RMIT University

Rachel Wilson

The ASPERA digital online archive/repository: key issues facing the contemporary moving image archive

Abstract:

I have been conducting research into the feasibility of (and developing the schema for) the establishment of an online digital archive/repository designed to house the screen production and research outputs of the staff and students of Australian Screen Production Education and Research Association (ASPERA) member institutions. In conducting this research a number of significant issues facing the archive sector as a whole, and moving image archives in particular, have been identified. As all information agencies begin to grapple with the logistical issues emerging from the impending 'digital deluge' the very nature of the role that the moving image archive plays within the contemporary audio-visual landscape is under increasing scrutiny. The sheer volume of data produced with each 'born digital' audio-visual title means the issues of collection policy and curatorship, search-ability and accessibility together with preservation strategy become even more imperative. This paper articulates the rationale behind the project and identifies a number of the historical, theoretical, philosophical perspectives (and challenges) underpinning the notion of the contemporary moving image archive with particular reference as to how these bear influence on the establishment of the ASPERA online digital archive/repository.

Biographical note:

Rachel Wilson works as a lecturer of media production in the School of Media & Communication at RMIT University in Melbourne. Rachel is currently undertaking a PhD by project to develop the schema required for the establishment of a national digital online repository/archive for the staff and student screen production outcomes. Rachel is also currently the national president of ASPERA.

Keywords:

Moving image archives – Curatorship – Access – Preservation – Digital deluge

Introduction

> Motion picture film forms an indispensable part of our cultural heritage and a unique record of your history and our daily lives. Film archives, both public and private, are the organizations responsible for acquiring, safeguarding, documenting and making films available to current and future generations for study and pleasure. (FIAF, 2008)

The need for a national ASPERA online digital archive/repository of staff and student screen production research has emerged as a result of a number of coalescing forces. Most significantly is the fact the that moving image history is now widely understood and accepted a significant instrument in the documentation of the history and culture of our society and as an important artistic expressions in it's own right (Gracy 2008a: 1). With the emergence in the last decade of 'Orphan Film' studies there is an increasing appreciation of how student films form an important (yet largely neglected) genre within our moving image and cultural heritage (see Cherchi Usai, Strieble, Prelinger). Alongside this specific work relating to particular content in (or indeed absent from) the official archive is the emergence of 'the archive' as a focus of research within the humanities and social science in more general terms. The release of Jacques Derrida's *Archive Fever: A Freudian Impression* in 1995 helped focus a renewal in analysis on the way in which archives act as cultural and historical gate-keepers and challenged us to begin to identify and examine gaps within the constructed histories embedded within official archives. Student films, often marginalised as merely tools for training, are interesting historical documents in their own right and yet they remain almost completely absent from the moving image historical record.

Another imperative informing the need for a national archive emerges as Australia embarks on a new system for recognizing, and subsequently funding, university generated research. With the implementation of the ERA (Excellence in Research Australia) access to the research generated in each discipline is of vital importance. Within the new system key assessment criteria for excellence pertain to both quality and too a lesser yet important extent, impact. Obviously (and especially) for research produced in the screen production disciplines it is imperative that this research first be 'seen' or 'viewed' for its impact to be accurately judged. The establishment of online archive will provide a vehicle (amongst many others) to further facilitate access to this material.

Furthermore the creation of an online digital archive/repository will better enable the tertiary screen production discipline's contribution to Australian cultural life to more be more readily accessed and recognized both with in national and international contexts. It will provide an avenue allowing us to highlight the variety of research the discipline conducts at a national level. The archive will also provide a focal point for cross-institutional collaboration for the discipline and it may be possible to use the archive/repository as a powerful tool in lobbying the federal government (and for that matter the broader industry) for greater access to scarce research and/or production funds. With the formalisation of collecting and archiving procedures, we will also be in a stronger position to develop and sustain meaningful strategic allegiances with the other key cultural and archiving organisations in our field such as the National Film

and Television Archive (NFSA) and the Australian Centre of the Moving Image (ACMI).

At the most pragmatic of levels it is my experience, through working in a number of tertiary screen production programs across a number of different institutions, that the majority of screen production research is stored in highly unsuitable and unstable environments. These include cupboards (quite often in lecturers' offices), under stairs or buried away in storage, and as a result are rarely accessed, viewed and commented on beyond the initial classroom screening. Each year the current cohort of students ask to see the previous students' work. This desire emerges from genuine curiosity through to a highly competitive desire to do better than the year before. It also has the potential to become a useful marketing tool for each participating institution and allow members to be better able to preserve their unique individual cultural heritage and identity whilst communicating and sharing their particular screen research outputs.

Overall the project is set to design a framework in which to preserve and make accessible the cultural heritage work of higher education staff and students producing screen content. By providing access to the material online, both within institutions and to the wider public, a number of pedagogical ambitions will also be realized. These include; providing access to new audiences, introducing and cementing the importance of preservation and archiving in relation to cultural heritage (and professional practice) and, provide cohesive links to and for the industry and community to the work undertaken within the discipline. The variety of material and content with which the staff and students engage provides a series of interesting insights into our cultural and institutional memory. For example what issues are considered important to eighteen or nineteen year olds today compared to 2020? Can a mode of practice influence the type of material produced from specific training institutes? What kinds of themes preoccupied our most recently celebrated film director when they were a student?

The model for the archive/repository is still, at this stage, a necessarily vague one. How it will eventually function will be largely dependant on the amount of funding the venture will attract. There is a proliferation of institutional repository software now readily available. Yet very few packages are able to focus on creating an environment that is both user friendly and is responsive into the changing agendas of screen research and how it might be made assessable and useful in the future. There are a number of fundamental affordances that the archive/repository will need to have in order to become, as envisioned, the primary space to house and preserve the Australian tertiary sector's staff and student work. The archive/repository will need to be web based and yet centralized enabling for robust preservation strategies, as required by any archive. It will live and run utilizing existing institutional infrastructure (including high speed fiber optic cables and large server farms) developed primarily for the high performance computing needs of the science disciplines. Each member institution will have a separate interface that will allow them to individually control who has access to the material and when (although the central aim is that it will be as open as possible), accessed via the central ASPERA

portal. Students will be the primary depositors and work can only be accessed for assessment via the archive—there in guaranteeing accurate deposit (if staff can't access it—it won't be assessed). The archive/repository will where possible operate using open access software mitigating some of the cost of future updates and sustainability. The schema for the cataloguing meta-data required for meaningful future access is still under investigation and will be dependant on user testing on how many fields it is feasible to expect students to accurately complete.

The hybrid space between archives and repositories

A major requirement in conducting research in this arena has been the need to become fluent in a range of jargon from different domains (library, archive, museum and repository). This is coupled with the development of a set of skills allowing more meaningful communication with computer scientists and network engineers. In attaining this fluency I have discovered that the traditional boundaries that previously existed between the different information and collection agencies are quickly disintegrating. Where archives once had little requirement to make material accessible beyond the selected few there has been a tremendous growth in the public's expectation that entire collections should be screened digitally via the internet—driven largely by the explosion of the information revolution. Libraries on the other hand need to create robust preservation strategies in order to deal with the many different (and constantly altering) forms of information they are required to make accessible on a daily basis. Museums are growing increasingly dependant on information managers in order to provide geographically disperse publics with better access their collections.

In the past the different roles of each information institution has reflected the way 'in which the information was created, used, valued, preserved and disposed of by individuals, organisations, and communities in the conduct of business, scholarship, learning, and personal affairs.' (Gilliland-Swetland 2000: 4). The table below (courtesy of Gilliland-Swetland) demonstrates overlapping and distinct duties in which each institution has traditionally engaged.

Libraries
- Identify, acquire, preserve, and provide access to the world's published knowledge
- Promote equity of access to information
- Promote intellectual freedom
- Support education and continuous learning and research
- Support the development of information literacy in society
- Serve as focal points for communities and promote community interests
Museums
- Identify, acquire, preserve, and exhibit unique, collectible, or representative objects

- Promote cultural, community, and familial identity and understanding
- Provide experiences where visitors can make connections between content and ideas
- Serve as memory institutions for a culture
- Support formal and informal learning and research
- Serve as focal points for communities and promote community
Archives   - Identify, appraise, preserve, and make available documentary materials of long-term value (essential evidence) to the organization or public that the archives serves   - Ensure the accountability of government by preserving public records and making them available to the citizenry as is legally and ethically appropriate   - Ensure the accountability of nongovernmental institutions to their shareholders, boards, and other constituents   - Preserve unique or collectible documents   - Serve as memory institutions for a culture   - Support scholarly, administrative, and personal research

Table 1. From Gilliland-Swetland 2000: 5

Since this table was published in 2000, institutional repositories have also emerged as an important player in the management of informational data—particularly as it relates to access to peer reviewed material. As a result of this growth, particularly within universities, much of the work in conceptualising and rationalising the ASPERA archive/repository has borrowed heavily from the rhetoric of those involved in the international push for Open Access digital repositories. There has been growing and substantial pressure globally (and in particular in Europe and North America) to provide ready and easy access to publicly funded research. This drive primarily stems from the need to get time sensitive scientific and medical research quickly into the public arena and how important this issue is considered was illustrated by 'The Federal Research Public Access Act of 2006' (FRPAA) bill which was presented to the US Congress in 2006. Australia's responses to these international currents was reflected in the May 2004 launch by John Howard of 'Quality and Accessibility Frameworks for Publicly Funded Research' that is part of the *Backing Australia's Ability—Building our Future through Science and Innovation* policy statement. These sentiments have continued to garner government support, even after a change of government, and whilst there is yet to be an adoption of a policy of the mandatory submission of publicly funded research to a national repository system this government continues to invest heavily in organisations set up to specifically to facilitate such an eventuality—such as the Australian National Data Service (ANDS).

The above table has proven particularly useful in order to understand the impact of the disintegration of boundaries between various institutions and how best to approach the organizations that the ASPERA archive/repository will be required to negotiate

and work with. For example as a museum ACMI has different access policies to that of the NSFA whose primary mandate has traditionally been to collect and preserve.

It is easy to naively enter into this research arena with idealistic notions that advances in computing, digital storage and the web now equalled the seamless and easily facilitated integration and interoperability of these types of collection agencies and all that is needed is to select an off-the-shelf application and begin to populate it with content. However, it is of course far more complex than that. Encouragingly there is a growing movement within these organisations in working towards a set of shared ideals across seemingly diverse projects—a genuine desire to provide meaningful access, where possible, to as much information, to the majority of people.

The table has also proven also useful in helping to identify where the ASPERA project might eventually fit within the 'spectrum' of agencies. At this stage the work positions the project as a hybrid between a traditional moving image archive and an institutional repository. The project will need to: both collect and preserve, be entirely digital, be open source and provide access for 'many' to use, be cost efficient, be sustainable and be easily connected to other screen archives and research collections.

Clearly each of these aspirations has embedded in them a complex set of histories, politics and technical issues influencing their feasibility. Whilst it is essential to be mindful of how we are quickly reaching a point in history where the technical, physical and financial constraints in relation to the housing of vast amounts of digital data has become much more feasible—what is emerging as the central issue is—how can we as information managers provide access and meaningful entry points into the vast and expediently growing sets of data? What policies need to developed in order for users to understand and know how to search the databases for the material they need, and how might the archivist best manage the massive volume of data each distinct 'object' generates?

The digital deluge and the challenges facing the contemporary moving image archive

> It is estimated that about one and half billion viewing hours of moving images were produced in the year 01999, twice the number made just a decade before. If that rate of growth continues, three billion viewing hours of moving images will be made in 02006, and six billion in 02011. By the year 02025 there will be some one hundred billion hours of these images to be seen. (Cherchi Usai 2001: 111)

The volume of material (data) that can be conceivably preserved and subsequently accessed is at the heart of debates surrounding the role of the contemporary moving image archive (not to mention its day to day running). Like any good debate, there is no consensus on how best to proceed. Preservation to the traditionally rigorous standards and providing meaningful access to the volume of data referred to above is almost incomprehensible. When (and if) archives are able to adhere to these standards the process of preservation and cataloging is extremely expensive and resource intensive. Subsequently the ratios of what is created as opposed is what is preserved is diminishing as time goes on. This is especially valid in non-western and developing

nations. Whilst digital technology offers 'the seductive promise of a real miracle: perfect vision, eternal moving images that can be reproduced as infinitum with no loss of visual information' (Cherchi Usai 2001: 113) there is a growing understanding that this is an unattainable dream. The archiving sector is then left to ponder if 'so many of those images come and go without our even hearing about them, then how in the world can we form an idea of what cultural heritage is?' (Cherchi Usai 2001: 111-112).

Overall the sector appears to suffering from a sense of confusion and a lack of direction regarding the role moving image archives must now serve. My research suggests that the significant divide (which has always existed within the sector) between access and preservation has begun to shift. Understandably, preservation has been the central driver for much of the past century but the rising costs associated with the process (in the face of such insurmountable odds), together with a shifting paradigm emerging from the advent of low cost digitization, has led many key players to suggest the sector needs to reconceptualize issues from new points of view.

> do we perceive ourselves as archives, as museums, or as libraries? How much has this blurring…[of] functions—preservation, presentation, and making accessible—been with the community from it's beginnings? Or is it only an effect Digital has on our self-image? (Cherchi Usai et al. 2009: 52)

It is becoming accepted wisdom that 'moving image stewardship may no longer be the exclusive province of institutions such as archives and libraries, and may soon be accomplished in part through the work of other interested individuals and organizations as they contribute to and define collections' (Gracy 2007: abstract). As Karen Gracy points out in her 2007 article 'Moving Image Preservation and Cultural Capital' the technologies being conceived and built within the current Internet environment have the potential to create a new paradigm where 'the curatorial or archival authority with which cultural heritage institutions are invested may diminish to the point where society may question the need for such entities to perform such work' (Gracy 2007: para. 9).

Traditional archives are also under significant pressure to justify the nature of their collections and curatorial policy. As Annemaree Lloyd illustrates in her paper 'Guarding Against Collective Amnesia? Making Significance Problematic: An Exploration of Issues', the material that is collected for the official record tells only one story. The material 'not collected' or 'excluded' tells another altogether different one. Although 'official' archives generally attempt to claim objectivity it is now widely accepted that the designation of any item as documentary heritage is a significant political act (Lloyd 2007: para.10). Many memory institutions are by their very structure and nature subject to political coercion and influence. According to Lloyd these influences are often downplayed and rarely debated within the sector. What is considered significant enough to preserve, order and structure is 'underpinned by notions of truth held by the powerful in society and by the decisions of the powerful about which truth, or which versions of truth, are valid and worth of preserving for the long term' (Lloyd 2007: para.15).

> There's an illusion being created that all the world's knowledge is on the web … material that is not digitized risks being neglected as it would not have been in the past, virtually lost to the great majority of potential users. (Erway and Schaffner 2007: 1)

In a presentation made to the Digital Library Federation in Philadelphia in 2007 well known and influential North American archivist, Rick Prelinger, pointed out that 'what we think constitutes an archive is being displaced by vernacular definitions of what archives are (I'm thinking YouTube)' (Prelinger 2007: 3). Prelinger argues that YouTube has quickly become the world's default media archive and in doing so it, and other similar online applications, have significantly heightened public expectations in terms of access, whilst correspondingly lowering expectations around quality and, he goes on to argue, its difficult to imagine how any publicly funded institution can ever hope to equal them. Traditional archives are besieged by what he calls 'moving image exceptionalism' (e.g. massive copyright problems, multiple and fragile formats, expensive, access contingent on high cataloging and preservation standards, excessive deference paid to powerful rights holders and donors, and the infestation of Hollywood user rights models into the archive world) and posits a radical set of principals that separate issues of preservation, cataloging and description from those of access.

Influenced by the work of Ricky Erway and Jennifer Schaffer in their report *Shifting Gears: Gearing Up to Get Into the Flow* Prelinger suggests archives should, where possible, 'expose unpreserved and uncatalogued collections online, ask for forgiveness rather than permission and default to access rather than enclosure (and ritual)' (Prelinger 2007: 37). He also argues that online visibility has the potential to drive preservation agendas within archives. If archives were to become more user focused they would be in a better position to rally users as allies particularly in advocating for better resources for preservation. Prelinger argues that, with the advent of easily accessible content from multiple web sources, archives need to shift from the traditional position of accumulating media forms that were easy to collect and seemed important at the time to the opposite and 'collect materials that are most challenging to source, to store, and to serve' (Prelinger 2007: 55). Furthermore he wonders why archives and libraries have for so long internalized and mirrored the constraints and hierarchies of the world around us and suggests, albeit idealistically, it is time for them to become 'a force for clarity and mindfulness when it comes to deciding how culture is made visible and distributed' (Prelinger 2007: 47). What is required is a decidedly more open and inclusive notion of access. Below is his useful chart depicting a spectrum of access from the least open archive practice to the most open.

Access is a spectrum, openness a practice

[LEAST OPEN]

content archived and unavailable
content available for in-house research access with permission
content available for in-house research access without permission
content circulates within controlled world (e.g., print loans)
content licensable by special arrangement
low-resolution streaming video copies available online
high-resolution streaming video copies available online
content not interoperable with content in other collections
quotation allowed without reproduction
quotation allowed with reproduction (e.g., frame grabs)
reference-quality, time-coded, Elmo, defaced study copies furnished
broadcast-quality or projectable copies furnished
user-controlled copying in-house
low-resolution downloadable copies available online
high-resolution downloadable copies available online
content licensable by simple arrangement
content interoperable with other collections
content freely available for reuse with restrictions (e.g., noncommercial use)
content freely available for sharing online
content freely indexable, crawlable, navigable by noncommercial web services
content freely indexable, crawlable, navigable by anyone
content freely available for any reuse without restriction
component parts of works (e.g., shots, segments, audio tracks, edit lists) available freely online without restriction

[MOST OPEN]

(Prelinger 2007: 49)

In his 2009 book *Film Curatorship: Archives, Museums, and the Digital Marketplace* leading archivist and scholar Paolo Cherchi Usai suggests that for traditional archives to survive there needs to be an ongoing dialogue and conversation between institutions and their user groups. A new, more contemporary notion of the role of the archive curator needs to be developed and accepted and that development of a distinct selection policy for each archive is the key to contemporary archival practice. At the 7th Orphans Film Symposium in New York City April Cherchi Usai asserted that film historians and scholars need to at the heart of archival curatorial selection policy—particularly in relation to the selection of which titles ought to be slated for preservation.

Another key challenge for all archives and repositories alluded to above is that of searchability. The capacity to locate and find items and objects of use within the vast amounts of content now available is becoming increasingly difficult. Much time, energy and resources are devoted to the establishment and maintenance of rigorous cataloging standards within traditional archives. As a result there are often massive backlogs as teams of highly trained professionals grapple with the volume of material submitted whilst establishing new standards for the emerging forms of digital media production. With the growing acceptance of 'folksomies' and the ability of 'users' to create meaningful connections to, and through work, there appears to be an opening for the inclusion of more 'everyday' forms of cataloging. As database theory begins to significantly influence the way cultural institutions do business, the role of the traditional relational cataloging database as the *only* way to catalogue material is being severely tested. Momentum is growing throughout the cultural collection community to provide access not only to their collections but to their catalogues and thereby allow (some) users the ability to add to the searchable metadata. A growing

demand for access to what has traditionally been seen and conceptualized as the 'backend' of collections is an interesting development. The sector will need to proceed with care in this approach but its potential to add value to collections for their users is only now just beginning to be appreciated.

A tangent to these debates surrounding searchability is the fast growing computational ability of machines to read video or digital moving images. My research has lead me to some interesting niches within the computer sciences where a number of international teams are working feverishly on (largely military) applications to develop fast paced content-based detection software. The ability of machines to conduct high quality searches based on both feature (facial recognition, text extraction) and event detection criteria is growing. They have already been testing these systems within moving image archives and are keen to share the software with institutions charged with the job of collecting and maintaining large moving image collections such as the BBC, ABC (North America), CNN, C-Span and the Netherlands Sound and Vision Archive (Oven et al. 2008: 4).

Implications for the ASPERA online archive/repository

One of the key aims of my research into the history and philosophical underpinnings of moving image archives has been to develop an ASPERA archive/repository in such away that it is highly compatible with the Australian National Film and Sound Archive. Initially I found discussions surrounding standards very intimidating. I worried that the vision of the archive/repository enabling students to catalogue and submit their own films was naïve and idealistic (even if it appears the most sustainable way to maintain the archive). Cataloguers at the NSFA are extensively trained and use highly specialised software—but I have discovered that there is a great deal of debate and change surrounding these issues within the collection community. Archives, museums, galleries and libraries are all in the process of developing strategies to manage the heightened expectations for access from the public, with many undertaking large-scale reviews of their policies in this area. There is a palpable sense that the timing could not be better for a project of this nature. The gaps in existing collections are becoming increasingly obvious (e.g. why hasn't the NSFA ever had a policy to comprehensively collect student work?) and there is a significant push amongst all the collection agencies within Australia that we will be in unique position to harness the opportunities afforded to the nation with the imminent implementation of the National Broadband Network (NBN).

There are, of course, still many issues to left to negotiate (such as copyright) before the ASPERA online archive/repository can be fully realised. But it is now much easier to imagine positioning the project as one of the 'new and vernacular' archives as described above. The radical shifts in how we access and view content is contributing significantly to shifting ideas of permanence and, like many others, I believe that what will emerge as acceptable preservation standards in the future will come to reflect these changing paradigms. As Paolo Cherchi Usai writes the notion of the authentic experience of cinema has always been an illusion (Cherchi Usai 2001: 101) and it is the drive and hunger to see and reuse content, using whatever device is most current,

that is driving the push to develop new models for archives—not the desire to engage within the rarefied surrounds of darkened cinemas. I believe that a model that invites as many people as possible to the table as stakeholders and as creators of their own heritage will have the longest life and prove to be the most useful to the community in the long run. This is naturally the ultimate aim of every archivist and memory institution.

Works cited

Cherchi Usai, David Francis, Alexander Horwath, Michael Loebenstein (eds) 2008 *Film Curatorship Archive, Museums, and the Digital Marketplace,* Pordenone: SYNEMA Publikationen

Cherchi Usai, P 2001 *The Death of Cinema: History, cultural memory and the digital dark age* London: British Film Institute

Cherchi Usai, P 2009 'Are All (Analog) Films "Orphans"?: A pre-digital appraisal' *The Moving Image* 9 (1), 1-18

Derrida J 1996 *Archive Fever: A Freudian impression* trans. Eric Prenowitz, Chicago: University of Chicago Press (*trans*. of *Mal d'archive*, Paris: Galilée 1995)

Erway, R and Schaffner, J 2007 *Shifting Gears: Gearing up to get into the flow*, OCLC Programs and Research, at http://www.oclc.org/programs/publications/reports/2007-02.pdf (accessed 15 June 2010)

FIAF Open Forum 2008 'Don't Throw Film Away—The FIAF 70th anniversary manifesto' *Journal of Film Preservation* 77 (78), 5-6, at http://www.fiafnet.org/uk/members/Manifesto.cfm (accessed 18 May 2010)

Gilland-Swetland, AJ 2000 *Enduring Paradigm, New Opportunities: The value of the archival perspective in the digital environment*, Washington DC: Council on Library and Information Resources, at http://www.clir.org/pubs/reports/pub89/contents.html (accessed 18 May 2010)

Gracy, KF 2007a *Film Preservation: Competing definitions of value, use, and practice* Chicago: The Society of American Archivists

Gracy, KF 2007b 'Moving Image Preservation and Cultural Capital' *Library Trends* 56 (1), 183-98

Lloyd, A 2007 'Guarding Against Collective Amnesia? Making Significance Problematic: An exploration of issues' *Library Trends,* 56 (1), 53-66

Paul Over, George Awad, Jonathan Fiscus, Brian Antonishe 2009 'TREC Video Retrieval Evaluation' in *TREVCID 2009 Participant Notebook Papers and Slides*, at http://www-nlpir.nist.gov/projects/tvpubs/tv.pubs.org.html (accessed 4 June 2010)

Prelinger, R 2007 'It's only a Moving Image: Archives, access and the social contract' *Digital Library Federation: Proceedings of the Fall Forum,* Philadelphia, 5 November 2007, at http://www.diglib.org/forums/fall2007/2007fallprogram.htm (accessed 10 June 2010)

Streible, D 2007 'The Role of Orphan Films in the 21st Century Archive' *Cinema Journal* 46 (3), 124-28

Streible, D 2009 'The State of Orphan Films: Editor's introduction' *The Moving Image* 9 (1), vi-xix