

# Parametric Curation in Digital Archives: Concept and Potential Benefits

CHRISTOPHER A. LEE and JEREMY LEIGHTON JOHN and KAM WOODS

**Abstract:** It is now common practice for professional photographers to use “raw” formats to store digital photographs. Previously, photographers took master digital photographs such as TIFF files and produced numerous variant files as they experimented with editing options (e.g. white balance, cropping, tweaked color profiles). With the emergence of raw photography, large photographic files are edited and handled through the recording of metadata that represent information about changes made to a single master file, thereby avoiding retention of many large edited variant files.

Underlying this technology is the concept and process of “parametric image editing” (PIE), which is geared toward nondestructive and reversible editing of digital photographs. This avoids excessive and inefficient management and storage of files. If an earlier version of a file is required, it is simply restored using the metadata. Such information may be explicitly available to the user or it may be held within the software.

Parametric editing can serve as a model for digital curation. “Parametric curation” is an approach that uses metadata about changes made rather than unduly replicating identical or similar information. One particular approach is based on the capture and retention of forensic disk images, which provide exact copies of all storage sectors on storage media. The files within the disk image may be retained in situ for future export as and when necessary. We will discuss the rationale for parametric curation in archival collections, as well as various practical implications for archival workflows, including the use of Digital Forensics XML (DFXML).

## **About the author:**

*Christopher (Cal) Lee* is Associate Professor at the School of Information and Library Science at the University of North Carolina, Chapel Hill. He teaches courses on archival administration; records management; digital curation; understanding information technology for managing digital collections; and acquiring information from digital storage media. He is a lead organizer and instructor for the DigCCurr Professional Institute, a week-long continuing education workshop on digital curation, and he teaches professional workshops on the application of digital forensics methods and principles to digital acquisitions.

Cal's primary area of research is the long-term curation of digital collections. He is particularly interested in the professionalization of this work and the diffusion of existing tools and methods into professional practice. Cal developed "A Framework for Contextual Information in Digital Collections" (*Journal of Documentation*), and edited and provided several chapters to *I, Digital: Personal Collections in the Digital Era* published by the Society of American Archivists.

Cal is Principal Investigator of the BitCurator project, which is developing and disseminating open-source digital forensics tools for use by archivists and librarians. He was also Principal Investigator of the Digital Acquisition Learning Laboratory (DALL) project, which investigated and tested the incorporation of digital forensics tools and methods into digital curation education. Cal has served as Co-PI on several projects focused on preparing professionals for digital curation responsibilities: Preserving Access to Our Digital Future: Building an International Digital Curation Curriculum (DigCCurr), DigCCurr II: Extending an International Digital Curation Curriculum to Doctoral Students and Practitioners; Educating Stewards of Public Information for the 21st Century (ESOPI-21), Educating Stewards of the Public Information Infrastructure (ESOPI2), and Closing the Digital Curation Gap (CDCG). In a project called Curation of a Forensic Data Collection for Education, Cal investigated and developed resources to enhance access and use of disk images to support digital forensics education.

**Dr Jeremy Leighton John** belongs to the Digital Research and Curator Team in the Department of Digital Scholarship at the British Library where he has been Curator of eMANUSCRIPTS at the British Library since 2003, having been Specialist Scientific Curator from 2000. Previously he worked as a cataloguer of bioacoustic collections. In 1996 he completed a DPhil in Zoology at Merton College, University of Oxford, concentrating on evolutionary and phylogenetic topics. He is a Fellow of the Linnean Society of London and the Royal Geographical Society.

As Principal Investigator of the Digital Lives Research Project funded by the UK Arts & Humanities Research Council, he promoted the use of digital forensics with an article in the journal *Nature*: 'The future of saving our past'. In 2012, as Project Lead for the Personal Digital Manuscripts Project at the British Library, he published a technology watch paper for the Digital Preservation Coalition entitled 'Digital forensics and preservation'.

Presently serving on the Committee of the Section for Archives and Technology of the Archives and Records Association of UK & Ireland. Memberships have included the Library Committee of the Royal Society and the Advisory Committee of the National Cataloguing Unit for the Archives of Contemporary Scientists.

**Kam Woods** is a Postdoctoral Research Associate in the School of Information and Library Science at the University of North Carolina at Chapel Hill. He is currently

Technical Lead on the BitCurator project, and works with Dr. Cal Lee developing techniques and tools to assist in long-term archiving of born-digital data.

Kam's research focuses on long-term preservation of born-digital materials. He is interested in interdisciplinary approaches that combine technologies and expertise in the areas of archiving, computer science, and digital forensics for the purpose of enabling and maintaining access to digital objects that are at risk due to obsolescence. Prior to his current work at UNC, Kam worked with Cal Lee on the development of educational materials to support the use of realistic forensic datasets in professional training and to identify and explore novel uses of forensic data and tools in the context of digital archives.