Start Making Sense: Database, Media, and Narrativity

John Fass
Royal College of Art
Kensington Gore
London, UK
john.fass@gmail.com

Abstract
This paper discusses the exploration of multimedia archives, and suggests ways of structuring navigation experiences conducive to the production of meaningful narratives. As archives begin to accumulate and catalogue extensive digital records methods of access and engagement are also adapting. Two categories of digital public space are explored; extensive curated public collections such as the BBC archive, and less obviously curated user-generated data such as Flickr[1] photos and browsing data. The use of the term archival is here intended to refer to any group of materials organised in a structured way. The difference between traditional means of information structure and digital database ordering is also examined. Explorations of interface, narrative, and interaction frame the research, and an overview of relevant projects is provided.

CHI'13, April 27 – May 2, 2013, Paris, France.

Author Keywords
Archive science; narrative construction; public archives; user trails, digital patina.

ACM Classification Keywords
H.5.m. Information interfaces and presentation

1. http://www.flickr.com/

Introduction

National public collections, such as the BBC, hold many millions of items in a wide variety of different formats. Private digital archives consist of videos, photographs, and music but also the more informal data that is captured when people click around their favourite websites and participate in social media networks. A distinction is made between digital version of non-native formats (e.g. digitised VHS videos) and native digital formats (tweets or Facebook updates). Both types of data occupy digital public space and both embody specific challenges for meaning-making, navigation, ownership and curation. Silos of data do not produce meaning without cognitive filtering and the volume and variety of data poses a serious challenge to categorisation. Curating practices are in the process of adapting to a new definition of visitor generated content. This paper proposes to integrate the two categories by suggesting narrative construction as an organising principle.

Background

Archive science has concentrated on what Breakell calls 'the physical or notional locus of the archive' [1]. A key principle is that archive materials should be arranged and presented with neutrality. The archive continuum refers to the way in which archival records adapt to different types of enquiry and method over time [11]. Archives allow users to be selective, to choose those items that reflect personal concerns or confirm hypotheses [3]. While formal media archives can adapt these principles to virtual objects, informal digital records are less easy to categorise.

New types of digital material such as the hundreds of millions of tweets (and associated metadata) archived by the Library of Congress every day are changing the way archives are understood and used. The Twitter archive represents 'more observations, recorded at the same time, by more people, than ever preserved in any medium before'[19]. Parallel to the data deluge produced automatically by the social media network, personal data collections now include biometric measurement, medical records, cellphone movies and computer game scores. As these new data types appear in formal archives and personal collections in a more organised and curated way, new methods of structuring and understanding them through narrative production are necessary.

Database

While databases can drive narrative experiences (e.g. computer games), they come pre-loaded. The narrative environment is generated by a games designer. The game's database is subsequently configured to deliver the correct narrative elements at key points in the user journey or in response to specific user input, e.g. turning left in a virtual cityscape. Ryan [14] claims narrative elements, in contrast to database entities, 'cannot be freely permuted, because they (must be) held together in a sequence by relations of cause and effect, and because temporal order is meaningful'. Databases are inherently hierarchical configurations, they are designed for non-linear atemporal expression since they must provide for a (constrained) multitude of different outcomes.

Some theorists establish an opposition between database and narrative claiming: 'new media objects do not tell stories; they don't have beginning or end…they

are collections of individual items' [12]. These items are brought forth (performed) by the interface in a series of operations. The interface (symbolising the system) acts on the database (in the form of a sequential list) in a series of steps carried out by the user. The database structure contains an algorithm [12], a set of instructions, that translates interface actions into output, (e.g. clicking the search button). This is seen as merely a simulation of narrative since there is no actor or narrator, and there are no progressively connected events. Public media archives can be seen as vast databases waiting expectantly for interaction, this paper proposes some ways of extracting meaning through narrative from them.

Databases could be designed as self-replicating organisms. Synthetic biologists [10] try to create programmable artificial cell units, their intention is for artificial cells to carry out predictable functions in the same way software produces consistent outcomes. Building on von Neumann's system of self-reproducing automata [21], individual database cells could be designed to react to their neighbours, to varying query patterns and to mutate in response to new entries. For example, as material is added to the archive, its closest thematic neighbours could cluster towards it automatically.

Databases drive the digital [9] and, whether expressed physically or virtually, the digital world has a bias towards abstraction, centralisation, and atemporality [16]. An examination of the material properties of databases, and their effects, can redirect attention towards the humanising (and uniquely human) nature of narrative production. By augmenting archival systems with the digital conditions for human narrative construction (rather than simulations or abstractions) systems design and interfaces can orient towards meaning.

Interface

Traditional archive materials are organised around familiar interfaces, (such as books, DVDs, and physical objects), the institution and its staff, policies and location can also be seen as an interface. Physical spaces have clear affordances for archive work: tables, lights, chairs, shelves, but they also have staff with embodied knowledge and both procedural and navigational expertise. The visitor's experience is mediated through human structures that reflect centuries of knowledge management.

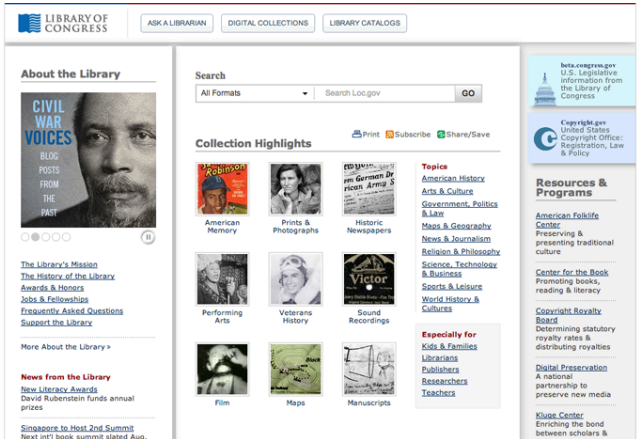

Figure 1. Library of Congress http://www.loc.gov/index.html

Digital collections rely on a visual representation of this knowledge, and an interaction model to deliver their materials. (see Figure 1) In the same way that media archives often feature digital versions of analogue

objects, their interfaces can be seen as digital versions of the knowledge and data architecture of the institution.

User interfaces perform their own model of abstraction and representation on the database. They act as (more or less symbolic) metaphorical querying mechanisms. Direct manipulation metaphors abound i.e. 'pushing' buttons, 'opening' files, and 'sending' emails. The resulting narratives are similarly metaphorical. The implicit nature of digital activity; connecting, interacting, querying, and retrieving is hidden. By making these operations visible the elements of narrative construction can be made more apparent.

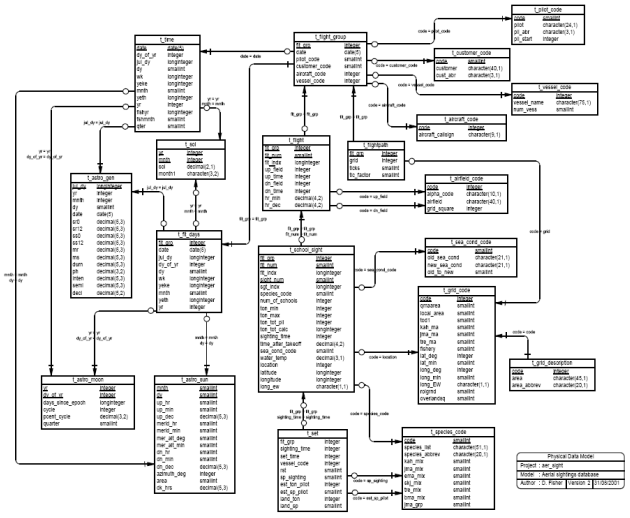

Figure 2. Database entity relationship diagram.

Archive queries are often presented as a list, the default form of the database. A series of on/off choices lead people through the archive in a reflection of the binary software processes that deliver archive materials up to the interface. Some researchers [2] have suggested inferring user queries from browser data, obviating the need for explicit user queries altogether. Bypassing the conscious process of database interrogation has an implication for interface, learning and for the generation of meaning. It leaves decision making to an algorithm, cuts out the interface, and limits results to those formats and entities easily retrievable by software standards. This paper proposes a more conscious and transparent mechanism in which data entities are represented not by the choices they offer but by their relationships (see Figure 2). Interactions are captured and shown to users, building up a genome of possible meanings.

Narrative

Story making and narrative creativity is fundamental to humanity [13]. Stories are an essential mechanism in transmitting information, performing identity, 'managing complexity, and expressing context awareness' [4]. Much of how we understand the world depends on our ability to tell stories about it [17]. Using the five basic narrative queries; *Who, What, When, Where* and *Why*, one approach [5] outlines how personal digital devices could develop context awareness. *When* and *where* are answered by the automatic time stamping and GPS metadata that accompanies most digital data capture. *What* people are doing is harder to measure but could be addressed by the concept of programming by demonstration (PBD) [5], where the details of a task are extrapolated to overall task identification. Combined with spatial and

temporal data, PBD goes some way to developing contextual awareness and therefore narrative complexity for machine intelligence.

Who and *why* present much more of a challenge for digital devices. Identity awareness involves a complex navigation of personal relationships [4], while answering the question of *why* a person acts as they do is considered impossible for a machine [5]. Devices are currently limited to demonstrating to a user the reasons for their *own* behaviour, i.e. running out of battery power or searching a particular database. Structuring not only collection records around narrative questions, but also database queries could prioritise narrative creation and understanding.

Walter van de Velde applies a similar idea to the 'techno-social inhabited space' [20] around us, which is seen as programmable by our behaviours. By acting in the world we create a 'behaviour landscape' populated by information. This information demands, above all, attention, 'without which it is dead' [20]. Without interaction or engagement, the contents of electronic databases cannot be said to be embodied. In the same way, the archive contains the elements of all possible narrative probabilities and it too requires attention. What might the potential behaviour (or meaning) landscapes embodied in the archive become?

One way objects and systems communicate their function is through patina. The doorframe with hand marks indicates which way the door opens, the most scratched record is the most popular etc. Patina can be defined as the accreted residue of activity along with the subtractive wear of subsequent use. Digital patina is represented by the trails content has followed (reblogged, retweeted) and the less measurable ways interest paths have influenced the archive and then faded from collective consciousness. If these trails can be made more permanently visible, their patina can show a direct correlation with narrative construction in the same way stories can take shape over generations.

Interaction

An important consideration for interaction with archives and databases is how meaning is made, how interactions are structured, and what kind of affordances are present. Kirk et. al. [8] show how arranging physical objects in 3D space (in this case photographs) differs from arranging digital objects on a 2D table top. Unlike the (screen based) digital world, physical objects can be held in either hand and viewed at different distances to the eyes. The height and weight of a group of photographs gives a clue as to their number, and direct comparison can be done in different modalities.

Taking the comparison further, recent research [15] explores how people make meaning from physical versus digital objects in the home. Physical mementos were found to have spatial meaning, accruing different association depending on where in the home they are placed i.e. private or public rooms. They were also seen to afford immediate access and to be curated, (highly selected) reliable, and permanent. Digital mementos were difficult to organise, (people save multiple copies in multiple folders), often hidden from view, (on the home computer) and considered unstable and impermanent. These findings have implications for the future of archives both personal and public and how (and where) they should be configured for maximum narrative potential.

Dominant forms of digital interaction design focus on the user experience (UX). They prioritise productivity, efficiency and throughput over meaning, connectivity and playfulness. Interaction structures are based on the cognitive tradition in human factors research and interaction designs reflect this with much use made of hierarchical menus and linear scrolling. The spatial desktop metaphor prevails: files are 'opened', collections 'searched', and apps 'launched'. Interactions possible in the context of archive collections remain tied to the page based web paradigm even though some of the archived material itself is generated in a very different way. e.g. pulse readings or financial data. Interactions are not seen as providing an experience beyond the immediate context of the institution: they do not provide for downloading, remixing, reconfiguration or sharing.

Three practical examples illustrate alternatives: *Clouds* [7], is a documentary film made using computer vision (see Figure 5). The film is conceptualised as a database of conversations and is a fully searchable, tagged narrative digital archive. *Clouds* represents a new paradigm of image making. Instead of being the transcription of a temporal process the film is an entirely native digital format. It does not need *post hoc* tagging, processing, or categorisation since metadata is automatically generated in the same way cellphone photographs are time stamped and geolocated. If fully meta tagged stories could be told about (or within) multi media archive items, the process of story and meaning making could be made more immediate and tangible.

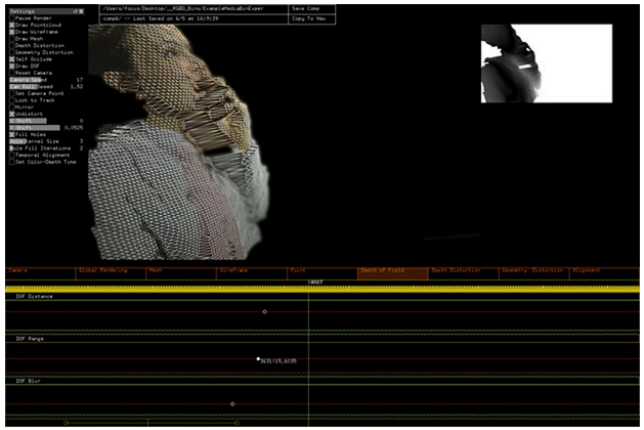

Figure 3. Clouds http://vimeo.com/42852185#at=0

Lifestreams [6] is a novel software architecture designed to provide an alternative metaphor for handling information density and file organisation than the desktop model. Lifestreams proposes a time-ordered stream of documents combined with a highly constrained number of operators for locating, summarising, monitoring and annotating (see Figure 4). Lifestreams emphasises the temporal over the spatial providing an obvious mapping with the most basic category of metadata; the time stamp. The system allows for diverse user interactions such as sharing transversally across the stream e.g. what did BBC archive users look at on November 23 at 11.00? If archive collections, or a personal selection of them, were organised in this way the concealed and unconscious nature of digital interactions could be revealed and the personal integrated with the public.

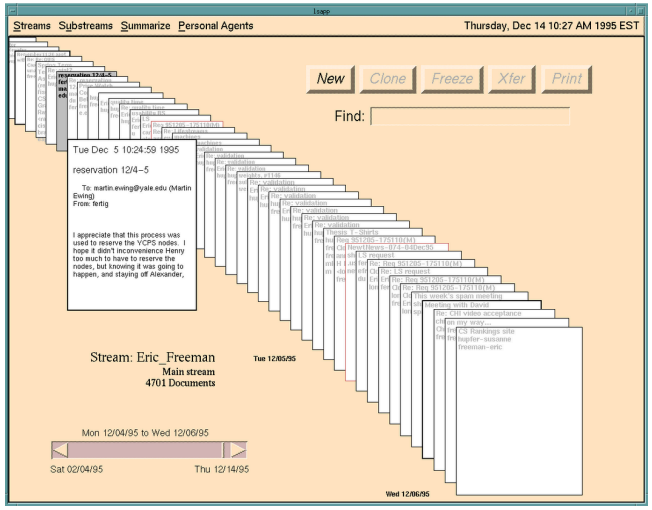

Figure 4. Lifestreams
http://www.sigchi.org/chi96/proceedings/videos/Fertig/etf.htm

SP-ARK [18] is an online platform for exploration of a personal media archive. The archive relates to the film *Orlando* (1992) by Director Sally Potter and represents all her personal materials, both formal and informal, generated during the making of the film. The SP-ARK interface offers a detailed breakdown of the filmmaking process from development to distribution, with diverse archive items such as notes, photographs and call sheets assigned across six categories. People can record their research interest pathways through the material, annotate items, and send messages to other users. A type of digital patina emerges with indications of which items are most searched and commented. Consciously implementing this strategy to bring about a user interest patina where public and private archives intersect is a way of allowing meaning making to take place.

Conclusion

This paper has outlined a background for the narrativisation of digital public space as a way of drawing meaning from discrete data objects. I have proposed some ways of informing a new field of enquiry by engaging with work from biology, information science and narrative research. The paper also defines the problem of new and ubiquitous media types, the disconnect between public and private collections, the lack of means to experience them fully, and the need for curation of the data stream. The next phase of archive development will see a blurring of the boundary between public and private records. This distinction is being rendered less meaningful by the highly granular nature of automatic digital data capture.

Systems designers aim for increasing levels of individual personalisation and the exploitation of personal context, behaviour and activity. For example the information ecosystem that integrates medical records with self-measured blood pressure or weight data is an increasing focus for personal informatics research. As systems design begins to deliver seamless hybrid digital/physical experiences, the need for coherent and meaningful narrative frameworks to make sense of them will increase. Directions for further research in this field include: curation (amateur and professional), the nature of digital storytelling, database as self organising entity, and the affordance of digital patina.

References

[1] Breakell, S., *Perspectives: Negotiating the Archive*. Tate Papers, 2012.

[2] Brenna, L., Gurrin, C., Johansen, D., Lee, H., Smeaton, A. F., Zagorodnov, D., *Supporting Mobile Access to Digital Video Achives*. Proc. MobileHCI'06. 2006.

[3] Brockmeier, J., *After The Archive: Remapping Memory*. Culture and Psychology. 16(1), 5-25. 2010.

[4] Brooks, K., *The Context Quintet: Narrative Elements Applied to Context Awareness*, Proc. Human Computer Interaction International, 2003. Lawrence Erlbaum Associates, Inc.

[5] Brooks, K., Do story Agents Use Rocking Chairs? The Theory and Implementation of One Model for Computational Narrative. ACM Multimedia 96, Boston MA, USA.

[6] Freeman, E., *The Lifestreams Software Architecture*, PhD Dissertation, Yale University, 1997. http://www.cs.yale.edu/homes/freeman/dissertation/etf.pdf accessed 19/12/12.

[7] George, J., Minard, J., *Clouds*. In WIRED, UK Edition, Condé Nast Publications Ltd. Dec. 2012.

[8] Kirk, D., Izadi, S., Sellen, A., Terrenghi, L., *Affordances for Manipulation of Physical versus Digital Media on Interactive Surfaces*, Proc. CHI 2007.

[9] Linton, P., Schuchhard, P., (Eds) The Digital World in 2025 Indicatirs for European, European Internet Foundation, 2009.

[10] Macia, J., Munteanu, A., Rodriguez-Caso, C., Solé, R. V., *Synthetic Protocell Biology: From Reproduction to Computation*, Phil. Trans. R. Soc. B, 2007. Publiched online 1 May 2007.

[11] Manoff, M., *The Materiality of Digital Collections: Theoretical and Historical Perspectives*. Libraries and the Academy, Vol. 6, No, 3. The Johns Hopkins University Press, Baltimore, 2006.

[12] Manovich, L., *The Database Logic*, Ch 5, *The Language of New Media*, MIT Press, 2001.

[13] Minsky, M., *Society of Mind*, New York. Simon and Schuster. 1988.

[14] Paul, C., *The Database as System and Cultural Form: Anatomies of Cultural Narratives*, Database Aesthetics, Victoria Vesna (Hg.) (2004).

[15] Petrelli, D., Whittaker, S., *Family Memories in the Home: Contrasting Physical and Digital Mementos*, Personal and Ubiquitous Computing, February 2010, Volume 14, Issue 2, pp 153-169.

[16] Rushkoff, D., Program or be Programmed, Soft Skull Press, 2010.

[17] Ryan, M-L., Beyond Myth and Metaphor: The Case of Narrative in Digital Media. The International Journal of Computer Game Research. Vol. 1, Issue 1, July 2001.

[18] SP-ARK http://www.sp-ark.org/sectionPage.php Accessed 18/12/12.

[19] Stross, R., *When History is Complied 140 Characters at a Time*, New York Times, May 1 2010.

[20] Van de Velde, W., *The World as Computer*, A Quest for New Horizons in the Information Society European Visions for the Knowledge AgePaul T Kidd (Ed.) Ch. 17. 2007.

[21] von Neumann, J. *Theory of Self-reproducing Automata*, University of Illinois Press, 1966.