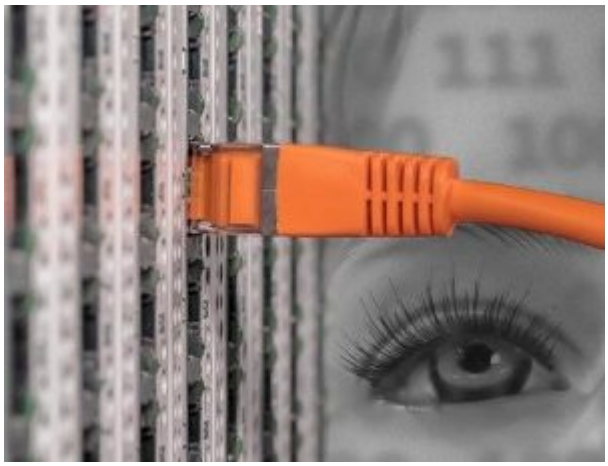
The Digital Derangement of Archives

written by Jeremy Huggett | 12/12/2019



Modified from original by Michael Schwarzenberger via Pixabay

Bill Caraher has recently been considering the nature of 'legacy data' in archaeology (Caraher 2019) (with a commentary by Andrew Reinhard). Amongst other things, he suggests there has been a shift from paper-based archives designed with an emphasis on the future to digital archives which often seem more concerned with present utility. Coincidentally, Bill's post landed just as I was pondering the nature of the relationship between digital archives and our use of data.

So do digital archives represent a paradigm shift from traditional archives and archival practice, or are they simply a technological development of them? Digital archives are commonly understood to be a means of storing, organising, maintaining, and making data accessible in digital format. Relative to traditional archives they are therefore not limited by physical space or its associated costs and so can make much more information available more easily, cheaply, and widely. But a consequence of this can be a kind of 'storage mania', in which data become easier to accumulate than to delete because of digitalisation, and where data are released from the limitations of time and space through their dematerialisation (Sluis 2017, 28). This is akin to David Berry's "infinite archives" (2017, 107), who suggests that "One way of thinking about computational archives and new forms of abstraction they produce is the specific ways in which they manage the 'derangement' of knowledge through distance." (Berry 2017, 119). At the same time, digital archives represent new technological material structures built on the performativity of the software which delivers large-scale processing of these apparently dematerialised data (Sluis 2017, 28).

There are perhaps three key areas where archives and digital data interact: the digital infrastructure itself, preservation practices within that infrastructure, and the effects of these on the digital data we subsequently use.

For example, Ina Blom argues that there is a fundamental shift between traditional and digital archives: where once content was distinct from infrastructure, “once the archive is based on networked data circulation, its emphatic form dissolves into the coding and protocol layer, into electronic circuits or data flow” (Blom 2017, 12). Of course, this is seen as advantageous in many ways, not least in terms of making archives more efficient and enabling greater access and availability. But in effect the archive is reconfigured as a database with management algorithms controlling internal data manipulation and search algorithms governing data access and retrieval:

> Search engines can be designed to find the proverbial needle in the haystack, or even to create a haystack where there are only needles, that is, build patterns where there seemed to be only fragments … As soon as new information enters a networked database, the structure of the database can reorganize itself, just like old songs change over time with changing audiences and changing social, political or cultural circumstances. Flexibility and instability have become technical qualities instead of problems to be controlled. Digital archives are unstable, plastic, living entities, as stories and rituals were in oral cultures. (Brouwer and Mulder 2003, 5).

Elsewhere, I’ve discussed the blurring of boundaries between platform and infrastructure and the implications of gatekeeping, arguing that these tools and facilities both provide a means of knowledge creation and at the same time act as (largely covert) constraints on our practice. But what becomes clear here is that the digital archive, as representing our collective endeavours as archaeologists, is a ‘living’ dynamic structure which changes its nature as the content is recontextualised by this digital environment even if the original source remains intact (Dekker 2017, 16). Yet we retain a traditional perspective of an archive as an unchanging, stable, permanent, structure while embracing the new flexibility and accessibility they offer. As Berry suggests, digital archives are “deeply computational in structure and content because the computational logic is entangled with the digital representations of physical objects, texts and ‘born digital’ artifacts” (Berry 2012, 13).

These digital representations themselves bring about a shift in archival practice: where traditional archives primarily endeavour to secure their records in the form and condition that they were in when they were collected, digital records are maintained by combination of refreshing them to reduce the risk of data loss (by reproducing exact copies of their content and format/structure) and periodic migration to overcome technological obsolescence (by reproducing their content as far as possible but changing elements of their format/structure in the process of translation to new software versions).

> With digital technologies, nothing is stored but code: the mere potential for generating an image of a certain material composite again and again by means of numerical constellations. Forget to update the software through which an encoded material is made visible, and there is little left – at least from the point of view of the cultural interface. This is not because information is ‘immaterial’ but because visibility is not a measure of its specific forms of material inscription: inscription is simply some kind of modification of an electromagnetic substratum. (Blom 2017, 12)

A related distinction is in the separation between the digital bitstream itself and the archaeological content within: the data are encoded in different binary formats which require software intervention to make their semantic content usable. In the digital context, therefore, preservation consists of maintaining their meaning and trustworthiness as records (are the copies authentic and exact, or have there been changes?) whereas in traditional archives, preservation is concerned with protecting the media themselves (e.g. Duranti 2001, 46). As a result, truly original digital data only exist as long as they remain accessible by current technology, but as time passes it is copies which persist and the 'original' is likely lost, living on in reproductions.

This transitory and mutable nature of digital data does not lend itself to benign neglect (e.g. DeRidder 2011). However, the demands of digital preservation standards and best practices make benign neglect the likely destination for a great deal of archaeological data – the cost, effort, and quantity of data can make it difficult for anyone – even the digital repositories themselves – to do anything other than store data in the hopes that there will be sufficient resources to preserve it in the future, perhaps through a demonstrable demand for those data indicating a need for prioritising its proper archiving.

All this would suggest that the very act of placing data in, and subsequently using data from, digital archives changes our relationship with not only those digital data, but also non-digital data too. The supremacy of the digital archive with its apparent flexibility and accessibility makes data that are not either accessible or digital effectively invisible: "A paradox of digitality is the way in which its convenient surfaces serve to conceal that which is not digital." (Berry 2017, 106). In an archaeology increasingly defined by and generated using digital data, understanding the potential and limitations of archaeological knowledge requires a critical understanding of the nature of our digital archives.

References

David Berry 2013. Introduction: Understanding the Digital Humanities. In David Berry (ed.) *Understanding the Digital Humanities* (Palgrave, Basingstoke), 1-20. https://doi.org/10.1057/9780230371934_1

David Berry 2017. The Post-Archival Constellation: The Archive under the Technical Conditions of Computational Media. In Ina Blom, Trond Lundemo and Eivind Røssaak (eds.) *Memory in Motion: Archives, Technology, and the Social* (Amsterdam University Press, Amsterdam), 103-125. https://doi.org/10.2307/j.ctt1jd94f0.8

Ina Blom 2017. Rethinking Social Memory: Archives, Technology, and the Social. In Ina Blom, Trond Lundemo and Eivind Røssaak (eds.) *Memory in Motion: Archives, Technology, and the Social* (Amsterdam University Press, Amsterdam), 11-38. https://doi.org/10.2307/j.ctt1jd94f0.4

Joke Brouwer and Arjen Mulder 2003. Information is Alive. In Joke Brouwer and Arjen Mulder (eds.) *Information is Alive: Art and Theory on Archiving and Retrieving Data* (V2_, Rotterdam), 4-7. Available via https://v2.nl/publishing/information-is-alive

Bill Caraher 2019. Legacy Data, Digital Heritage, and Time. *Archaeology of the Mediterranean World* 9th December 2019.

https://mediterraneanworld.wordpress.com/2019/12/09/legacy-data-digital-heritage-and-time/

Annet Dekker 2017. Introduction: What it Means to Be Lost and Living (in) Archives. In Annet Dekker (ed) *Lost and LivIng (In) Archives: Collectively Shaping New Memories* (Valiz, Amsterdam) 11-25. Available via https://monoskop.org/images/1/11/Dekker_Annet_ed_Lost_and_Living_in_Archives_2017.pdf

Jody DeRidder 2011. Benign Neglect: Developing Life Rafts for Digital Content. *Information Technology and Libraries* 30(2), 71-74. https://doi.org/10.6017/ital.v30i2.3006

Luciana Duranti 2001. The Impact of Digital Technology on Archival Science. *Archival Science* 1, 39-55. https://doi.org/10.1007/BF02435638

Katrina Sluis 2017. Accumulate, Aggregate, Destroy: Database Fever and the Archival Web. In Annet Dekker (ed) *Lost and LivIng (In) Archives: Collectively Shaping New Memories* (Valiz, Amsterdam), 27-40. Available via https://monoskop.org/images/1/11/Dekker_Annet_ed_Lost_and_Living_in_Archives_2017.pdf