

March 2018

# Overview of a born-digital archives access workshop

Held at Wellcome Collection on  
24 November 2017



The first part of the document discusses the importance of maintaining accurate records of all transactions. It emphasizes that every entry, no matter how small, should be recorded to ensure the integrity of the financial data. This includes not only sales and purchases but also expenses, income, and transfers between accounts.

The second part of the document provides a detailed breakdown of the accounting process. It starts with the identification of the accounting period, followed by the collection and classification of transactions. The next steps involve the recording of these transactions in the general ledger, the calculation of debits and credits, and the preparation of the trial balance.

The third part of the document focuses on the preparation of financial statements. It explains how the data from the general ledger is used to create the income statement, the balance sheet, and the cash flow statement. Each statement provides a different perspective on the company's financial performance and position.

The fourth part of the document discusses the importance of reconciling accounts. It describes how to compare the company's records with bank statements and other external sources to identify and correct any discrepancies. This process is crucial for ensuring the accuracy of the financial statements.

The fifth part of the document covers the final steps of the accounting cycle, including the closing of the books and the preparation of the final financial statements. It also discusses the importance of archiving records and maintaining proper documentation for future reference.

In conclusion, the document emphasizes that accounting is a systematic process that requires attention to detail and adherence to established principles. By following the steps outlined in this document, businesses can ensure that their financial records are accurate and reliable, providing a clear picture of their financial health.

# Background

This report outlines the findings of a workshop organised by the Collections Information Team at Wellcome Collection and the Department of History at the University of Sussex, to bring together researchers and archival professionals to explore methods for providing access to born-digital archives.

In early 2017 Dr James Baker from the Department of History, University of Sussex contacted the Collections Information team at Wellcome Collection to express interest in an article written by one of the team's archivists setting out their efforts to process born-digital archives.<sup>1</sup>

We met to discuss our mutual interests in more depth. Wellcome was looking for ways to build up evidence of user requirements for access to born-digital archives without being able to provide particularly sophisticated access, nor knowing many researchers who were actively using born-digital archives in their research. James was interested in the impact born-digital archives have on the historical method and in promoting born-digital archives to researchers, making it clear that they exist and are there to be used.

We decided to arrange a workshop to bring together researchers who use (or might use) born-digital archives, archivists who process this material and Wellcome staff who will have input into the design and development of future access systems. The plan was to be as hands on as possible, to provide real archival records to explore in various ways and to get the discussion started.

We organised the 16 participants into four groups which rotated around four different case study scenarios, examining each one for 30 minutes before moving on to the next. Each scenario focused on a particular aspect of user access and the participants were prompted to consider a set of questions specific to each case study. The organisers stationed themselves at each case study and both supported the participants and took notes on how they interacted with each scenario and how the discussions developed.

<sup>1</sup> Victoria Sloyan: "Born-digital archives at the Wellcome Library: appraisal and sensitivity review of two hard drives", *Archives and Records*, 37:1 (2016) p20-36, DOI: [10.1080/23257962.2016.1144504](https://doi.org/10.1080/23257962.2016.1144504)

# Case Studies

## Case Study A: Exploring records in the “original” environment

Participants explored a set of records copied from the scientist Ian Dunham’s hard drive.<sup>2</sup> The records were selected by Dunham and copied onto an external hard drive for transfer to Wellcome. For the workshop, the records (bar any that are closed for sensitivity reasons) were copied onto a laptop running Windows 7, the same operating system used by Dunham at the time of transfer.

The directory structure and original file and folder names were retained. The laptop also had open an internet browser displaying the online catalogue for the collection. In line with our current practice, catalogue records were created for every folder, sub-folder and digital file in the directory. The catalogue records for individual digital files contain the original filename and the last modified date but no other metadata, and for resource reasons they are not enhanced with additional information. However, catalogue records for the directory sub-folders do have additional information such as a date range and brief description of the content.

Participants were asked to consider the following questions:

- How would you expect and/or like to interact with this method?
- How do you find searching and browsing?
- How does it compare to case study B?

The four groups varied in how they split their time between the digital files and the catalogue, but they all found it difficult to consult the two in tandem and easily switch from viewing the digital files to finding the corresponding online catalogue record. Consequently, many participants expressed a desire to have more context presented alongside the digital files rather than have it presented separately in the catalogue.

The participants were also divided between those who enjoyed being able to browse and explore the directories and those who were frustrated by the difficulty in quickly finding specific information. This was caused partly by being asked to navigate a personal filing system that did not necessarily chime with their expectations, but also by the use of opaque filenames that did not reveal the content. Most participants saw the value in preserving the original directory structure and filenames as a way of providing insight into the creator’s logic and working practices, but they also wished to have alternative arrangements and filenames that enabled quicker searching and easier identification of topics and data. A desire was also expressed for the catalogue to be more helpful in identifying the location of specific themes, subjects and file formats.

There was some recognition that this form of access enables researchers to search not just within discrete files but also across them and hence allows for statistical analysis. However, this was mainly proposed by the minority of participants already engaged in this form of research. The majority did not consider this type of use without prompting and their approach was very much to transplant analogue research practices to the digital sphere.

## Case Study B: Accessing records through Wellcome’s Viewer

Participants explored records from the entirely digital Art and Health Witness Seminar archive, comprising three PDFs and one audio file.<sup>3</sup> The records were accessed through Wellcome’s online Viewer, the interface for accessing digitised content. The catalogue record and hierarchical tree were displayed underneath the Viewer and the only metadata available was that included in the catalogue.

<sup>2</sup> Ian Dunham archive: hard drive, catalogue ref. [GRL/IDN/F](#), held at Wellcome Collection

<sup>3</sup> Arts and Health Witness Seminar, catalogue ref. [ART/WIT](#), held at Wellcome Collection

Participants were asked to consider the following questions:

- How would you expect and/or like to interact with this method?
- What would the experience be like for a more diverse collection?
- How does it compare to case study A?

Many participants felt that this access method provides a clear link between the digital file and the catalogue record holding the contextual information, but it gives a more sanitised view of the records than case study A. By effectively turning digital files into digitised content, the Viewer also closes down some research avenues, especially as the original metadata is not accessible. It was recognised that many current researchers would probably be familiar and comfortable with this method of accessing records. It was also thought that it could cater well for both those wanting to quickly find specific information and those who might otherwise be intimidated by born-digital archival research. There was a suggestion that this method could be used for providing off-site access, with on-site access offering more options.

### **Case Study C: Levels of curation**

Participants explored three series of records: the supplementary digital material in MS.9225 David Beales, *The Road to the Asylum* and two series from the Oliver Wrong archive.<sup>4</sup> All three series were copied onto a laptop and accessed within the Windows 7 environment. A browser displaying the catalogue records was also open on the laptop.

MS.9225 comprises over 3,000 files in a range of formats held together in a single folder. Original filenames have been retained. One series of Wrong files were posthumously migrated from WordPerfect to Word by Wrong's family and renamed using the following convention: "Description of text, date, subject matter". The other series remains in the WordPerfect format with Wrong's original filenames.

Participants were asked to consider the following questions:

- Is there enough to work with?
- Is additional archival curation necessary (or desirable)?

As with case study A, a lot of the discussion revolved around navigation problems and the difficulty in finding information of interest. Many participants found it hard to know where to begin and found original filenames hard to interpret, though there remained recognition that original filenames and directory structures have research value. A popular suggestion was to have multiple presentations of the records, allowing researchers to see the original structure and filenames but also explore curated versions designed with specific research interests in mind.

One important point raised was that the debated around original versus curated filenames was not just about the user experience when viewing the records. It is also a concern much earlier in the research process because researchers need to be able to identify records that are worth requesting and visiting Wellcome to view. If original filenames are so opaque as to be a barrier to this, then questions around curation and access are moot as there will not be any researchers accessing the records.

### **Case Study D: Value of the physical media**

Participants examined a data CD from the John Sulston archive and a DVD output from one of Wellcome's public engagement grants.<sup>5</sup> The CD contains records concerning a trip to Japan and originally sat within a paper file on the same subject. The DVD grant output was deposited in the corporate archive alongside a musical programme, lyric book and evaluation report.

The CD records were copied onto a laptop running Windows 7 and the DVD was playable using the laptop's DVD drive. A browser displaying the catalogue records was also open. The related paper files and the physical media were next to the laptop for the participants to view.

Participants were asked to consider the following questions:

- Do the physical media hold any value?
- How well are the paper record, physical media, digital records and online catalogue connected?

<sup>4</sup> David Beales, *The Road to the Asylum* supplementary digital material, catalogue ref. [MS.9225/2](#) and Professor Oliver Wrong: Wrong digital material, catalogue ref. [PP/WRO/B/11](#), held at Wellcome Collection

<sup>5</sup> John Sulston archive: The Secrets of John Sulston Jul-Aug 2002 Hiroshima/Tokyo, catalogue ref. [PP/SUL/C/3/6/9/1](#) and Wellcome Trust Corporate Archive: 097573/Z/11/Z Epidemic, catalogue ref. [WT/C/6/1/58](#), held at Wellcome Collection

The consensus amongst the participants was that the physical media can contain extra information that is often missing from catalogue descriptions and can be used in areas of research that are not always covered by catalogue descriptions, such as graphic design. It was felt that photographs of media were a good substitute to retaining the physical item in most, though not all, cases. However, if photographing the media was adopted careful thought would need to be given to the location of the photograph in relation to the catalogue record to ensure it was correctly understood.

# Conclusion

Following the workshop the conveners met to discuss what we had observed at each case study. Whilst recognising that our findings are specific to the group of participants and we should be wary of extrapolating too far, we noticed some common themes.

Firstly, directory structures and filenames are data and they have recognised value as a way of understanding the creator's logic and thought processes. Nevertheless, there was an expectation that the records would undergo some level of curation by the archivist. Participants wanted multiple presentations alongside the original, based on clusters of themes, file formats and file sizes. They also wanted a better overview at various levels in the directory. To some extent what they were requesting, particularly regarding overviews and context, already exist in the catalogue. But the fact the participants did not find this information easily (or at all) suggests there are improvements to be made in the catalogue.

A key finding from the workshop was that many researchers have limited experience of using born-digital archives in their research and so can only speculate about how they would like to access and use such records. We need researchers to start using born-digital archives for them to know and be able to articulate their requirements, but in large part limited use is due to archives either not being available or not being discoverable, meaning researchers do not know the records exist. The Collections Information Team at Wellcome can immediately start to address this latter issue and attempt to increase researcher use of our born-digital archives.

We are already aware of some of the limitations of our online catalogues, but the workshop made very clear the fact that whilst a lot of the information about the content and context of digital records is captured in the catalogue, it is not easy to find. Therefore, we are taking practical steps to improve the cataloguing process and enable more researchers to find the digital records that are relevant and useful to their research. The hope is that increased use of our born-digital archives will firstly enable us to continue developing our understanding of user requirements and secondly show there is demand for a more scalable access method than we currently employ.

Firstly, we are changing which catalogue records we make visible online. Wellcome currently makes visible all the catalogue records created for a set of born-digital files, from the highest level of the directory structure right down to the individual digital files. These catalogue records are automatically created as part of the digital preservation process, though as previously mentioned most of them not enhanced. The thinking up to now has been that since they exist we may as well make them visible online as they might prove useful to some researchers. However, the workshop suggested they may actually be a hindrance. In the days of paper catalogues a researcher would start at the top of the entire archive and drill down into the various sections. They would reach the catalogue record describing a set of files before reaching the descriptions of each individual file. With online keyword searching, researchers can now enter a catalogue at any level in this hierarchy and so could view the catalogue record for an individual digital file (which has no information beyond a filename and last modified date) before viewing the catalogue record for the directory which provides a description of the contents of that directory and other important contextual information. Given that access is provided to sets of files (the deliverable unit), not individual ones, it appears that making visible catalogue records for the digital files creates an unhelpful distraction from the enhanced catalogue records that provide the information researchers are looking for. Consequently, we are going to trial only making visible catalogue records down to deliverable-unit level; that is the catalogue record that describes a discrete set of records, be that a floppy disk or a folder on a hard drive. The descriptions for these records will be more detailed than previous and if appropriate will list the filenames of the files contained within that disk or folder.

One other alteration we are making is to the extent field. Here we will list the total number of files broken down into format types rather just the number of files. For example, “10 document files, 15 audio files, 20 image files” instead of “45 digital files”.

Finally, we are changing the access condition wording in each catalogue record. Previously it stated that born-digital records were not available to researchers. This is being modified to state that people can contact the Collections Information Team to request access to born-digital records. We hope this change in particular will encourage researchers to start using our born-digital collections.

The workshop brought together a mix of experienced researchers, archivists, and digital developers to come together to share experiences and further understand each other’s requirements and expectations regarding born digital archives. One of the main aims of this workshop was to place user expectations (in this case academic researchers with varying degrees of experience in using born digital records) at the centre of the development of access to born digital archives at Wellcome Collection. By using practical exercises, we were able to appreciate the various forms that access could take, while also encouraging those researchers to think of born digital archives as relevant, and potentially integral, to their research.

Changes to how born digital archives are catalogued and described at Wellcome Collection are an immediate outcome from the workshop, and will hopefully result in increased requests to use this material. Other actions will take longer to develop and build, but staff tasked with designing and improving Wellcome Collection’s catalogues and collection discovery platforms now have a clearer understanding of researcher needs and expectations regarding born digital records. Furthermore, the workshop attendees constitute a potential user group that could be involved in the development of these platforms, whether through testing or future development workshops, as Wellcome Collection looks to remove barriers to access to born digital archives.



The first part of the document discusses the importance of maintaining accurate records of all transactions. This includes not only sales and purchases but also any other financial activities that may occur. It is essential to ensure that all entries are properly documented and supported by appropriate evidence.

In addition, the document emphasizes the need for regular reconciliation of accounts. This process involves comparing the company's internal records with the bank statements to identify any discrepancies. By doing so, the company can ensure that its financial statements are accurate and reliable.

Another key aspect of financial management is the timely payment of bills and invoices. Failure to do so can result in late fees, damaged relationships with suppliers, and potential legal action. Therefore, it is crucial to establish a system for tracking and paying all obligations on time.

Finally, the document highlights the importance of budgeting and financial forecasting. By creating a budget, the company can set financial goals and monitor its performance against those goals. Forecasting allows the company to anticipate future financial needs and make informed decisions about investments and capital expenditures.

**March 2018**

**Wellcome exists to improve health for everyone by helping great ideas to thrive. We're a global charitable foundation, both politically and financially independent. We support scientists and researchers, take on big problems, fuel imaginations and spark debate.**

**Wellcome Trust, 215 Euston Road,  
London NW1 2BE, UK  
T +44 (0)20 7611 8888, F +44 (0)20 7611 8545,  
E [contact@wellcome.ac.uk](mailto:contact@wellcome.ac.uk), [wellcome.ac.uk](http://wellcome.ac.uk)**

The Wellcome Trust is a charity registered in England and Wales, no. 210183. Its sole trustee is The Wellcome Trust Limited, a company registered in England and Wales, no. 2711000 (whose registered office is at 215 Euston Road, London NW1 2BE, UK).