

Digital Preservation at Oxford and Cambridge

A collaborative research project to evaluate and provide sustainable recommendations for our digital preservation programmes

DPOC visits the Wellcome Library in London

Posted on [24 January, 2017](#) by [ehalvarsson](#)

A brief summary by Edith Halvarsson, Policy and Planning Fellow at the Bodleian Libraries, of the DPOC project's recent visit to the Wellcome Library.

Last Friday the Polonsky Fellows had the pleasure of spending a day with Rioghnach Ahern and David Thompson at the [Wellcome Library](#). With a collection of over 28.6 million digitized images, the Wellcome is a great source of knowledge and experience in working with digitisation at a large scale. Themes of the day centred around pragmatic choices, achieving consistency across time and scale, and horizon scanning for emerging trends.

The morning started with an induction from Christy Henshaw, the Wellcome's Digital Production Manager. We discussed digitisation collection development and Jpeg2000 profiles, but also future directions for the library's digitised collection. One point which particularly stood out to me, was changes in user requirements around delivery of digitised collections. The Wellcome has found that researchers are increasingly requesting delivery of material for "use as data". *(As a side note: this is something which the Bodleian Libraries have previously explored in their [Blockbooks](#) project, which used facial recognition algorithms traditionally associated with security systems, to trace provenance of dispersed manuscripts).* As the possibilities for large scale analysis using these types of algorithms multiply, the Wellcome is considering how delivery will need to change to accommodate new scholarly research methods.



- Brain teaser: Spot the Odd One Out (or is it a trick question?). Image credit: Somaya Langley

Following Christy's talk we were given a tour of the digitization studios by Laurie Auchterlonie. Laurie was in the process of digitising recipe books for the [Wellcome Library's Recipe Book Project](#). He told us about some less appetising recipes from the collection (such as three-headed pig soup, and puppy dishes) and about the practical issues of photographing content in a studio located on top of one of the busiest underground lines in London!

After lunch with David and Rioghnach at the staff café, we spent the rest of the afternoon looking at Goobi plug-ins, Preservica and the Wellcome's hybrid-cloud storage model. Despite talking digitisation – metadata was a reoccurring topic in several of the presentations. Descriptive metadata is particularly challenging to manage as it tends to be a work in progress – always possible to improve and correct. This creates a tension between curators and cataloguers doing their work, and the inclination to store metadata together with digital objects in preservation systems to avoid orphaning files. Wellcome's solution has been to articulate their three core cataloguing systems as the canonical bibliographic source, while allowing potentially out of data metadata to travel with objects in both Goobi and Preservica for in-house use only. As long as there is clarity around which is the canonical metadata record, these inconsistencies are not problematic to the library. 'You would be surprised how many institutions have not made a decision around which their definitive bibliographic records is', says David.

