NATIONAL ENDOWMENT FOR THE HUMANITIES

OFFICE OF **DIGITAL HUMANITIES**

Narrative Section of a Successful Application

The attached document contains the grant narrative and selected portions of a previously funded grant application. It is not intended to serve as a model, but to give you a sense of how a successful application may be crafted. Every successful application is different, and each applicant is urged to prepare a proposal that reflects its unique project and aspirations. Program guidelines also change and the samples may not match exactly what is now required. Please use the current set of application instructions to prepare your application.

Prospective applicants should consult the current Office of Digital Humanities program application guidelines at https://www.neh.gov/grants/odh/digital-humanities-advancement-grants for instructions.

Applicants are also strongly encouraged to consult with the NEH Office of Digital Humanities staff well before a grant deadline.

Note: The attachment only contains the grant narrative and selected portions, not the entire funded application. In addition, certain portions may have been redacted to protect the privacy interests of an individual and/or to protect confidential commercial and financial information and/or to protect copyrighted materials.

Project Title: *Distant Viewing Toolkit (DVT) for the Cultural Analysis of Moving Images*

Institution: University of Richmond

Project Directors: Lauren Tilton and Taylor Arnold

Grant Program: Digital Humanities Advancement Grants, Level II

2. List of Participants

Principal Investigators

- Taylor Arnold, University of Richmond
- Lauren Tilton, University of Richmond

Technical Team

- Nathaniel Ayers, University of Richmond
- John Bell, Dartmouth College
- Mark Williams, Dartmouth College

Case Studies Team

- Annie Berke, Hollins University
- Claudia Calhoun, New York University
- Jenny Oyallon-Koloski, University of Illinois
- Bret Vukoder, Carnegie Mellon University

Advisory Board

- Paul Achter, University of Richmond
- Jeremy Butler, University of Alabama
- Karen Cariani, WGBH / American Archive of Public Broadcasting
- Trevor Muñoz, University of Maryland
- Miriam Posner, UCLA
- Holly Rushmeier, Yale University

4. Narrative

Enhancing the Humanities

Introduction Moving images have served as a dominant form of cultural expression in the United States since the beginning of the twentieth century. Millions flocked each week to the theaters after World War I and then tuned into television after World War II. Today, Americans spend over 1.5 hours each day, on average, streaming digital media over the internet (Nielsen 2016). Extensive scholarship in media studies has established how formal elements of moving images — such as camera angles, sound, and the construction of narrative arcs — reflect, establish, and challenge cultural norms (Mulvey 1975, Braudy 2002). In other words, moving images offer a lens into a community's ideals and values. The study of culture in the twentieth century requires, therefore, considering media as a serious source of historical evidence. However, with currently available tools, the formal analysis of moving images has been restricted to a small set of works capable of being studied by close analysis. With a rapidly growing set of large archives containing digitized time-based media, computational approaches are now needed to take full advantage of the available archival material. Tools capable of applying algorithmic approaches to moving images stand to open exciting new avenues of research in digital humanities, cultural analytics, and media studies.

Project Goals We seek Level II start-up funding to build the Distant Viewing Toolkit (DVT), an open source software library for studying large collections of moving images. The project is in collaboration with the the Digital Scholarship Lab (DSL) at the University of Richmond and the Media Ecology Project (MEP) at Dartmouth College. Funding is requested for the development and extension of cutting-edge techniques in computer vision to facilitate the algorithmic production of metadata describing the content (i.e., people/actors, dialogue, scenes, objects) and style (i.e., shot angle, shot length, lighting, framing, sound) of time-based media. This information can be used to visualize, summarize, search, and explore digitized corpora of moving images. These techniques will allow scholars to see content and style within and across moving images such as films, news broadcasts, and television series, revealing how moving images shape cultural norms.

To illustrate how the DVT Toolkit might be used in practice: Once completed, we will make a "software library" of tools available for free download. For example, humanities scholars, media librarians, and even students will be able to download the toolkit and install it on their personal computer. Once installed, the user can either load their own content for analysis or import preprocessed datasets from our website. The DVT software is able to automatically build a report (i.e., the "metadata") describing formal stylistic and content-driven aspects of moving images (see pages 19-22 for examples). The user can then explore the metadata using intuitive tools included in the DVT library to gain a better aesthetic understanding of how the media is constructed and to compare how these aesthetics change across various dimensions such as over time or across different series. Ultimately, the software will allow for the analysis at scale of formal elements that are currently only accessible through close analysis.

Because it is critical that the DVT Toolkit is useful to a wide variety of scholars, the toolkit will be designed to work with a wide variety of content, including news broadcasts, dramatic series, musicals, and comedies. To ensure this, during the grant period we will be testing the toolkit with a variety of scholars studying different kinds of moving images. Specifically, the corpora to be examined as case studies are: early Hollywood feature films (Jenny Oyallon-Koloski, Illinois-Urbana Champagne); mid-century educational films produced by the US Government (Bret Vukoder, Carnegie Mellon University); Network Era comedies (Annie Berke, Hollins University); police procedurals (Claudia Calhoun, NYU); and local television news broadcasts (Lauren Tilton, University of Richmond). These case studies not only demonstrate the utility of the toolkit, but also provide our team with valuable feedback on how it should function.

Problem Statement There is an increasing call to take seriously visual culture and moving images as objects of study in the humanities (McPherson 2009; Posner 2013; Acland and Hoyt 2016). Over the past 20 years, we have seen the development of exciting digital humanities tools that focus on text and, increasingly, methodologies such as distant reading and macroanalysis that have produced exciting interventions in areas such as literary studies (Jockers 2013; Moretti 2013). However, tools for analyzing non-textual materials have been lacking. This shift to turn the humanities' focus to visual culture comes as an increasing number of archives, including the Internet Archive and the Library of Congress, have digitized and published large corpora of culturally important time-based media. Simultaneously, recent shifts in media studies have questioned the logic of the traditional canon of works, which privileges film over other mediums and is highly concentrated on a small set of well-known directors and time periods. New research has turned to the importance of areas such as television, further highlighting the necessity for new methods that can analyze large sets of moving images at scale. The recent launch of the *Journal of Cultural Analytics*, an open-access journal for the computational study of culture, offers another testament to the growing field of scholars eager to use computational techniques to study a variety of cultural forms.

Excitement surrounding the use of computational techniques for studying visual culture is certainly warranted. The analysis of movie images at scale offers great potential for addressing important humanities-driven research questions. Consider, for example, a scholar interested in studying the way in which news was disseminated within different communities in the 1980's. Corpora available from the American Archive of Public Broadcasting now provide access to thousands of hours of material from local news shows such as *Evening Exchange* (WHUT-TV, Washington, D.C.), hosted by Kojo Nnamdi, and shows with a national distribution such as *The MacNeil/Lehrer Report.* Are there noticeable patterns in how a local show is blocked relative to a show produced for a national audience? Where are MacNeil and Lehrer framed relative to guests on the show? Does this framing change over time? How does it compare to the framing of Nnamdi and his guests? Does the framing reflect or contrast with power dynamics between the hosts and their guests? How are race, gender, and class performed in the visual and sonic spaces of the shows? Analyzing questions such as these requires a careful analysis of subtle patterns over thousands of hours of material — a difficult task to perform by hand. Addressing questions regarding the formal aesthetic qualities of large moving images corpora, in contrast, is a task ideally suited for study by computational analysis.

Yet, extending existing digital humanities methods intended for the analysis of text to other media forms is not trivial. There are many aspects of textual data that have made it particularly accessible to computational approaches. For example, relative to other forms of media, digitized text takes up a minimal amount of storage space. Consider the collected works of William Shakespeare. The digitized text of all his plays, sonnets, and other materials can be saved in a file taking up under 7 megabytes of storage space. In comparison, the standard size of just a single photo on a modern smart-phone is currently between 5 and 12 megabytes. Even heavily compressed, a half-hour video file digitized in standard definition requires over 1000 megabytes of storage. The small file sizes of textual data makes it possible to run analyses without requiring particularly fast or efficient code. In contrast, non-textual data's scale often far exceeds that of text, requiring highly specialized techniques to summarize and access the data on standard computers.

Another challenge is the development of methods for analyzing non-textual data. The way in which text is natively stored in digital form lends itself to computational analysis. Digitized text is stored as a stream of characters (letters and punctuation marks) and simple algorithms are able to group these characters into individual words, a process known as tokenization. Words are important because they have an extrinsic meaning within the context of a language. Techniques such as 'keywords in context' (KWiC) and topic modeling operate directly by locating and counting words. Unfortunately, the analogous approach to studying sound and image data is significantly more difficult. An individual pixel within an image file

has almost no meaning out of context. Grouping pixels into objects, the task of image segmentation, is known to be a difficult task currently undergoing active research (Zagoruyko 2016). Even when applied successfully, grouped pixels do not uniquely define an object in the same way that groups of characters do. For example, there are only a few words that describe the concept of a dog (i.e., 'dog', 'canine', 'hound') but an infinite combination of pixels that can depict a particular instance of a dog. As a result, new specialized methods and techniques are necessary to analyze non-textual data.

Proposed Approach The DVT software library addresses the challenges of working with moving images by summarizing media objects through the automated detection of stylistic and content-driven metadata. It algorithmically approximates the ways in which humans process moving images by identifying and tracking objects, people, sound, and dialogue. As a result of advances over the past two years in deep learning and computer vision, it is now possible to build models capable of automatically performing these annotation tasks with human-like accuracy (He 2016; Szegedy 2017). These advances, combined with the recent increased interest in and access to large corpora of moving images, make this the perfect time for building an automated annotation tool specifically designed for application in the humanities.

The DVT software library will work by allowing users to input raw media files in a variety of formats. The input files will then be analyzed to detect the following features: (1) the dominant colors and lighting over each shot; (2) time codes for shot and scene breaks; (3) bounding boxes for faces and other common objects; (4) consistent identifiers and descriptors for scenes, faces, and objects over time; (5) time codes and descriptions of diegetic and non-diegetic sound; and (6) a transcript of the spoken dialogue (see pages 19-22 for examples). These features serve as building blocks for the analysis of moving images in the same way words are the foundation for text analysis. From these extracted elements, higher-level features such as camera movement, framing, blocking, and narrative style can be derived and analyzed.

DVT offers two output formats. The first provides data stored as a self-contained interactive website. The page can be opened locally in any web browser and requires no technical programming expertise from the user. This format provides tools to explore and visualize the extracted information. The second output format consists of a collection of plain-text JSON files. These files are optimal for technical users looking to integrate the output into larger analytic pipelines such as building a visual search interface within a public-facing archive. In either output format, the generated metadata is significantly smaller in size compared to the raw media files, making it easy to share and publish the resulting data files.

In order to extract the metadata elements, DVT will make direct use of several specific deep learning frameworks and models. The toolkit utilizes the architecture of three open source programming libraries: dlib (King 2009), ffmpeg (Tomar 2006), and TensorFlow (Abadi et al. 2016). Within these frameworks, novel computer vision and sound processing algorithms extract the required features. Specifically, the project draws from OpenFace (Amos et al. 2016) for face detection; YOLO9000 for object detection (Redmon 2017); the Places-CNN (Zhou et al. 2016) for scene detection; Colorization (Zhang et al. 2016) for working with black and white images; GOTURN for object tracking (Held 2016); and CMUSphinx (Lamere 2003) for converting sound to text. These specific algorithms were chosen due to their open-source licenses, use of the most up-to-date techniques, and the institutional support behind the algorithms at CMU, MIT, and Berkeley. Our work in building DVT consists in modifying and stitching together these six models for our specific humanities-centric needs. For example, only one of these models works directly with moving images, taking only still images as inputs, and only one is able to process black and white images. Our toolkit will extract individual frames, colorize if necessary, apply each algorithm to the frames, and then intelligently combine the results into a single cohesive structure.

The six models we are building in the DVT software library will allow for the use of new, domain spe-

cific data to improve the performance of a generically trained algorithm. This process, known as transfer learning, is one of the key reasons for the popularity of deep learning in machine learning. A major part of building the DVT library will be applying transfer learning to tweak the open-source computer vision algorithms to better function on moving images. This will be done by first hand-labeling a training set of 10,000 still frames with information about common objects and characters found in each frame (tasks 3 & 4). Likewise, sound and shot break information will be hand-recorded from several hours of raw material (tasks 2 & 5). Time coded closed captioning is available for some of our source materials and can be used to create training data for speech recognition (task 6). These hand labeled datasets can be used to apply transfer learning to the base model, updating them for our specific humanities-focused application tasks. Putting the models together, along with extensive documentation, yields the DVT software library, capable of creating summary metadata directly from raw media files.

Testing with Scholars/Case Studies Along with developing the toolkit, we will also apply our methods to a series of case studies in cooperation with scholars on our Case Studies Team. The case studies allow us to illustrate how the approach offered by the DVT software can be used to address a set of humanities questions in novel ways. It will also provide valuable feedback for ensuring that design decisions address the way it will be employed by other scholars.

The Case Studies Team consists of four media scholars investigating questions from a wide variety of corpora. Jenny Oyallon-Koloski (Illinois - Urbana Champagne) studies the use of movement in cinematography. She will be applying the toolkit to early feature-length films from both France and Hollywood. Her particular focus will be studying how Florence Lawrence, often credited as the 'The First Movie Star', was blocked and framed over her career starring in nearly 300 silent films. Bret Vukoder's (Carnegie Mellon University) work focuses on the cultural analysis of mid-century American feature-films. Vukoder will be applying the toolkit to a collection of over 2,500 informational films from the 1950's produced by the Federal Government and held by the National Archives with a focus on editing practices and use of stock footage. Claudia Calhoun (NYU) is a scholar of television studies, with a focus on the formal analysis of race and gender in police procedurals. Annie Berke (Hollins) specializes in the intersection of gender studies and comedic television. Calhoun and Berke will, in collaboration, apply the DVT library to a set of 22 US series from the Network Era of television. Their focus will be on the use of non-diegetic sound for comedic and dramatic effect. Tilton will additionally join the case studies team by applying the DVT library to a comparative analysis of the news shows *Evening Exchange* and *The MacNeil/Lehrer Report.*

The various specializations of the Case Studies Team ensures that the DVT library will assist in producing significant scholarly contributions to the field of media studies. Each scholar will provide us with valuable feedback to help improve the DVT system. The team will also serve as liaisons to the community in order to publicize the project's work and the DVT toolkit. In the final phase, they will be co-authors on the humanities-focused white paper.

Environmental Scan

Beginning with the early work of Barry Salt (Salt 1974), quantitative methods for analyzing moving images have focused predominantly on the distribution and patterns of shot lengths. Prominent examples include Yuri Tsivian's Cinemetrics project (`http://www.cinemetrics.lv/`), *Arclight Guidebook to Media History and the Digital Humanities* (Acland and Hoyt 2016), and Jeremy Butler's ShotLogger (`http://shotlogger.org/`). These projects demonstrate the feasibility of distributing extracted metadata from copyrighted materials and the power of computational techniques to extract useful information over a large collection of moving images. However, there is much to be learned from other extractable metadata beyond shot detection.

A limited number of prior attempts have been made to produce a set of tools for extending the quantitative approach of Salt to other parameters. Examples include the (now defunct) *Videana: A Software Toolkit for Scientific Film Studies* (Ewerth et al. 2009), *ACTION: Audio-visual Cinematic Toolkit for Interaction, Organization, and Navigation* (Casey and Williams 2011) for detection of camera movement, and the language and color analysis of Burghardt, Kao & Wolff (2016). These prior tools pre-dated recent advances in deep learning and focused only on a small set of extracted features. DVT expands on these by incorporating the significantly richer features including face and object detection, sound identification, and dialogue extraction. Capturing a much wider and more granular set of features, the DVT project gives a more complete view of the formal decisions made by actors, writers, camera operators, directors, and editors.

History of the Project

The project has developed over the past year and a half. Following informal discussion at several conferences including the American Historical Association 2016, DH2016, and the Joint Statistical Meetings 2016, the idea was more fully presented at the IMLS-funded *Image Processing and Reunification Workshop* held at the University of Maryland in September 2016. Arnold and Tilton then began an initial prototype of the DVT library, presenting early results at DH2017 and the Joint Statistical Meetings in August, 2017. An initial repository holding a preliminary version of the DVT library can be found at `github.com/statsmaths/dvt`; early results are documented in blog posts published at `distantviewing.org`. Recently, we began collaborating with the Media Ecology Project (MEP) to make use of their software for manual feature annotation, the NEH-funded Semantic Annotation Tool (SAT; PR-234316-16). The tool will be used by student research assistants to manually tag training datasets for tuning the deep learning algorithms for humanities applications. Further, the DVT team is able to utilize MEP's large network of media archives and scholars.

Work Plan

The grant period will be split into three distinct phases (see page 23 for a graphical description). The six months of Phase I will focus on building training datasets for transfer learning (by Tilton) and producing the high-level workflow of the DVT tool using the dlib, ffmpeg, and TensorFlow frameworks (by Arnold). Training data will be annotated using the Semantic Annotation Tool (SAT) of Williams and Bell. Mid-way through Phase I, the technical team will meet for two days at the University of Richmond. The focus of the meeting will be on developing the output data formats. A draft agenda is included in the appendix of the grant (page 24). Phase II will last 9 months, kicked off by virtual meetings with members of the advisory board. These will consist of showing a demo of the tool and initial results, with extensive time to collect feedback. The remainder of the phase consists of implementing and applying transfer learning to the vision and sound processing algorithms (by Arnold). At the end of Phase II, the Case Studies Team will convene a two-day meeting at the University of Richmond. Phase III of the project focuses on documenting the DVT tool, writing up the results of the case studies, and publishing the toolkit. The core team will work collaboratively to draft a paper focusing on the technical aspects of the project; along with the Case Studies team, a report on the humanities aspects of the project will also be produced. Virtual meetings with the advisory board will present an initial version of these papers and the toolkit. DVT will be officially launched at the 2020 annual meeting of the Society of Cinema and Media Studies.

Staff

Taylor Arnold and Lauren Tilton are the grant's co-directors and co-principal investigators. They have a considerable history of successfully blending humanities scholarship and computational methods. Their NEH (HD-51421-11) and ACLS-funded project *Photogrammar* applied image, mapping, and textual analysis to the study of Depression-Era photography, and NEH-funded (MD-234145-16) *Participatory Media* is

producing a digital platform for the curation of community film from the 1960s through the present day. They also wrote the book *Humanities Data in R* (Springer 2015) and several research articles addressing the power of working at the intersection of statistics and the humanities.

Taylor Arnold, a computational statistician focused on applications to text and image analysis, will spend all of his research time during the grant's 18-month duration working on the project. He has considerable experience building and deploying machine learning software in government, industry, and academic settings. In order to make the tools accessible to other projects, all of the modules will be extensively documented and made publicly available under an open source license. Arnold will also oversee the building of the javascript visualizations and public-facing website.

A scholar of moving images and digital humanities, Lauren Tilton will spend her research time during the grant working on the project. In the first two phases of the grant, this will consist of working with the students who are labeling the training sets that will go into the DVT library and assisting with the case studies. During the final phase of the grant, she will be focused on interpreting the humanities impact of the results and synthesizing these as talks, interactive visualizations, and two scholarly papers. She will engage in a case study comparing runs of two local television news broadcasts. Tilton will also take the lead on working with the project's advisory board.

Mark Williams and John Bell (Media Ecology Project at Dartmouth) along with Nathaniel Ayers, the technical head of the Digital Scholarship Lab, complete the technical project team. Williams is the Director of The Media Ecology Project – a network of researchers and tools working to provide access to moving images in order to facilitate archival discovery and new scholarship. John Bell, along with Williams, has designed the Semantic Annotation Tool (SAT) for the manual annotation of moving images. The DVT team will be using this tool to manually annotate images as an input to the DVT machine learning algorithms. Ayers has extensive experience developing public digital projects including *Mapping Inequality*, *Renewing Inequality,* and *Forced Migration.* He brings expertise in graphic design, user experience, and front-end web programming languages. Ayers will spend 5% of his time on the project during the grant period, focusing on the design and deployment of the interactive output format of the DVT library.

The Case Studies Team, highlighted previously, brings the additional expertise of four prominent media scholars to the DVT team. Furthermore, the advisory board consists of 6 additional experts from both humanities and technical backgrounds. The board's primary role, as laid out in the work-plan, will be participating in two video calls with the core team to offer feedback and suggestions on the project's progress.

Final Product and Dissemination

The primary final product at the end of the grant period will consist of the DVT software library. Additionally, there will be a public website showcasing the metadata extracted from our five case studies and two open access scholarly journal articles. One article will focus on the toolkit and the other will center on the results of the case studies. A blend of these two papers will form the project's final white paper. These elements will be disseminated to various communities as we present the project at: DH 2019, the Joint Statistical Meetings in 2019, and the Society for Cinema and Media Studies in 2020. In order to publicize the toolkit and website to a wider public audience, we will utilize the social media and journalistic network from the various public and digital humanities centers that we and our advisory board belong.

A downloadable application version DVT library will be available from the project's website. Source code for the library will hosted on GitHub under the GPLv2 license for ease of discovery and use. Usage of DVT software library requires no prior programming experience.

5. Biographies

Project Directors

Taylor Arnold is Assistant Professor of Statistics at the University of Richmond. A recipient of grants from the NEH and ACLS, Arnold's research focuses on computational statistics, text analysis, image processing, and applications within the humanities. His first book *Humanities Data in R* (Springer, 2015) explores four core analytical areas applicable to data analysis in the humanities: networks, text, geospatial data, and images. His second book, the forthcoming *A Computational Approach to Statistical Learning* (CRC Press 2018), explores connections between modern machine learning techniques with theories of statistical estimation. Numerous journal articles extrapolate on these ideas in the context of particular applications. Arnold has also released several open-source libraries in R, Python, Javascript and C. Visiting appointments have included Invited Professor at Université Paris Diderot and Senior Scientist at AT&T Labs.

Lauren Tilton is Assistant Professor of Digital Humanities in the Department of Rhetoric and Communications at the University of Richmond and a member of Richmond's Digital Scholarship Lab. Her current book project focuses on participatory media in the 1960s and 1970s. She is the Co-PI of the project *Participatory Media*, which interactively engages with and presents participatory community media from the 1960s and 1970s. She is also a director of *Photogrammar*, a web-based platform for organizing, searching and visualizing the 170,000 photographs from 1935 to 1945 created by the United States Farm Security Administration and Office of War Information (FSA-OWI). She is the co-author of *Humanities Data in R* (Springer, 2015). She is co-chair of the American Studies Association's Digital Humanities Caucus.

Technical Team

Nathaniel Ayers is a member of the Digital Scholarship Lab at the University of Richmond, serving as the head of the Lab's technical work and providing technical assistance to faculty and students. His most recent project is *Mapping Inequality*, which introduces viewers to the records of the Home Owners Loan Corporation (HOLC) through 150 interactive maps and roughly 5000 individual area descriptions. Previous projects under the American Panorama umbrella include *ForcedFDH Migration* and *Atlas of the Historical Geography of the United States*. A graduate of Virginia Commonwealth University, School of the Arts, Nathaniel has previously done programming and visualization work for the University of Virginia.

John P. Bell is currently the Lead Applications Developer in the ITC department at Dartmouth College. Bell serves as a software developer, data artist, and teacher at both Dartmouth College and the University of Maine. Ranging from straightforward and utilitarian, the running theme across his projects is a desire to merge concept, design, and technology to convey a single, compelling argument to his audience. The diversity of his work is reflected in the variety of places it has been featured, spanning arts festivals including Ars Electronica and ISEA, research groups like DOCAM and Harvard's Berkman Center for Internet & Society, and publications ranging from WIRED to the Chronicle of Higher Education. John Bell is the Architect of The Media Ecology Project.

Mark Williams is Associate Professor of Film and Media Studies at Dartmouth College. He has previously taught at USC, Loyola Marymount, UC Santa Barbara, and Northwestern. His courses at Dartmouth include surveys of U.S. and international film history, television history and theory, and new media history and theory. He has published in a variety of journals and anthologies. He directed the Leslie Center Humanities Institute entitled Cyber-Disciplinarity. In conjunction with the Dartmouth College Library, he is the founding editor of an e-journal, The Journal of e-Media Studies. With Adrian Randolph, he co-edits

the book series Interfaces: Studies in Visual Culture for the University Press of New England. He founded and has twice directed the Dartmouth off-campus program in Los Angeles for The Department of Film and Media Studies. In 2014 he received an award for Scholarly Innovation and Advancement at Dartmouth for directing The Media Ecology Project. In 2015 he received an NEH Tier 1 Research and Development grant with John Bell to build the Semantic Annotation Tool (SAT) for use in The Media Ecology Project. His book *Remote Possibilities, a History of Early Television in Los Angeles*, will be published by Duke University Press.

Case Studies Team

Annie Berke is Assistant Professor of Film at Hollins University with areas of expertise in American film and television, media industry studies, and gender and sexuality. Her book project is *You Just Type*: Women Television Writers in 1950s. She has been published in such journals as the *Historical Journal of Film, Radio and Television*, *Film Quarterly*, and *Feminist Media Histories*. She has organized numerous conferences, including the *Auteurs in the 21st Century* and the *On Television* conferences. She is currently the faculty co-chair of the Comedy and Humor Studies Scholarly Interest Group (Society of Cinema and Media Studies). She completed her Ph.D. in Film & Media Studies and American Studies, with a concentration in Women's and Gender Studies, at Yale University, as well as her Masters in Film Studies at Columbia University.

Claudia Calhoun is Visiting Assistant Professor of Cinema Studies at NYU. Her research asks how media objects work as agents of historical change. She is currently working on her first book manuscript. Tentatively titled, *More Than Facts: Dragnet and Postwar Culture*, the project uses the radio and television series Dragnet (1949 - 1959) to show how the police procedural genre served a pedagogical purpose after World War II, bringing Americans into the justice system and instructing them in their role as citizens. In addition to her work on *Dragnet*, Claudia has written for academic and popular outlets on a range of subjects, including the George Stevens's 1956 epic *Giant*, gender in the academy, and teen idols. Her next scholarly project looks at narratives of race and integration in U.S. postwar film and television.

Jenny Oyallon-Koloski is Assistant Professor of Media and Cinema Studies at the College of Media at the University of Illinois. Her work applies a sophisticated interdisciplinary methodology to study human movement in media with a historical and medium-specific focus. Her current project provides a critical intervention in the way film scholars study figure movement and dance on screen. Oyallion-Koloski combines the methodological tools of film studies with those of dance studies, specifically Laban Movement Analysis (LMA), in which she is certified by the Laban/Bartenieff Institute of Movement Studies. As case studies for this methodology, she analyzes key films by the French director Jacques Demy and offer a revisionist history of his career and relationship to genre cinema and the Hollywood musical. Oyallion-Koloski also serves as the movement analysis expert for the Media Ecology Project, based at Dartmouth College. She have developed an annotation system for scaled close and distant analysis to analyze and document gestural patterns for MEP's pilot project on Florence Lawrence and performance in silent cinema.

Bret Vukoder is a doctoral candidate at Carnegie Mellon University in the Department of English. He received a B.S. in Business Administration with majors in Public Administration, English, and Political Science and a minor in Cinema Studies at the University of Tennessee. He also earned a Masters in English Literature at Tennessee. Broadly, Vukoder's research concerns film —- particularly American film -— and the extent to which it forms, reinforces, and disseminates cultural practice. More specifically, he is interested in both Hollywood and state-sanctioned/produced films that were constructed and appropriated as ideological weapons during the Cold War, the conditions of production behind many of these films, and the channels by which they were conveyed.

Advisory Board

Paul Achter earned a Ph.D. in Speech Communication from the University of Georgia in 2001 with an emphasis on rhetorical theory and criticism, media criticism, and political culture. He then began work as a postdoctoral associate with Drs. Celeste Condit and Roxanne Parrott, conducting research and managing teams working on CDC and NIH grants about public understandings of race and genetics. Achter is now associate professor and Chair in the Department of Rhetoric and Communication Studies at the University of Richmond. His work has appeared in scholarly journals including the *Quarterly Journal of Speech*, the *Southern Communication Journal*, and *Critical Studies in Media Communication*. He has also written in popular media such as CNN.com and Huffington Post. At Richmond, he teaches courses such as public speaking, war rhetoric, and television criticism.

Jeremy Butler is Professor of Telecommunication and Film at the University of Alabama. He has taught television, film, and new media courses since 1980 and is active in online educational resources for television and film studies. He is the author of *Television Style* (Routledge 2009) and *Television: Critical Methods and Applications* (Routledge 2011). He is the creator and maintainer of the digital projects *Shot Logger* and *Laugh Logger*.

Karen Cariani is Director of the WGBH Media Library & Archives. Cariani oversees the American Archive of Public Broadcasting (AAPB) as the project's director, in collaboration with the Library of Congress. In addition to duties overseeing WGBH archive, she is also responsible for fundraising for the MLA and AAPB, managing and promoting projects that move our initiatives forward. As always there is great emphasis on making the materials accessible for use by researchers, scholars, and educational purposes. Cariani is the co-director of two grants awarded from the Mellon Foundation to assess the educational value of the WGBH collection and one to build a prototype using Fedora as a digital repository to deliver media materials to scholars. Features added to Open Vault included faceted search, ability to annotate and tag video or stills, searchable text transcripts synced with videos.

Trevor Muñoz is Assistant Dean for Digital Humanities Research at the University of Maryland Libraries and an Associate Director of the Maryland Institute for Technology in the Humanities (MITH). He works to foster digital projects that involve close collaboration between librarians, archivists, and other digital humanities researchers. As part of this work, he has written, spoken, and consulted about the strategic opportunities and challenges of doing digital humanities work within the institutional and cultural structures of academic research libraries. Muñoz holds an MA in Digital Humanities from the Department of Digital Humanities at King's College London and an MS in Library and Information Science from the Graduate School of Library and Information Science at the University of Illinois, Urbana-Champaign. He currently serves as a core team member for the "Synergies among Digital Humanities and African American History and Culture" (AADHum) initiative.

Miriam Posner is an Assistant Professor in the Information Studies department at UCLA. She also serves on the executive committee of the Association for Computers and the Humanities. A media scholar who's interested in science and technology, Posner current book project, *Depth Perception*, on American medical filmmaking is under contract with UNC Press. She is currently working on a new project related to supply-chain capitalism.

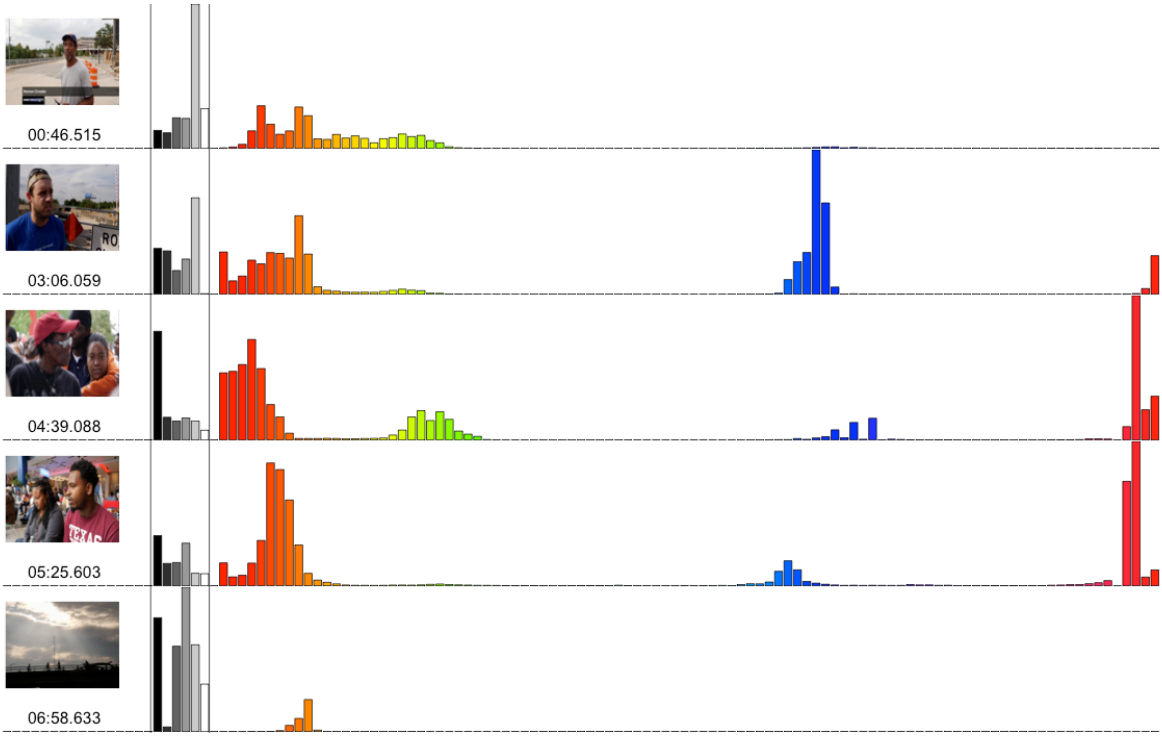
Holly Rushmeier is a professor in the Yale Department of Computer Science. Her research interests include shape and appearance capture, applications of perception in computer graphics, modeling material appearance, and developing computational tools for cultural heritage. Rushmeier was Editor-in-Chief of ACM Transactions on Graphics from 1996-99 and co-EiC of Computer Graphics Forum (2010-2014). She has also served on the editorial boards of IEEE Transactions on Visualization and Computer Graphics,

ACM Journal of Computing and Cultural Heritage and IEEE Computer Graphics and Applications. She currently serves the editorial boards of ACM Transactions on Applied Perception, ACM Transactions on Graphics, the Visual Computer and Computers and Graphics. In 1996 she served as the papers chair for the ACM SIGGRAPH conference, in 1998,2004 and 2005 as the papers co-chair for the IEEE Visualization conference and in 2000 as the papers co-chair for the Eurographics Rendering Workshop. She has also served in numerous program committees including multiple years on the committees for SIGGRAPH, IEEE Visualization, Eurographics, Eurographics Rendering Workshop/Symposium, and Graphics Interface.

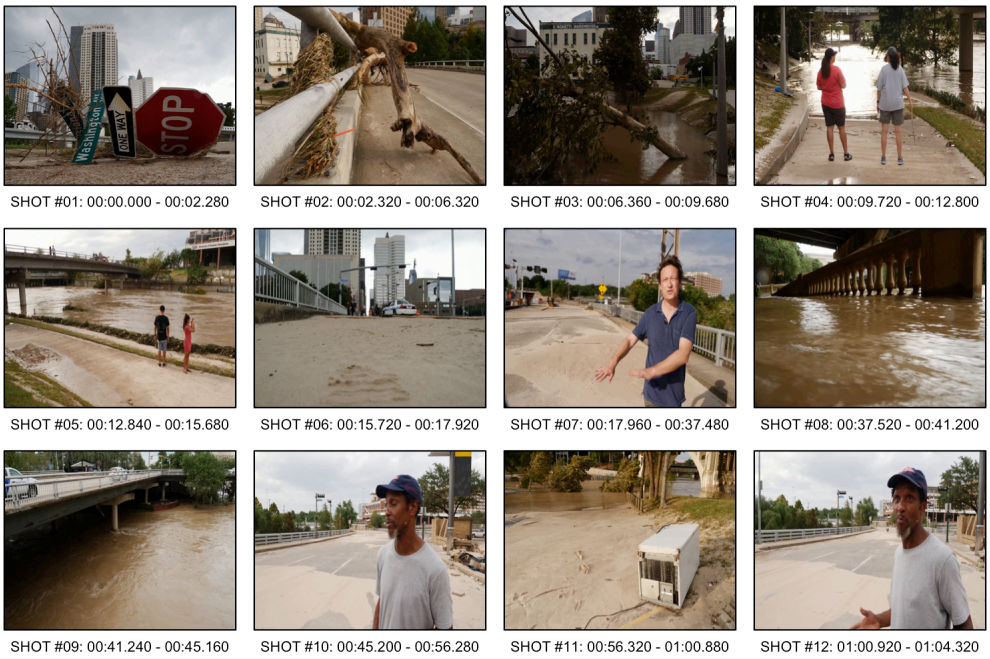
7. Appendices

Example Input Video An mp4 video file showing a segment from BBC's Newsnight program reporting from Houston on the aftermath of Hurricane Harvey (31 August 2017). Total runtime: 6m59s

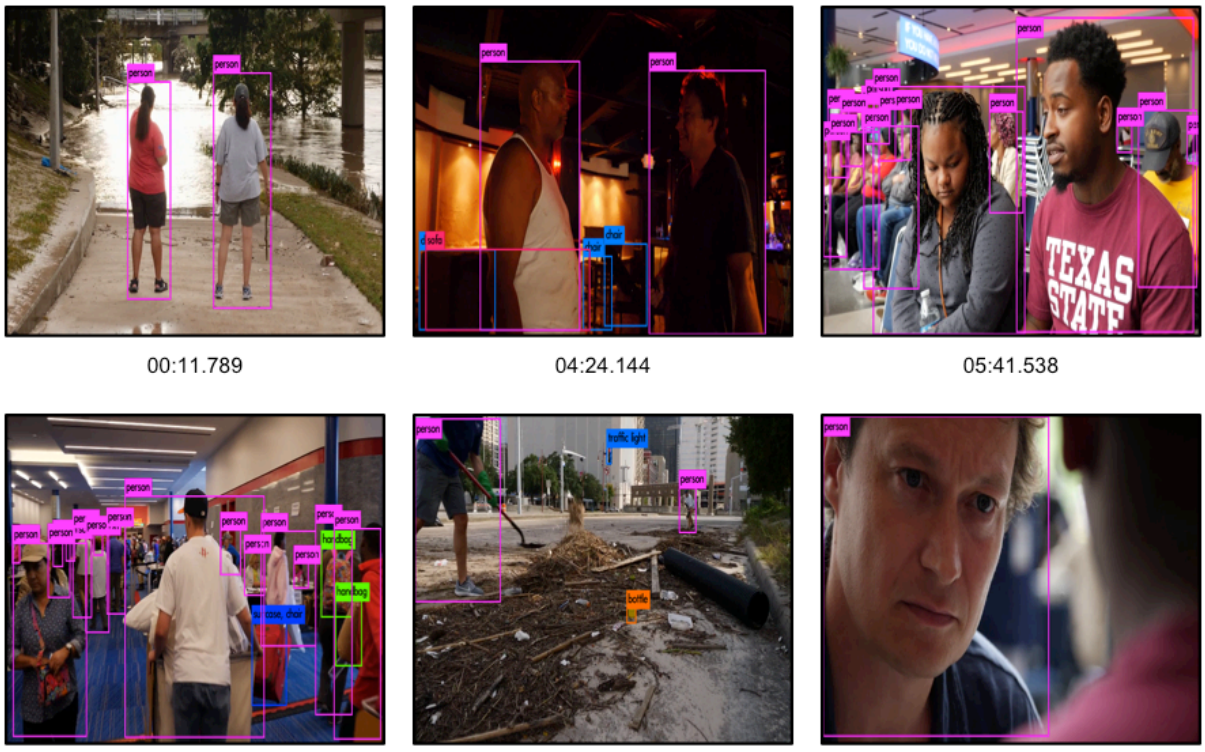
Example Output; Feature #1 Dominant colors from each frame. Examples of colors from 5 frames:



Example Output; Feature #2 Shot and scene breaks. Representative frame from the first 12 shots are given here:



Example Output; Feature #3 Object and face detection. Examples from 6 frames:



Example Output; Feature #4 Face tracking and identification. Examples of two tracked faces are shown here:

Example Output; Feature #5 Detect and classify diegetic and non-diegetic sound. The first two minutes of sounds are captured in this table:

time start	time stop	diegetic?	description
00:00.400	00:18.960	no	background music (guitar)
00:17.000	00:30.360	yes	car noises
00:37.360	00:37.880	yes	car horn
00:31.120	00:58.200	yes	water noises
01:10.160	01:15.160	yes	heavy wind
01:22.240	01:30.560	yes	car noises
01:38.200	02:11.080	no	background music (ambient)
02:11.760	02:12.800	yes	scraping noise
02:13.600	02:13.840	yes	scraping noise
⋮	⋮	⋮	⋮

Example Output; Feature #6 Automatically produce textual transcript by speech to text algorithms. Extracted transcript from the first 26 seconds of the example:

time start	time stop	speaker	text
00:03.400	00:10.560	narrator	the flood waters vanished as fast as they'd come in the heart of America's fourth largest city
00:10.600	00:13.200	narrator	they are just beginning to contemplate the aftermath of an unprecedented storm
00:18.240	00:26.360	Gabriel Gatehouse	well this sand here tells you that the water was flooding right over this busy intersection here in downtown Houston
⋮	⋮	⋮	⋮

Detailed Workplan The figure below extrapolates on the workplan description in the project narrative. Each task is associated with a particular subset of team members and assigned to a contiguous time window.

Agenda for Core Team Meeting I: 04-05 January 2019, University of Richmond

Friday, 04 January 2019 (Bell and Williams arrive prior evening):

08:30 - 08:45	Pick up Bell and Williams at Linden Row and drive to campus
08:45 - 09:00	Coffee and visit logistics (wifi, reimbursement, ect)
09:00 - 10:15	Evaluate the proposed metadata schema for each series, episode, character, and scene. Describe the ideal visualizations to present these data elements and identify metadata (or specific formats) required to serve these visualizations.
12:00 - 13:15	Lunch on campus
10:30 - 12:00	Meeting and demo with student assistants working on coding the training sets for the project.
13:30 - 15:30	Prior to the meeting, everyone will be assigned two public digital tools to evaluate. During this time, each of the 5 core team members will show their examples and give critical and constructive feedback on both (1) the public website and (2) any related written scholarship. A list of lessons learned from these will be accumulated.
16:00 - 17:15	Evaluate the proposed flow of the toolkit. Write a list of the types of users and discuss how each one would navigate the site. Collect explicit feedback about what type of organizational cues (visual, symbolic, textual) would be most appropriate for every action.
17:30 - 19:00	Dinner (downtown Richmond) and informal discussion

Saturday, 05 January 2018:

08:30 - 08:45	Pick up Bell and Williams at Linden Row and drive to campus
08:45 - 11:30	Discussion of visual design of the graphic output of DVT. Nathaniel Ayers will provide a few proposals to ground the discussion on particular options.
11:30 - 12:30	Lunch on campus
12:30 - 15:00	Plan meeting with advisory board. Create an agenda of the video meetings with each member, with particular focus on the questions we would like to ask. Required slides and other materials for the video meetings will be outlined and delegated to a particular team member.
15:15 - 16:30	Run through tasks and owners of all Phase II items. For each task, explain the process and desired outcome to other members of the team. If one part depends on previous actions being completed, make these explicit.
16:45 - 17:15	Verify that all notes created from the meeting are posted and available on the internal team GitHub wiki pages. Make particular notes for new action items that came out of the meeting.
17:30	Bell and Williams depart

Agenda for Case Studies Team Meeting: 18-19 October 2019, University of Richmond

Friday, 18 October 2019 (Berke, Calhoun, Oyallon-Koloski, and Vukoder arrive prior evening):

08:30 - 08:45	Pick up Case Studies Team at Linden Row and drive to campus
08:45 - 09:00	Coffee and visit logistics (wifi, reimbursement, ect.)
09:00 - 10:30	Team introductions. Each member of the case study will have twenty minutes to give an informal introduction to their work and the questions that drive their research.
10:45 - 12:00	Arnold will guide a demo of the current toolkit and demonstrate ideas for how the toolkit can be incorporated into research.
12:00 - 13:00	Lunch on campus
13:00 - 14:45	Discussion of avenues for integrating the toolkit into existing and developing lines of research for each participant.
15:00 - 17:00	Workshop-style deep dive into the inner workings of the toolkit and how it was built. Attempt to explain the technical aspects as clearly as possible to a non-specialists.
17:30 - 19:00	Dinner (downtown Richmond) and informal discussion

Saturday, 19 October 2019:

09:00 - 11:30	Ayers presents a prototype of the DVT frontend tool. Whiteboard ideas for tutorials, design decisions, and points of clarity in the user documentation.
11:45 - 13:00	Lunch
13:15 - 15:00	Discuss venues for publishing humanities-focused articles using the toolkit will be explored. Where possible, deadlines for following up on these will be set.
15:15 - 16:45	Construct timeline and deadlines for the final phase of the project, in which the Case Studies Team will test and write-up short reports about the tool's application to their corpus of interest.
17:00 - 17:30	Verify that all notes created from the meeting are posted and available on the internal team wiki pages. Make particular notes for new action items that came out of the meeting.
17:45	Case Studies Team departs

Agenda for Core Team Meeting II 08-09 February 2019, University of Richmond

Friday, 07 February 2020 (Bell and Williams arrive prior evening):

08:30 - 08:45	Pick up Bell and Williams at Linden Row and drive to campus
08:45 - 09:00	Coffee and visit logistics (wifi, reimbursement, ect.)
09:00 - 10:00	General discussion of project progress. Formally check-off actionable items and list any open issues or roadblocks.
10:15 - 12:00	Workshop-style demo of the completed and documented DVT library will be facilitated by Arnold. Other members of the Digital Scholarship Lab will also attend. Any difficulties or interesting points will be documented; a version of the workshop will be posted to the website and code repository.
12:00 - 13:00	Lunch on campus
13:00 - 14:30	Discussion of avenues for promoting the project. During this time period we will actively draft e-mails to send to professional contacts and organizations to promote our project. These will be ready to send off as soon as the toolkit is launched in April 2020.
14:45 - 17:00	Building off of talks given at DH2019, outline humanities white paper as a group. Determine the best process for moving forward with completing the paper during the final phase of the grant.
17:30 - 19:00	Dinner (downtown Richmond) and informal discussion

Saturday, 08 February 2020:

09:00 - 11:30	Plan meeting with advisory board. Create an outline of the video meetings with each member, with particular focus on the questions we would like to ask. Required slides and other materials should be outlined and delegated to a particular team member.
11:45 - 13:00	Lunch
13:15 - 15:00	Prior to this meeting, other sources of funding for the next phase of the project will be identified. During this time period the team will outline the required documents for these funding opportunities. Each item in the grant proposals will be assigned to a team member with a tentative due-date for the draft completion.
15:15 - 16:45	Show the prototype library, which should a this point have a fully functioning graphical front-end. We will discuss and document any final modifications that need to be made prior to the beta-testing in the next and final phase of the grant period.
17:00 - 17:30	Verify that all notes created from the meeting are posted and available on the internal team wiki pages. Make particular notes for new action items that came out of the meeting.
17:45	Bell and Williams depart

Bibliography

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., & Kudlur, M. (2016, November). TensorFlow: A System for Large-Scale Machine Learning. In OSDI (Vol. 16, pp. 265-283).

Acland, Charles R. and Eric Hoyt, Editors. *The Arclight Guidebook to Media History and the Digital Humanities.* REFRAME Books, 2016.

Ajmera, Jitendra, Iain McCowan, and Hervé Bourlard. "Speech/music segmentation using entropy and dynamism features in a HMM classification framework." Speech communication 40.3 (2003): 351-363.

Amos, Brandon, Bartosz Ludwiczuk and Mahadev Satyanarayanan. "OpenFace: A general-purpose face recognition library with mobile applications." 2016.

Baraldi, Lorenzo, Costantino Grana, and Rita Cucchiara. "Measuring Scene Detection Performance." Iberian Conference on Pattern Recognition and Image Analysis. Springer International Publishing, 2015.

Baughman, James L. "Television Comes To America, 1947-1957." Illinois History 46.3, 1993.

Baughman, James L. *The Republic of Mass Culture: Journalism, Filmmaking, and Broadcasting in America Since 1941.* JHU Press, 2006.

Braudy, Leo. *The world in a frame: What we see in films.* University of Chicago Press, 2002.

Buckland, W. "What Does the Statistical Style Analysis of Film Involve? A Review of Moving into Pictures. More on Film History, Style, and Analysis." *Literary and Linguistic Computing*, 23(2): 219-30.

Burghardt, M., Kao, M., Wolff, C. "Beyond Shot Lengths – Using Language Data and Color Information as Additional Parameters for Quantitative Movie Analysis." In: *Digital Humanities 2016: Conference Abstracts.* Jagiellonian University & Pedagogical University, Kraków, pp. 753-755.

Butler, Jeremy G. "Toward a Theory of Cinematic Style: The Remake." Morrisville, NC: Lulu, 2003.

Butler, Jeremy G. *Television: Critical Methods and Applications*, 4th Edition. New York: Routledge, 2012.

Caldwell, John T. *Televisuality: Style, Crisis, and Authority in American Television.* Rutgers University Press, 1995.

Casey, Michael and Mark Williams. *ACTION: Audio-visual Cinematic Toolkit for Interaction, Organization, and Navigation.* 2011.

Cervone, Alessandra, et al. "Towards Automatic Detection of Reported Speech in Dialogue Using Prosodic Cues." Sixteenth Annual Conference of the International Speech Communication Association. 2015.

Cosentino, S., et al. "Automatic Discrimination of Laughter Using Distributed sEMG." Affective Computing and Intelligent Interaction (ACII), 2015 International Conference on. IEEE, 2015.

Dow, Bonnie J. *Prime-Time Feminism: Television, Media Culture, and the Women's Movement Since 1970.* University of Pennsylvania Press, 1996.

Ewerth, R., Mühling, M., Stadelmann, T., Gllavata, J., Grauer, M. and Freisleben, B. *Videana: A Software Toolkit for Scientific Film Studies.* In: Ross, M., Grauer, M. and Freisleben, B. (eds.), *Digital Tools in Media*

Studies –- Analysis and Research. An Overview. Bielefeld: tanscript Verlag, pp. 100-16.

Fiske, John. *Reading Television.* Routledge, 1978.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). "Deep residual learning for image recognition." In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).

Held, David, Sebastian Thrun, and Silvio Savarese. "Learning to track at 100 fps with deep regression networks." European Conference on Computer Vision. Springer International Publishing, 2016.

Horwitz, Jonah. "Visual Style in the 'Golden Age' Anthology Drama: The Case of CBS." Cinémas: Journal of Film Studies 23, no. 2–3 (2013): 39–68.

Jockers, Matthew L. *Macroanalysis: Digital methods and literary history.* University of Illinois Press, 2013.

Kar, T., and P. Kanungo. "A Texture Based Method for Scene Change Detection." 2015 IEEE Power, Communication and Information Technology Conference (PCITC). IEEE, 2015.

King, Davis E. "Dlib-ml: A Machine Learning Toolkit". *Journal of Machine Learning Research* 10, pp. 1755-1758, 2009

Kumar, Rupesh, Sumana Gupta, and K. S. Venkatesh. "Cut Scene Change Detection Using Spatio-temporal Video Frames." 2015 Third International Conference on Image Information Processing (ICIIP). IEEE, 2015.

Pulver, Andrew, Ming-Ching Chang, and Siwei Lyu. "Shot Segmentation and Grouping for PTZ Camera Videos." 10th Annual Symposium on Information Assurance (ASIA 2015). 2015.

Lamere, P., Kwok, P., Walker, W., Gouvêa, E. B., Singh, R., Raj, B., & Wolf, P. (2003, September). Design of the CMU sphinx-4 decoder. In INTERSPEECH.

Manovich, Lev. *The Language of New Media.* MIT press, 2001.

Mittell, Jason. *Complex TV: The Poetics of Contemporary Television Storytelling.* NYU Press, 2015.

Moore, Barbara, Marvin R. Bensman, and Jim Van Dyke. *Prime-Time Television: A Concise History.* Greenwood Publishing Group, 2006.

Morley, David. *Television, Audiences and Cultural Studies.* Routledge, 2003.

Moretti, Franco. *Distant Reading.* Verso Books, 2013.

Mulvey, Laura. "Visual pleasure and narrative cinema". Screen. Oxford Journals. 16 (3): 6–18, 1975.

Nadimi, Sohail, and Bir Bhanu. "Physical Models for Moving Shadow and Object Detection in Video." IEEE transactions on pattern analysis and machine intelligence 26.8 (2004): 1079-1087.

Nielsen. "The Total Audience Report: Q1 2016.", 2016. Retrieved from: http://www.nielsen.com/us/en/insights/reports/2016/the-total-audience-report-q1-2016.html.

Newman, Michael Z. *Video Revolutions: On the History of a Medium.* Columbia University Press, 2014.

Posner, Miriam "Digital Humanities and Film and Media Studies: Staging an Encounter." Workshop della

Society for Cinema and Media Studies Annual Conference, Chicago. Vol. 8. 2013.

Redmon, Joseph, and Ali Farhadi. "YOLO9000: better, faster, stronger." CVPR 2017.

Sanders, Jason, Gabriel Taubman, and John J. Lee. "Background Audio Identification for Speech Disambiguation." U.S. Patent No. 9,123,338. 1 Sep. 2015.

Salt, Barry. "Statistical Style Analysis of Motion Pictures." Film Quarterly 28.1 (1974): 13-22.

Silverstone, Roger. *Television and Everyday Life.* Routledge, 1994.

Spangler, Lynn C. *Television Women from Lucy to Friends: Fifty Years of Sitcoms and Feminism.* Greenwood Publishing Group, 2003.

Spigel, Lynn. *Make Room for TV: Television and the Family Ideal in Postwar America.* University of Chicago Press, 1992.

Sun, Yi, Ding Liang, Xiaogang Wang, and Xiaoou Tang "DeepID3: Face Recognition with Very Deep Neural Networks." arXiv preprint arXiv:1502.00873 (2015).

Szegedy, Christian, et al. "Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning." AAAI. 2017.

Thomson, Kristin. Report of the Ad Hoc Committee of the Society For Cinema Studies, "Fair Usage Publication of Film Stills". 2010.

Tomar, Suramya. "Converting video formats with FFmpeg." Linux Journal 2006.146 (2006): 10.

Tsivian, Yuri, and Gunars Civjans. "Cinemetrics: Movie Measurement and Study Tool Database." (2011).

Xu, Peng, Lexing Xie, and Shih-Fu Chang. "Algorithms And System For Segmentation and Structure Analysis In Soccer Video." ICME. Vol. 1. 2001.

Zagoruyko, S., Lerer, A., Lin, T. Y., Pinheiro, P. O., Gross, S., Chintala, S., & Dollár, P. (2016). A multipath network for object detection. arXiv preprint arXiv:1604.02135.

Zhang, Richard, Phillip Isola, and Alexei A Efros. "Colorful Image Colorization". European Conference on Computer Vision (ECCV), 2016.

Zhou, Bolei, Aditya Khosla, Agata Lapedriza, Aude Oliva, and Antonio Torralba. "Learning Deep Features for Discriminative Localization." Computer Vision and Pattern Recognition (CVPR), 2016.

9. Data Management Plan

Expected Data

The project expects to produce six types of data: (a) software source code, (b) website source code, (c) reports and articles, (d) metadata from case studies, and (e) closed captioning data and (f) still images extracted from television episodes.

Data elements (a) through (d) will be saved as git repositories and published on GitHub. All of these data will be publicly available as they are built by the team. The commit history in the git repository provide a complete summary of the process used to construct all of the data elements. Also, the code for creating the extracted metadata (d) is included in the examples directory of the source code in (a).

Data elements (e) and (f) are being stored using cloud storage provided by Box under an institutional license from the University of Richmond. These elements are under copyright protection and we will therefore not distribute them outside of our core team. If other researchers require access to this material, there is an easy approach. We provide reproducible code in (a) for extracting the copyrighted material from commercially available media formats (typically DVDs). External groups can then recreate (e) and (f) after acquiring a copy of the raw source material.

Period of Data Retention

All of the data that will be made public, elements (a) through (d), will be accessible on GitHub as it is created. There will be no embargoing of the data and no need for a separate process of publishing the data.

At the end of the grant phase, a machine readable version of elements (a) through (d) will also be placed on the University of Richmond Scholarship Repository. The UR Scholarship Repository is a service of the University Libraries at the University of Richmond. Published materials and data included in the digital repository reflect the research and scholarly work of the university community and are openly available to the general public for download and use. The University Libraries provide this service to University of Richmond faculty, staff, and students free of charge and are committed to providing perpetual access to deposited content.

Data Formats and Dissemination

The source code (a) and (b) will be written in code readable by any plain text editor. The DVT Toolkit will utilize Docker for ease of use, but could also be run without Docker by manually installing all software dependencies. The research reports and articles (c) will be written in markdown; this is also readable by any plain text editor and can be converted into a large number of other formats using the open source software *pandoc*. Markdown is also seamlessly converted to HTML by the GitHub platform.

Extracted metadata (d) is saved as plain text files using a comma separated value scheme. This format can be read and parsed by any programming language and most statistical and visualization software. The closed captions (e) are also stored in plain text files. Still images (f) are stored in the JPEG format, which is readily readable and convertible by all major web browsers, image programs, and email clients. We anticipate no need to convert these to another format inside of the next 10 years. All text files will be saved using the UTF-8 encoding.

Data Storage and Preservation of Access

Elements (a) through (d) will be stored on GitHub, with a duplicate copy of the entire repository stored on our internal Box account and on GitLab. The copyrighted material will continue to be hosted on our institutional account with Box. Migration to a new cloud platform in the long-term, if necessary, will be managed with the help of the University of Richmond's IT department. A machine readable version of elements (a) through (d) will be placed on the University of Richmond Scholarship Repository for long-term storage.