Centring Ideas of Personal Digital Context on Behaviour and Mindset:

Presentation by Catherine Hobbs for UBC Symposium, 17 Feb. 2012

Frame

-position piece to challenge archivists and archives studies students in the audience to centre on the particular contexts of individuals' fonds: archives that cohere around individual people

In the end, while the paper approaches the creator's mindset and behaviour in creating and keeping the fonds, it directs us as archivists to change our own mindset to focus on them.

I have discussed before that broad strategies and methodologies developed for organizational archives are a poor fit for personal archives where the motive and the context emanate from an individual him/herself. I have criticized archivists for not taking a close enough look at the features of personal archives in order to adequately capture their contexts. The particular circumstances of these archives' creation include minute decision-making based on the habits and predilections of a particular individual. However, there has been a steady move by archivists toward broader social contexts, macro-thinking about records and an attempt by some to align approaches to personal archives with those used for organizational records. These other approaches take us immediately out of the context of creation. Because personal archives are fundamentally different in nature and call on us to treat them accordingly, I've been asking a simple question: Why do we want to veer away from this personal context immediately?

Personal archives force us to reckon with the use, attitudes and habits of their makers. If we are to understand them archivally, we must grapple with the personal context of their creation and the implications of their creation and keeping for the creator him or herself. [Creator vision]

Theorists first approaching the digital have also tended either to emphasize large scale models for mapping or technological solutions for dumping. The emphasis has been until relatively recently on technical challenges focused on data integrity and safeguarding the links of a record to provenance. There have been significant developments concerning digital archives such as ingest tools and workflows, migration strategies, structuring metadata and planning pre-custodial intervention with archives' creators. The emphasis on the aspects of authenticity and trustworthiness, though has implied that the qualitative nature of these archives and their personal contexts are secondary.

We have been hearing for a number of years that technology has fundamentally changed our interactions and modus operandi of life. Individuals have been migrating toward a more robust digital existence in ever-increasing numbers; proving every day that they live qualitatively different lives by virtue of technology. The relationships between individuals and their technologies as well as their attitudes towards documenting have been creeping into archival literature in only tangential ways of yet. This may be the traditional archival lag or inertia inherent in our profession. We have yet to focus wholeheartedly on how attitudes, habits and individual choices form the grain of digital existence and add context to personal digital archives.

Personal archives in digital form are different from organizational records in that they are dependent on their personal context for any sort of cohesion as storage and removable media make these personal information ecologies ever-more dispersed. The habits of mind outlined by Catherine Marshall, in lives involving multiple devices, decentralized creation and storage, bring with them new attitudes toward proliferation, keeping and loss. Marshall describes a sanguine attitude about potentials for loss, inabilities to think in scale, and ideas that dispersed storage and the Cloud will store it or maintain it over time for you (i.e. a sense of non-responsibility). The individual is the only one in this dispersed context who has a grasp of the extent of the fonds and the fonds is bound by his/her conceptions, however inaccurate. The individual will be the one who can locate the boundary of the fonds and, similarly, because the arrangement

is provisional and unclear and the fonds relies on what the individual thinks he/she is doing, this makes the individual equally the only one for whom the order is apparent. Technology is allowing personal and organizational archives to drift further apart because of the freedom of movement and customization inherent in personal technological environments. In fact, because we are reliant on individual activities which define and map these fonds, this may be when this fundamental difference between personal and institutional archives achieves a full airing.

Recognition of this reliance on personal context is gradually creeping into archival projects. With reference to the Salman Rushdie Digital Archives project, Peter Hornsby a software engineer at Emory University was quoted as saying, "the imprint of the writer's personality lies within his computer.[1]" Archivists and others are hinting at elements of original order in ways that ways that evince the personality and choices of the creators of those archives. They are recognizing the technologies' potential for detailing documenting habits of the creators, though not necessarily in a holistic way.

What might we think of as orders of personal meaning in the digital realm? I am suggesting the phrase orders of personal meaning to give us a bit more flexibility in a digital world: where practices are provisional, proliferate and change constantly and where we have to deal with the non-physical.
-away from debates about original order and concepts of origin as well as notions of authenticity from this which have tended to dominate
-4 dimensional (i.e. including time) and multilayered notion of orders of personal meaning (i.e. orders can overlap and phase out)

Digital archives present a new twist about the classic elision between the intellectual and physical components of original order and their links to archival arrangement. Unfortunately, these links had never been fully discussed in the realm of paper which only compounds our present confusion or leads to a desire to jettison the concept on

[1] "The Author's Desktop" by Mary J. Loftus *Emory Magazine* Winter 2010
http://www.emory.edu/EMORY_MAGAZINE/2010/winter/authors.html

the part of some. However, there are some elements of our traditional interpretations of original order which might still produce interesting interpretations in a life with digital technology.

In the physical world, we were inclined to see proximity and grouping together as indicating meaning. Now, where we once had a physical site of creation of archives, this has been augmented by a virtual desktop environment and further complicated by the use of portable technologies that may or may not relate to other devices and technologies. The confusion about which evidence of original order takes precedence (physical or intellectual) is flipped back on itself in a virtual world. So are we talking about the juxtaposition of documents anymore? Did we ever really do that? Is there a physical or a virtual proximity?

But technologies still rely predominantly on visual/spatial referencing and hierarchical access modes like file trees. So there are places where we can consider whether detailing issues of grouping and proximity still have meaning.

On the other hand, there are moments in digital contexts when arrangement can become purely intellectual because the site of the record's creation is completely out of sync with its creation or where the platform doesn't matter because the creator's devices are networked. This implies a disembodied or floating interrelationship among records.

Other confusions rear their heads like the classic confusion between a file folder and an archival file is further extended, complicated or for some rendered mute by the tension between virtual folders and physical carriers. Though this may be construed as a Digital Stone Age issue, the physical storage device or the hardware might represent everything an author took with her to Paris to work on a particular literary work. In such a case, the physical carrier is probably working both like a folder and an intellectual file or sub-series and the laptop might work like a series of its own. So we still need to keep our eye on the intent to group from the creator and not just on groupings that seem

apparently to be folders or series. The key is in tracing intention and use. In the reverse, as Jane Zhang's recent thesis on original order and the digital[2] mentions, "the act of filing, including the decision to file (keep), is obviously missing in the uncontrolled accumulation of email in the inbox." i.e. The intent is not in evidence which is itself a clue.

Personal context involves the motivation for placing things together and issues of action and intention (as fraught as these concepts are after Postmodernism). Personal archives primarily stand as evidence of what someone was doing and thinking. So, really it's a simple question: what were they doing and thinking within this personal information ecology? How much of this is traceable? If we don't look for that, we get something denuded of personal context. If we don't do that, we get something more akin to data than personal archives.

Digital personal lifelogs are automatically generated data about human existence created by cameras and other devices. For example, Chen, Kelly and Jones (University of Dublin School of Computing[3]) have discussed how individuals can search images automatically generated by a webcam. Their "Personal Information Archives" uses a timeline and particular search functions to augment and trigger memories about events and they tested memory cues that were automatically extracted. This is a fascinating area where this type of information is seen in the context of biometrics, such as the intelligent insole that measures responses of the foot.

It might be useful here to distinguish on this basis between digital personal archives and other things such as these automatically generated data sets, based on archival sensitivity to provenance: i.e. there was a creator carrying out activities, intent....] Again, why obliterate that?

[2] Zhang, Jane. *The Principle of Original Order: the Organization and Representation of Digital Archives* dissertation, Simmons College Graduate School of Library and Information Science, 2010.
[3] Chen, Kelly and Jones SenseCam 2010 Proceedings http://doras.dcu.ie/16340/1/SenseCam2010cheny.pdf

Personal contexts include use of technology and attitudes about it. The digital realm lends creators even further opportunities for choice, nuance and personalizing of workspace, apps or technologies. You can say you use a Mac Workbook like you prefer a certain pen but the implications of how you use it are both modifiable and traceable.

"Personal archiving is, by nature, a personal system" this is a quote from a study dealing with the hybrid office environments of scholars (CHI 2006[4]) A survey centring on the question "why archive?" and which came up with a number of motivations for this activity playing out in the physical and the virtual office spaces. These were: to create a legacy, to keep objects of symbolic value ("tokenism"), sharing, anxiety and identity construction. These scholars cited their subjects as engaging in "value-laden archiving" and they claimed that flexibility and fine-tuning facilitate extreme use and hyper-customization that can be used to construct and project a sense of identity. All of this is prescient, in that given the date (2006), the article studied only environments with desktops, laptops and optical media (not the more customizable and transportable devices than are prevalent now).

Studies of Human Computer interaction are important for understanding possible relationships toward technology in this fast-changing landscape of possibilities. Abigail Sellen and the Socio-Digital systems team at Microsoft study everyday behaviour in order to design better systems. Sellen, Odom, Harper and Thereska recently published a paper reflecting on how materiality and practices toward possessions are affected by storage of these possessions in the Cloud[5]. Again, some of the same values studied by earlier personal information management studies make a showing: "Like physical possessions, virtual ones too play an important role in how people assert their identity,

[4] Kaye, Joseph et al. "To Have and To Hold: Exploring the Personal Archive" *CHI 2006 Proceedings* Personal Information Management April. 22-27, 2006, Montreal, Quebec, Canada.

[5] William Odom, Abigail Sellen, Richard Harper, and Eno Thereska, Lost in Translation: Understanding the Possession of Digital Things in the Cloud, in *ACM SIGCHI Conference on Human Factors in Computing Systems* , ACM, 5 May 2012

realize their aspirations and interconnect with the lives of others. It is no wonder, then, that as users of contemporary technology increasingly engage with their digital stuff, seeking to place it in secure storage, sharing it with others, and sometimes wanting to know 'who has it' or 'where it has gone', that they end up worrying about rather profound issues." The emphasis of this scholarship is to inform building of a better, more responsive Cloud. The archivist should use this type of research for intelligence gathering about potential attitudes and concerns of creators but on a very practical level the archivist would aim to be situation-specific and trace particular behaviours of the individual creator. We can use these generalizable principles in terms of heightened consciousness of those possibilities.

Erika Farr at Emory University used the term "digital biostructure" to designate the hardware and software environment that the creator interacts with in life and the environment which sustains the digital life[6]. Emulation has been used by Emory to represent interior provenance and the look and feel of documents within a workspace taking the way the creator saw the environment at a given time as a base point. The Salman Rushdie digital archives project emulation is a highly focussed presentation of the Mac Performa 5400 computer as the creator saw and used it in the mid-1990's. This approach joins together the facets of the original platform with provision for reference service.

In terms of *modus operandi* of document creation, one of the most useful veins of inquiry at the moment is in Digital Humanities scholarship particularly dealing with the concepts around digital materiality: the ability for a digital object/platform to be meaningful for its digital format or use as a tool as much as for its content.

Paul Leonardi, a professor of communication studies presents two definitions for "material" which are helpful: first, materiality is seen as "the use of tools and their affordances" (i.e. abilities to affect, or facilitate action on the part of an individual) 2) the second definition employs the term "technology-in-practice" , that is that when used, an

[6] "The Author's Desktop"

artifact becomes different things to different people (e.g. a chair can be to sit on or act as a ladder...etc)[7].

The NEH project on born digital literary material (which involved Emory University, the Harry Ransom Centre, and the University of Maryland) focused on hybrid archives and on literary creators and, although their findings deal only with what they refer to as "creative content originators" (i.e. artists), some of this certainly has a broader application.

In reference to this question of digital parameters that constrain or facilitate composition, the White Paper from the NEH project (2009[8]) says, "we would gloss this as a curatorial sensitivity toward the uniqueness of individual instances of both hardware and data objects, coupled with an awareness of how the affordances of particular systems, environments and technologies can all impact the creative process. For example, knowing how much of a document would be visible on a screen at one time... ...can be critical to understanding aspects of an author's composition process". (21). The NEH group emphasize not only capturing the digital, then, but also the hardware and carriers: machines as artifacts themselves in an holistic vision of their use. This is something that galleries working with new media artists have been dealing with for a while. The NEH participants interviewed creators to understand habits of composition and opportunities for intervention that the creator has into creative space. One would ideally wish to capture all the functionality of the StorySpace software used by author Michael Joyce in his hypertext literary work, for example. Literary authors and writers stand at one end of the spectrum in terms of the weight of their artistic intention within the archives but there are ways in which personal archives creators are all, knowingly or unknowingly, affected by the affordances of the tools before them. This curatorial sensitivity the NEH reports outline is key in approaching a particular individual for a robust sense of their archives. We can look further at the crossroads of archives and the constraints affordances of

---

[7] Leonardi, Paul M. "Digital materiality? How Artifacts Without Matter, Matter" *First Monday,* Vol. 15, No. 6-7, June 2010. http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/3036/2567

[8] Kirschenbaum, Matthew G. et al *Approaches to Managing & Collecting Born-Digital Literary Materials for Scholarly Use*, White Paper to the NEW Office of Digital Humanities Start-Up Grant, May 2009. http://drum.lib.umd.edu/bitstream/1903/9787/1/Born-Digital%20White%20Paper.pdf

technology. Digital Humanist Matthew Kirschenbaum in his *Mechanisms: New Media and the Forensic Imagination*[9] suggests the concept of system opacity: that most people don't see the cogs turning in their computers "creating" as he calls it, "an open-ended symbiotic exchange between computation and representation." He says, "many of the plain truths about the fundamental nature of electronic writing [remain] apparently unknown at a simple factual level, or else overlooked, or their significance obscured." While Kirschenbaum uses this concept to discuss computers in new media literature, from an archivist's perspective this type of opacity would play a central role in a creator's understanding of his/her platform and use of technology. Computer literacy is not the only thing at play here, then, but an imagined understanding of the technology which the creator acts upon.

Right now, most archivists still seem to be acquiring and processing acquisitions off of removable media or hard-drives. Jane Zhang's 2010 thesis focused on original order and its relevance to digital archives. She emphasizes that archivists will not have time to do processing beyond the series level and that records will be automatically mapped in an original order and stay that way[10]. This is a digital parallel to leaving paper records in original physical order in the boxes as received. The forensic recovery projects I am familiar with have emphasized the issues of processing and preservation: with the disk image as a starting point. But a disk image is a random storage of bits of information: from with a logical list of files and file formats with metadata is created. It is from this view that files are then selected and put into series in a way that seems to mimic traditional approaches to disordered paper records with some help from file directory structures and naming conventions. We seem to be dealing with only the physical order/disorder counterpart and not with the conceptual elements of order as the creator saw them within archival arrangement. There needs to be a bridge between the confirmation of data integrity and trustworthiness and file format identification on the one hand (a model that centres on the technology) and these individual personal contexts

[9] Kirschenbaum, Matthew *Mechanisms: New Media and the Forensic Imagination* Cambridge Mass.: MIT Press, 2008 p. 62?
[10] Zhang, Jane. *The Principle of Original Order: the Organization and Representation of Digital Archives* dissertation, Simmons College Graduate School of Library and Information Science, 2010, p.88-89. 104, 162, 167

centralizing how it was used and envisioned by the creator on the other. Basically forensics centralizes systems, not people even in situations where the archivist has a heightened sensitivity to the personal digital biosphere and if it is not our intention or within our budgets to do emulation all the time, we have an even greater need to capture these contexts in some way. Even though emulation takes the creator's perception of the workspace as the underpinning of the emulated environment the hope is that this is tacitly transmitted to the creator by their experience of the workspace. Perhaps this too falls short if, instead we can provide an articulation from the donor about their use of platforms and devices explicit references and in their own words.

The sense of the whole resides with the individual. We recognize that this whole, based on Catherine Marshall's research is a blurry whole with assumed or forgotten boundaries. But this appears to be an extension of the normal facets of human memory and provisional storage of documents such as was seen in the analogue world. This is a question of proliferation, not a qualitative difference--if we look at the paper and hybrid offices studied by earlier PIM studies. As with so much else within the stream of constant information, we have to give up on ideas of complete control or totalization. There is (and probably never was) a totalizable fonds, perhaps with some parts of it we need to rely simply on the map to getting there and the creator's say so about what they have and what it means. Archival appraisal, though, is a one-shot deal in terms of the "what does this mean to you" question. This is the context that will disappear if the archivist fails to ask the right questions.

With reference to all of these issues, we can't rely on harvesting metadata to get to them. We have to do them in a semantic way and then we can represent them in terms of visualization or narrative form. The question of personal contexts is not perhaps fully or immediately "solvable" by archivists. It takes a leap or gesture of understanding on our part and questions "what might have personal meaning". Our processes are fallible and there are a multitude of potential traces. We don't have new tools to deal with this new reality in terms of capturing context so this approach is not as sellable as metadata.

It involves effort and reorientation but it might be our large step that renders minute detailing within description a question of choices and priorities.

What I hope we have in the short-term future is a well-defined and accepted practice of gathering and disseminating information about these particular contexts. There are rich ways to combine images of physical workspaces with virtual workspaces; to add maps or modelling of relationships between technologies and how these relationships change through time (which could be interactive or multi-faceted depending on the depth of understanding of the relationship of these platforms); we can also easily adopt a practice of outlining the habits and attitudes toward with technology in a documentary or technological sketch which would be included in archival appraisal and later transferred to archival description. There are also other practices that the archivist has access to now but could make better use of: for example detailing arrangement issues from either physical or virtual environments and about the interaction between these two environments in an arrangement note.

Some of our old approaches would become new again by addressing contemporary realities. Such complex and individualized settings for archives are asking archivists for a nuanced understanding of digital personal archives and require them to go forward with an approach to appraisal and arrangement and description that takes these into account.