloud service managed by the company Reload. Code base for the prototype is kept at the Royal Danish Library
Implement adjustments to Kaltura Mediaspace	Working interface hosted at the Royal Danish Library (http://vdm.statsbiblioteket.dk). The design and functionality is documented as screen shots in appendix 3.
Second round of design and implement	The design and implementation steps were repeated based on feedback from the first round of testing.

Table 4: Activities and Deliverables Implementation of Prototype

2.4. Pilot testing of prototype

Pilot testing was conducted in two rounds:

Pilot 1 Test

The first pilot tested the prototype version 1 and was carried in out in July - September 2018 by researchers from the three research teams involved in the project (see deliverables). Five tests were carried out in total. The individual feedback from the testers was collected and summarised in a list of 16 issues (see deliverables). These issues were analysed and prioritised by importance from a user’s perspective and the ability to implement the change within the scope of the project.

Based on the feedback, the majority of issues were addressed by improving the prototype version 2, or by providing clearer guidance to the user about the functionality and the capabilities of the application. A few comments were asking for changes too advanced for the prototype to be solved within the scope of the project. Comments on the handling of issues are described in section 2.3.

Pilot 2 Test

The improved prototype version 2 was released at the end of September and immediately tested by external testers. Researchers from Arhus University and two international universities, the University of Oslo, Norway and University of Waikato, New Zealand, tested the second prototype. The researchers represented different user profiles, a researcher with a technical background, a researcher focussing on educational research, and a researcher with focus on the dissemination of research. The results from this round were also summarised. The main issues to pay attention to when going beyond the prototype are also highlighted below.

The overall feedback was very positive based on the intention to provide a video research platform that handles access rights, storage, and enables small editing (trimming) and sharing of videos in a flexible setup using “projects” with specific rights.

Issues to take into consideration are:

- To ensure that sensitive data can be handled safely and access logged a re-implementation of the frontend and the code accessing the Kaltura backend will be needed.
- WAYF user access should be included to ease the login for new users, and to ensure the identity and affiliation of users accessing possibly sensitive research data.
- Easy linking to consent forms or allowing for details about the content of the consent forms: what is allowed for which purposes: e.g. “Can be shared in project group”, “Can be used in education”.
- Consider to have a kind of workflow build into the system: upload (batch or single file), adding (detailed) metadata, curation by other person (PI or data steward) and publication/release of data for use.
- User Guide, help text and tips will be required together with a general introduction to what the system are focussing on. Text has to be easy to understand.
- The interface layout has to go through a review using fewer different fonts and symbols.
- The search function should be more advanced making it easy for users to find material based on different kind of searches, e.g. on tags, names that almost match.

Users have different needs and wishes about working with metadata. Perhaps the suggestion having a metadata template to pre-fill in for a project could solve this issue.

WAYF integration was not implemented in the VDM prototype, but using WAYF login service is known technology for both The Royal Library and the CLARIN group from University of Copenhagen. No major obstacles are expected to implement WAYF login.

Furthermore, the implementation did not address all aspects of handling sensitive data, neither the GDPR required logging. We did not conduct severe security tests to explore if sensitive data could be accessed.

Activities	Deliverables
Performing pilot tests using WAYF-login access (see notes above)	Test templates for user test in phase 1 Test template for user test in phase 2 Test template for CLARIN pilot testing Prototype Diagrams Notes (appendix 4) Notes on CLARIN pilot implementation (appendix 5)
Evaluation of pilot test and pilot results report	Summary report of test phase 1 (appendix 6) Summary report of test phase 2 (appendix 7)

Table 5: Activities and Deliverables Pilot Testing

3. Synthesis of findings and outcomes of the project

In this section we will present the synthesis of our findings. This synthesis is based on the comparison of the actual performance of the platform with the potential or desired performance. We have described "where we are" in this pilot project (at the present state) and will show now "where we want to be" (our target state). This allows us also to present

functionalities that have been left out, and functionalities which would need to be developed for a working video sharing platform.

The pilot project set out to explore and expand the following functionalities

1. Exploration of existing infrastructure, including: Kaltura, Edumedia2, DMPonline to identify extensions and develop & test prototype modifications.
2. Metadata schemas that fit the needs of video based researchers
3. Sharing mechanisms that will allow groups of researchers to share sensitive data with each other in a secure way
4. Persistent identifiers (PIDs) using the DanPID
5. Long term preservation of original video material and attached metadata using the National Bitrepository Software
6. Web interface
7. Download functionality within this interface
8. Integration with clarin.dk as a test case
9. Simple analytics to be able to report usage of the platform to relevant institutions
10. Added transcoding server to increase the capacity
11. Extra storage space
12. Sharing platform with simple video editing capabilities
13. Legal framework and necessary legal agreements between end users and the Royal Danish Library for the secure deposit of sensitive data.

The VDM project was an exploratory project and managed to carry out most of the planned work. Two of the originally anticipated aims were not possible to be developed. However, those are also significant findings and outcomes of the project:

- **Long term preservation in the National Bitrepository and usage of DanPID** as persistent identifiers: Because the analysis of existing infrastructures including Kaltura as the main technical platform in the focus in this project turned out to show that Kaltura is not suitable for long term preservation, this part of the original plan was skipped. Instead a permanent identifier was added for publication purposes through the test integration with the CLARIN.dk infrastructure.
- **Legal framework and necessary legal agreements** between end users and the Royal Danish Library for the secure deposit of sensitive data was not implemented. Unfortunately, the Royal Danish Library lost their internal legal advisor before this part of the project could be carried out and the position could not be filled in time.

Recommendations as a direct outcome from this pilot project to implement for a production-ready environment:

- **Metadata** was discussed and published to show both the researchers and the developers point of view (see publications) A simple version of VDM metadata was implemented in the resulting prototype.
- **WAYF-login:** Simple and known technology. Has been integrated with other front ends in other projects and provides security layers including updates of passwords etc. One future challenge may be if data is shared with researchers outside of Denmark without WAYF-login. Edugain may allow in such a case international logins. However, this may raise questions on how projects and storage is funded and who may pay for the service and especially if storage requirements are very large.
- **Separate frontend instead of Kaltura Mediaspace:** This project has tweaked Kaltura Mediaspace in many ways, and through this has it has identified certain limitations. We find that Kaltura Mediaspace is a generalised frontend that for this purpose lacks some features and in other areas has too many features. These limitations of Kaltura Mediaspace are not changed easily if at all. To build a well-functioning and user-friendly interface for working with video research data, the resources used to tweak MediaSpace further could be better used to implement a new tailored interface from scratch. The new implementation should include the functionality of the “manage video research project”-prototype interface (the Reload prototype development). We estimate that an initial version of such new interface will take several man months of expert developers work.
- **PIDs:** Implementation of PIDs for publishing / sharing. This was only implemented in the prototype implementation to the CLARIN-platform. The handle technology is well known and standard technology so it would be possible in the future in a tailored interface to add PIDs to relevant content inside the Kaltura platform. This could be either for all content or only for parts of the content, e.g. content on the “publication” channel of a project to be able to link with a permanent identifier in a research paper.

- **Guidance inside DMPonline:** During the pilot work we reviewed DMPonline and developed a list of recommendations that could be added to the DMPonline guidance. However, we recommend that necessary information could stay hidden within the DMPonline template. However, guidance inside DMPonline could link directly to other information portals to be able to maintain text snippets only once (for example what is contained in the DataFlowToolkit or the Interactive Informed Consent Template).
- **Further work with DataFlowToolkit:** DataFlowToolkit can be easily extended with more information on how to handle data. One initiative could settle around making extended and more detailed text between themes (metadata, security etc.) and processes (create, edit etc.). This would enhance the guidance. If necessary, a split could be made into different data object types, e.g. separating guidance for data from “simple” video observation from more advanced technology like 360 cameras and the like.
- **Legal framework** specific to the setup of a working version of a video sharing platform. Questions to be addressed: who owns what kind of rights to which kind of data on the platform? Both on the shorter time frame as well as in connection with any long term preservation functionality. Video data for educational research is highly sensitive data so new GDPR rules need to be addressed.
- **Front door landing page:** To be able to collect and publish valuable resources in a single place. This could potentially be implemented using a standard CMS-system depending on the amount of information and the numbers of editors. Could maybe be part of existing knowledge portals like the DeIC eScience Vidensportal (<https://www.deic.dk/da/vidensportal>). An advanced version would include personalised information depending on institutional relations.

4. Sustainability and vision for permanent infrastructure

We see that a sustainable future implementation of a sharing platform for video based research has to first establish financial responsibility to run such service at national level. It would seem natural that DeIC could play a major role in the setup and running and the Royal Danish Library with regards to long term preservation. We recommend that the latter should also be discussed with the National Archives and other relevant parties. DIGHUMLAB might also play a role in the organisation of such a service — e.g. on service desk, training etc. DIGHUMLAB or DeIC could also play a role in providing the platform for support services that provide guidance on video research more generally, video based data management, and ethical practicalities (how to prepare informed consent). For those auxiliary services the VDM project has already produced resources that can act as blueprints to be used as is or these resources can be updated/ further developed (see example ethics template). These activities could be developed by some of the existing partners in the VDM project. Since the vision is to develop infrastructure that supports researchers it is advisable to involve researchers not only to, inform the development of such tools, but to also raise the awareness of the research community on how to work with a video sharing platform, observe the codes of conduct regarding the handling of sensitive data, and manage video data at a sophisticated level.

The VDM project team

Aalborg University

Kathrin Otrell-Cass (project leader, lead AAU), Lone Dirckinck-Holmfeld, Thomas Ryberg, Karsten Kryger Hansen

The Royal Danish Library

Bjarne Andersen (lead KB), Frank Lund, Erik Bertelsen, Kåre F. Christiansen

Copenhagen University

Lene Offersgaard (lead KU), Mitchell John Seaton, Costanza Navarretta

University of Southern Denmark

Max Roald Eckardt (lead SDU), Julia Ruser

List of appendices

Appendix 1 VDM project publications State of the art and VDM publications

Appendix 2 VDM Needs Analysis

Appendix 3 Screenshots of Prototype for video sharing

Appendix 4 VDM Prototype Diagrams Notes

Appendix 5 VDM-CLARIN-pilot_Short_notes_about_implementation

Appendix 6 TestPhase1_UserTest_Summary

Appendix 7 TestPhase2_UserTest_Summary_20181113