

---

## Digital Archiving and its Neutrality: Questions on logic, longevity, and discovery in Knowledge Production

---

**Sujith K G**

**National Institute of Fashion Technology, Kannur**

### *Abstract*

The proposed paper critically engages with the paradigm shift in knowledge production brought in by the digital media. It focuses on the element of accidental discovery in knowledge production across disciplines, particularly the medium of still photography. The logic of digital data archiving and its limitations in terms of facilitating random search and finding, as against analogue media where data is stacked in a traditional and human friendly manner, or the 'natural' ordering of archived information in analogue will be scrutinized in the paper. Philosophical questions on the scope, future and utilitarian aspects of archived digital data/information alongside its scopes of retrieval for purposes that they are or were not deemed to be for, will be evaluated critically. The evolution of technology and fate of digital data produced and archived at various junctures of the evolution require critical and philosophical interrogations to see what they were intended to be and what they have become specifically due to digitalization. The new meaning of eternity in the light of the deemed ability of the digital data to stay intact for long or forever against the phenomenon of becoming obsolete too soon with the change of technology that nullifies or challenges the notion of eternity will also be looked in to. Thinking specifically about the medium of photography, random discovery of photographic images in various formal and informal archives and associated research have been a very prominent form of knowledge production. The paper would be an attempt to engage with the discrepancy between the huge corpus of digital images produced and most of it being pushed beyond the scope of retrieval. This poses many pertinent questions to the concept of digital archiving of photos, retrieval of which is highly determined by the personal and political choices of the archiver. The projected neutrality and immense potential of the digital medium thus becomes a farce and will open before us, as a medium fluctuated by infinite political conditioning and biases. The paper will investigate this aspect of the digital medium using the digital archiving of photographs as a field.

---

Accidental discovery has been a pivotal component of knowledge production and with the normalisation of 'digital' as the mode of production and archiving of knowledge or data, there has been a paradigm shift that reduces the chance of accidental discovery and puts up curated and

moderated knowledge on top of the available repository, making it easily found and accessed. Minimal or little accidental discovery or discovery of certain data by chance in the process of searching for something that was not very specifically intended for and the abundance of digitally produced and archived, produced in analogue and archived digitally knowledge, that are archived not based on its own characteristics or materiality, but on a description of it or metadata, makes the knowledge and archives highly curated and influenced. This kind of knowledge is highly predisposed to the biases of the archiver and its scope is limited by the social, political and technological preferences of the person who wishes to access that particular piece of knowledge.

There are assorted classes of knowledge that were produced in various modes before the digital – in terms of production, archiving and the way it was accessed. The analogue mode of production as well as archiving constituted the primary database in museums, both formal as well as informal. The kinds of knowledge that have been produced in analogue forms such as celluloid, photographic paper and paintings – their archival has partially or completely shifted to digital. However, it is not to say that the analogue photography is not practiced any more at all. It is still practiced in certain exceptional cases such as by certain ardent film photography enthusiasts and fine art photographers.

### **Phases of Digital Inception**

A structuring of different phases of digital data production and archiving under broad categories is attempted here based on certain signposts in the field. This categorisation follows a linear temporality based on the invention as well as obsolescence of different modes of data production/archiving:

#### ***Phase 1, the Pre-Digital or the Digital Foetus***

This was when the ‘digital’ was being conceived, both as a technology as well as an idea. For laymen, it was just occasional news and science fiction – film and literature. Analogue production, Analogue Post production, Analogue Archive had been the markers of the time. During the ‘Digital Foetus’ era, in Photography and Film making, the processes of production, post production, circulation and archiving were done in analogue in its entirety. The process of photography was very much chemical and not digital. The popular imagination and representation of photography was in the form of film rolls and negatives. Silver nitrate films were exposed

using analogue cameras, films chemically washed and developed in no negatives and positive impressions were printed on photographic papers. Even in the most creative of the imaginations of photography, a ‘digital’ future like the present we have today was never there.



*Analogue Film Archive*



*Analogue Photography Archive*

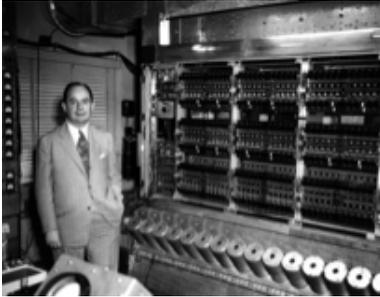
### ***Phase 2, the Digital Baby***

Analogue production, Analog Post Production, Analogue Archive Technologies such as Telecine, Nonlinear editing, reverse telecine, image scanners were invented, but in the very nascent phase and were seldom used. The newly invented technologies were not feasible technically and financially. The output was inferior to that of analogue counterparts and were mostly perceived as a not so viable, not so user friendly, and very unlikely to get popular technologies elsewhere. Most importantly, though people started hearing of machines called computers – yet the imagination that one day the world would revolve around it was beyond imagination. The knowledge and imagination that computers could be used in easing out the tasks was not very visual but numeric. Computers were machines that could help in doing mathematical calculations.

### ***Phase 3: The Digital Childhood***

Digital Technologies evolved and became known to people though were not quite accessible. Technologies of Digital Production (digital camera), Digital Post Production (computer based image processing), Analogue Archive (celluloid and magnetic tapes) got more popular, yet very expensive. Digital cameras and camcorders, nonlinear editing alongside arc lamp projection and printed images marks the ‘Digital Child hood’. Production remained mostly analogue, post production moved digital, archiving and circulation remained analogue mostly. CDS and DVDs got popular in the market over VHS tapes, but the logic of archiving stayed analogue. Computer based storages were mostly confirmed to the hard

drives and floppy disks. CDs and DVDs were popularly perceived as replacements of magnetic tapes and not storage devices.



*The first electronic computer built by the Institute for Advanced Study (IAS) in Princeton, New Jersey*



*One of the first-generation Apple computers from 1976*

#### **Phase 4: The Digital Adolescence**

Storage and memory moving online started here in to the ‘cloud’ in this era of Digital Production, Digital Post Production, Digital Archiving and digital circulation. Media forms moved out of CDs and DVDs with coming of internet. Another major move marked during the era was the upcoming of digital as a substitute of print publishing. Major publications around the world launched online versions along with the print anticipating a digital revolution. A few publications even stopped printing forecasting a paper free world. Platforms like amazon kindle was introduced attempting to replicate the experience of paper in digital.

#### **Phase 5: The Digital Adulthood**

The age that we live in where the paradigms are shifted in production, post production, archiving and circulation. It is the age of the ‘new’ where, most analogue modes of production, post production, archiving and circulation became obsolete. Most importantly, alongside most analogue technology, most of the initial digital technologies too met with the same fate. It is this era that marked a substantial departure in terms of the logic of archiving and retrieval of data, moving towards a more curated access from the natural orders of archival and storage as against the analogue.

#### **Phase 6: Digital ‘Old Age’**

This time we are unsure of future, only thing that is certain is a change that is inevitable. It is important to note that the progression from each age to the

next was mostly disruptive and forecasting the future on the basis of present was never possible. And the trend testifies it, the way things go throws the unpredictability and hence resulting uncertainty. There have been many films and books that imagined the future in lines of its contemporary times. One example is that of the *Blade Runner* (1982) the neo-noir science fiction film directed by Ridley Scott starring Harrison Ford. The 1982 film imagined a future world of 2019 where there is a conflict of survival between the humans and synthetic humans engineered to work in extra-terrestrial colonies of earth. Typically, every imagination or forecast of the future is based on the dominant social psyche of the present and in most cases, end up substantially different from the world that actually evolves. Another film that is worth citing in this context is *Black Panther* (2018), the American superhero movie directed by Ryan Coogler based on the Marvel Comics character Black Panther. The film imagines a technocratic utopia with unspecified temporality. Popular imaginations of utopias materialise here as both awe and fear. The utopias and dystopias and as imagined by Hollywood are reflexive of the popular imaginations of the future and they seldom live up to the reality.



### **Shift in materiality; Visuality to Digitality**

The most important cultural transformation brought in by the digital is the perceived materiality of a photograph. With the photography process going digital and the shift in materiality of the photograph, increased quantities of data are produced, circulated and archived. Instead of visuality, the most innate and independent characteristic of an image is subsided by certain descriptions of data that describes the image, something that is exclusively brought in by the digital mode of production and archiving – the metadata becomes more important than visuality. There arises excessive compulsion on tagging or description that becomes pivotal in archiving and wherever it is not sufficiently carried out, the images produced, though archived ‘safely’, remain inaccessible making it practically non-existent. Beginning from Walter Benjamin in his essay ‘The Work of Art in the Age of Mechanical Reproduction’ many cultural theorists have discussed shifts or changes in the materiality, resulting in boundless duplication.

It could be argued that the increased quantum of data is the gift of digital and without which the corpus would have been much smaller in comparison to what we have today. On the contrary, the counter question is where does the increased size of corpus leading to? What happens to the images produced? Particularly what is the purpose that they satisfy? Does the digital mode of production and archiving modify the purpose of the photograph from its capturing moment to that of its archiving and retrieval purposes? What are the kinds and levels of manipulations or modifications that it undergoes in the process? What are the implications these manipulations or modifications bring in on the research or knowledge production that involve the digitally archived and retrieved images?

### **Obsolescence of Technologies and Longevity of Data**

As mentioned earlier, like the analogue modes became obsolete, the digital technologies too face quick obsolescence in most cases. As the technologies of production and storage evolved or changed abruptly, the data produced and stored during the due course were largely not migrated effectively or lost permanently with the outdated or obsolete technology. Numerous technologies of digital data production, archiving and circulation have gone obsolete over time

Gadgets such as Digital diaries that could record digital data for personal use , single CCS magnetic tape camcorders, floppy disks, CD

ROMs, early generation mobile phone cameras and low resolution digital cameras and their inability to migrate the data from previous to new generation platforms due to the requirement of extra hardware with scarce computer facilities made the data remain where it was captured, stored and was eventually lost and ignored with time. A major quantity of it was just recorded and never played back again.

*Digital Diary**Handycam**CD ROM**Floppy Disk**Pager**ZIP Drive**Satellite Radio**DV Tape*

To sum up, none of the data archived from these gadgets made it to the age of high speed broadband internet and the one terabyte free storage on flickr.com and other similar cloud storages. With the evolution or obsolescence most or all of the data from the old digital gadgets were lost. We assume that the data that is digitally produced now has a possibility of becoming eternal owing majorly to the availability secure cloud storage, cheap and fast internet connectivity. Imagine the plight of storage devices which were so common during a time like the floppy disc, or even a CD rom in the contemporary moment.

Coming back to the question on photographs and their materiality, as the digital stripped photography of its traditional materiality – the film and the photo paper, the way they are archived and accessed went under tremendous change alongside the change in the production technology. However, one may argue that still the photograph can be visually searched on a computer screen through a preview or thumbnails – that is also to argue

that its primary characteristic of visibility can be kept intact. The argument against or the pertinent question is, what happens when the computer or any other interface that decodes the digital data and presents in a way we can see or experience it, is taken away. What happens to a photograph or a film that was a medium self-sufficient in itself to communicate what it intends to communicate without relying on a necessary supporting technology? In the absence of that interface or technology, the data or the knowledge archived in that storage, local or in the cloud, is as good as it was never produced or archived.

### **Biases of Cloud as an Archive**

Now I would like to discuss cloud services, and to mention specifically, the one many of us use the most owing to the abundant space that provide for free, [www.flickr.com](http://www.flickr.com). It is a service owned by Oath, previously owned by yahoo.com, the photo storage and sharing website that allows the user to store images – still and moving - up to one terabyte for free. The website lets the user to both share the images as well as to store them in private. There are so many other services too, mostly free offered by various companies who are multinational corporations such as Dropbox, Onedrive and Google drive. They come in different subscription plans both free as well as paid. Most of these can store any digital data, it works literally as an extended computer folder elsewhere ‘in the cloud’ and are supposed to be ‘safe’ and accessible from anywhere anytime.

### **Technocentrism**

There are people who keep infinite faith in these technologies. They trust the technology the way an ardent believer believes in God. They care not to question their faith in any degree : the technocentrists. For them, most technology is to be believed in without suspicion. They identify it to be the most logical and scientific, hence absolute trustworthy. Any possible question that may arise on the ethics or reliability of the technology will be conveniently skipped attributing a god like image to the technology that they use. This trust applies to the computer and mobile phone operating systems, email service providers, social media companies and so on. In absolute trust *technofanatics* share everything about them with such companies. The scariest questions are conveniently skipped, what if any of the companies that you trust absolutely breaks the trust? What if we are

locked out of our 'own' phones or computers? What if our access to the cloud is restricted one fine morning? What if flicker or google changes its policy and decides to charge you for the services? The possible questions are countless. What if the big databases where terabytes and terabytes of data are stored crash or made to crash one day? One crash may possibly erase the lion share of memories of not just technocentrists, but others too. Imagine the entangled situation where you remember that you had some memories, an idea of the kind of memories- of happiness, sadness, coloured or hazy – but not being able to access them or given a conditional and negotiated access, at a compromise of the privacy and safety of your data that you trusted to be safe and eternal.

With the ownership of the cloud – both present and future, the 'technocentrism' seems to have an unadulterated trust in the technologies that make lives easier. Most consumers don't think of the possibilities, both negative and positive, how and what these clouds can do to us. A technological coup, where a powerful corporation with vested interests taking control of the world through the technology that they own is never a distant possibility. Information is the key in invasion and occupation, that we have in plenty, and the ones who can control or restrict it can control or restrict human kind and could be the new imperialists who would span their empires beyond geographical boundaries.

Getting back to the metadata and description based archiving of photographic images, as the data is stacked and accessed based on the information or the description the archiver attributes, that data may get in to a phase where it will never be made available or be accessed in the same or similar sense the photographer intended it to make at the production stage. An image alongside its traditional materiality, could be stripped off its social and political implications and be something else from thereafter. This misrepresentation is possible in the analogue mode, but the in digital, it is more recurrent, overarching and seldom open to other interpretations. The meanings attached to digital forms are more emphasised, contextualised, associated and recurrent. Rather than letting the viewer interpret what the image's intention, the context in which the image is placed give them meanings that are more rigid and often deviant from what it was meant or intended to be.

The social, political and cultural affiliation and biases of the archiver is extremely pivotal and critical in determining the destiny of images produced. The keywords with which any digital data is archived is

susceptible to the whims and politics of the archiver. Though in case more structured and professionally managed organisational archives, key wording will have guidelines in place where the endeavour is to keep the information on images value neutral as best and effective as possible. Alongside the tagged information, most online and digital search algorithms do suggest similar data that are likely to be what you are searching in case you don't query for what exactly you are looking for. This way of recommendation with images works for people who are not so sure of what they are looking for and do expect the technology to help them with suggestions. This is the case with most millennials as their research is mostly internet based and keyword driven.

For example, as a teacher of photography, when I ask students to choose some renowned photographers' work to do a style analysis, they don't go and look for interesting images and find out who shot them. Instead they google 'top Indian photographers' or so and whatever comes up and appeals them are taken and analysed for the assignment given. This mode of enquiry not only limits the scope of discovery of images captured at various instances, but also does injustice to the realm of knowledge as a whole. Students encounter only what has been listed and indexed in the search engines and the possibility of accidental discovery of a certain image is very unlikely.

This way of archiving and retrieval, not only alters the way data is archived and retrieved, but also the cultures of discovery and knowledge production. For new scientific knowledge to be produced, it is mandatory that existing pieces of knowledge is accessed. In many instances, the existing knowledge is accidentally discovered and that leads to the production of intended or unintended knowledge. Both in case of intended as well as unintended, the nature of available knowledge is of vital importance. When the data that is made available to the researcher or producer is curated and restricted the scope of the produced knowledge also narrows down to specific and vested interests. Other multiples scopes and possibilities are curtailed by the higher levels of curation and interference. For any kind of research or study, particularly something of great importance, often a researcher need to access archived information. In case of a research that need the photography archives to be accessed, the mode of search and discovery in digital and analogue archives may generate different results. A visual search of printed photographs against the keyword and metadata based digital archive would mostly generate diverse outputs.

### Formal and Informal Archives and Public Memory

In an attempt to illustrate my argument, I cite the examples of two watershed events in contemporary Indian history. The Babri Masjid demolition and the Gujarat riots and image archives of both that show up on google search reflect the careful politics that run behind the digital data archiving/ tagging. Assuming that I as a researcher need something more than the images from the black day, I will have to dig deep. Most images that show up are of the demolished mosque, vandalism, supreme court buildings and so on. Images on public record prior to the black day are seldom available. The internet archives in general does not associate the past images to the keyword. On request, the librarian at the Hindu, Chennai commented that in the newspaper archive, there are images of the monument prior to the December 6. However, since they are not of much public interest, they do not show up. So, through a careful archiving what is achieved is an erasure of Babri Masjid as a historical monument, and its location in India's cultural past. Similarly, a search query for any kind of information from different parts of the globe on google yields different results. This kind of a moderated and curated archiving influences the material to be discovered for the purpose of analysis and knowledge production. When you search for Tiananmen square massacre, a search from Beijing yields images that has nothing much to do with the massacre in it

Considering the example of YouTube videos, we search for a particular video with whatever keywords we know about it. But what shows up stacked above and below is what is of interest here. A search for Tom and Jerry yields Tom and jerry videos alongside a couple of other cartoons. Here the algorithm assumes that I am looking for fun in general. Of the many that show is a Cinderella like video titled *sleeping beauty*. If I go for the *sleeping beauty*, what is suggested alongside has a couple of videos that are not typically meant for kids –one example is that of a video titled *Village beauty* and further pursuit leads you to adult content. This discovery depends on the popularity of the video as well as the way it is indexed to be discovered often and readily. In analogue archiving the categories are protectively rigid and possibilities of unwanted content creeping in is limited. To put it in a different perspective, we think we see what we want to see, but we see what they want us to see. The levels of deviation and progression of presentation of these images are decided by the algorithm of the search engine. Politics of ownership of the search engine and the archive is of utmost importance in this context. When these pre-planned and controlled deviation in access of archived data is considered normal, the popular schools of knowledge will also be hijacked. This would eventually lead to a scary future where minds

are hijacked – an involuntary hijack, where the hostages are unaware of the fact that they are captives and their intellectual capital is modelled and controlled by the archivers. Participants are hypnotised and act under the false consciousness of being the custodians and producers of free and independent knowledge.

In the subsequent Real vs Fake scenario, the idea of fake being more appealing through misrepresentation, description and metadata.

### **Meta Data determined reality:**

In Encoding / de coding Stuart hall presents:

"The level of connotation of the visual sign, of its contextual reference and positioning in different discursive fields of meaning and association, is the point where already coded signs intersect with the deep semantic codes of a culture and take on additional more active ideological dimensions." (Hall, 2001).

While Stuart Hall describes the ideological possibilities of visual signs mediated through the semantic codes of a culture, my proposal is to focus more on the encoding phenomenon itself that assembles data based on the hegemonic narrative. In a popular representation on metadata, Hal Draper's short story, "MS\_Fnd\_in\_a\_Lbry" (1961) is a satirical description of human civilization's fall due to information overload and the inability to catalogue and access the information. The story depicts human knowledge being compressed and stored in a drawer size box, which has to be accessed through complicated indices and bibliographies. This outgrows the size of all knowledge. The story follows the phantasmagorical narrative line of an anthropologist looking for thread to lost civilization of the humans. (Draper, <http://folk.uio.no/knuthe/kuriosa/draper.pdf>)

### **Future of Curated Archives?**

Draper definitely suggests the fall of human civilisation due to the excess dependency to digital archives. Many dystopic imaginations recreate the digital imperialism and control over the world motif, especially in Hollywood productions. But beyond these popular culture depictions it is quite important to rethink the taken for granted idea of the apolitical nature of digital archiving and the immense dependency to such archives. "Derrida

in archive fever: A Freudian impression” asserts the fact that archive cannot be independent of what is archived. The structure and form of records are informed by the contents and external bodies of knowledge. The archivists are aware of the conditions under which their data attains authenticity and propriety. (Derrida, 1996.)

---

## References

Benjamin, Walter. *The work of art in the age of mechanical reproduction*. Prism Key Press, 2010.

Brennan, Karen. "Beyond Technocentrism." *Constructivist Foundations* 10.3 (2015).

Brown, Richard Harvey, and Beth Davis-Brown. "The making of memory: the politics of archives, libraries and museums in the construction of national consciousness." *History of the Human Sciences* 11.4 (1998): 17-32.

Cook, Terry, and Joan M. Schwartz. "Archives, records, and power: from (postmodern) theory to (archival) performance." *Archival science* 2.3-4 (2002): 171-185.

Derrida, Jacques. *Archive fever: A Freudian impression*. University of Chicago Press, 1996.

Edwards, Elizabeth, and Janice Hart. *Photographs objects histories: on the materiality of images*. Routledge, 2004.

Edwards, Elizabeth, and Janice Hart. *Photographs objects histories: on the materiality of images*. Routledge, 2004.

Foucault, Michel. "The order of things: An archeology of knowledge." *AM Sheridan-Smith* (1970).

Garde-Hansen, Joanne. "MyMemories?: Personal digital archive fever and Facebook." *Save as... Digital memories*. Palgrave Macmillan, London, 2009. 135-150.

Hall, Stuart. "Encoding/decoding." *Media and cultural studies: Keywords* 16676 (2001).

Harris, Paul A., et al. "Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support." *Journal of biomedical informatics* 42.2 (2009): 377-381.

Introna, Lucas D., and Helen Nissenbaum. "Shaping the Web: Why the politics of search engines matters." *The information society* 16.3 (2000): 169-185.

Lager Vestberg, Nina. "ARCHIVAL VALUE: On photography, materiality and indexicality." *Photographies* 1.1 (2008): 49-65.

Logan, James, et al. "Audio and video program recording, editing and playback systems using metadata." U.S. Patent Application No. 10/165,587.

McLuhan, Marshall, and Quentin Fiore. "The medium is the message." *New York* 123 (1967): 126-128.

Papert, Seymour. "A critique of technocentrism in thinking about the school of the future." *Children in the Information Age*. 1988. 3-18.

Reed, Drummond Shattuck, et al. "Computer-based communication system and method using metadata defining a control structure." U.S. Patent No. 5,862,325. 19 Jan. 1999.

Tyacke, Sarah. "Archives in a wider world: the culture and politics of archives." *Archivaria* 52 (2001).

Walther, Joseph B., Celeste L. Slovacek, and Lisa C. Tidwell. "Is a picture worth a thousand words? Photographic images in long-term and short-term computer-mediated communication." *Communication Research* 28.1 (2001): 105-134.

Yee, Ka-Ping, et al. "Faceted metadata for image search and browsing." *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 2003.