

Digital Dunhuang: A Standard for Digital Preservation

Peter Zhou
University of California, Berkeley

ABSTRACT

The Digital Dunhuang project is enabling long-term preservation of cultural heritage of inestimable value, while providing a platform for sharing all digital assets generated in the act of preservation. This presentation will examine the major aspects of digital asset management and digital preservation implemented in this ongoing project.

KEYWORDS

Digital preservation, digital asset management, Digital Dunhuang, the preservation of world heritage

1 INTRODUCTION

The Mogao Caves at Dunhuang, a designated UNESCO World Heritage site, are a splendid treasure house of art from Ancient China. [1–2] The Dunhuang Academy has been devoted to the protection of the Mogao Caves since its founding in 1944 (originally called the Dunhuang Art Research Institute). The protection measures include slowing down the degeneration of murals with a variety of protecting techniques, as well as looking for digital tools to preserve the caves since 1990s. The Academy is currently undertaking the Digital Dunhuang project with cooperation from home and abroad. Work is being done to digitize three-dimensional structures of murals, sculptures and all the 492 caves, and also to construct a digital database of the Mogao Caves resources. The size and scope of the project is quite significant. Some exemplary statistics include:[3]

- Over 100 caves already photographed
- Over 14,000 square meters of murals captured
- More than 70 terabytes of data
- Over 87 caves in complete QuickTime Virtual Reality
- Over 20 years of climate monitoring data for 87 caves

The project has called on the expertise of a number of institutions over the years, including The Dunhuang Academy, Zhejiang University, Wuhan University, the Getty Conservation Institute, and Northwestern University. It is designed to incorporate a wide range of data types, including archeological and conservation data and files, documentation for a large number of artifacts, Dunhuang manuscripts, and scholarly publications.



Figure 1: Digitizing the Dunhuang Caves. Image courtesy of the Dunhuang Foundation.

2 FROM DIGITAL ASSET MANAGEMENT TO DIGITAL PRESERVATION

With the support of The Mellon Foundation, the Dunhuang Academy has been exploring building a permanent repository of all digital assets of Digital Dunhuang. It has also made persistent efforts to photograph the caves while conducting archeological study and conservation work. The Academy has made huge progress and accumulated large amounts of data in both analog and digital formats. The consensus seems to be: if the information gathered from the Mogao Caves is to be permanently preserved for future generations, the only way to ensure this is to integrate all the content that has been created in the past, is being created now, and will be created in the future into a large digital repository. [4] This digital repository can facilitate perpetual preservation, effective digital asset management operations, and easy access in a systematic way through the use of high technology delivery mechanisms. This presentation discusses major aspects of the design and development of this repository system. The ultimate goal is to store the totality as well as minutia of the Dunhuang Caves' digital content, preserve those digital surrogates, reconstructions, files, and digital assets perpetually, and provide a research platform for the study of mural art, history, geography, religion, economics, politics, ethnology, language, literature, science and technology in ancient China and Central Asia.



Figure 2: Dunhuang scientists preparing to digitally scan cave for 3D modelling. Image courtesy of the Dunhuang Academy.

There is a sense of urgency. More than half a million visitors come here every year. The historical and cultural significance of the caves, as well as their vulnerability- their age, instability of the terrain, and fluctuations of humidity caused by increased human presence- all add to the sense of urgency. The digital capture and restoration of Dunhuang will most likely to be the last records of today's Dunhuang.

2.1 Functional Components [5]

DAM: Facilitates asset creation, cataloging; image, video, and text management and delivery and version control. Tracks digital preservation actions. Pushes metadata and content to the Digital Dunhuang platform. Manages master high resolution files and original documents.

DIGITAL PRESERVATION: Managed digital preservation actions include creating checksums, validating files, and extracting technical metadata upon ingest; monitoring file format obsolescence; migrating file formats as well as content; tracking and copying files to LTO tapes.

PUBLISHING AND SHARING: Surrogates of master assets managed in the DAM will be pushed for external delivery.

2.2 Content Categories

Stitched/composite cave images, raw cave images, cave QTVRs, historical photos, videos, digital restorations, manuscripts from Cave 17, artifacts (approximately 10,000 objects), reproductions (copies) of images in caves, (digitized) microfilm of manuscripts, interactive panoramic of caves; research created by members of the Dunhuang Academy, scholarly publications; previously published bibliographies, indices, glossaries, and finding aids;

conservation data and materials, archaeological reports and drawings, CAD drawings, and 3D laser point cloud data

2.3 File Formats

TIFF, JPEG, JPEG2000 (still image), PSD, BMP, PSB (Photoshop large file format), CR2 (Canon raw format), DCR (Kodak raw format), DNG (Adobe/universal raw format), other RAW camera formats (list), CDR (Corel Draw), CAD, PTX (original 3D cloud points), DGN (Microstation Design File), PDF, CAJ, MOV (QTVR), MPEG2/35 Mbps (AVI wrapper), HD video files (format TBD), DPG (Nintendo video file format), Word, Excel, txt, MPEG4, H.264, FLV (Flash)

2.4 Cross-linking

All content (images, documents, etc.) must be displayed with metadata. For example, an initial search result might display thumbnails with basic data. The user can then select to see a large 3D image and fuller data.



Figure 3: An example of 3D rendering of cave interior paintings. Image courtesy of the Dunhuang Academy.

3 RESULTS AND DISCUSSION

Digital Dunhuang is an enterprise project. It enables us to explore ways to design and develop sophisticated databases, and perpetually preserve massive amounts of data. By extension, Digital Dunhuang has also shed light on what we need to do to preserve digital assets in libraries and museums around the globe.

ACKNOWLEDGMENTS

The work discussed was sponsored by the Dunhuang Academy, with the support from the Mellon Foundation. This presentation draws on a joint report by Linda Tadic and the author on the functional requirements of Digital Dunhuang in 2012; and on a panel discussion at the DH 2015 conference in Sydney, at which the author discussed this project with other panelists - Wang Xudong and Wu Jian of the Dunhuang Academy, and J. Stephen Downie of the University of Illinois. Their contributions are greatly acknowledged.

REFERENCES

- [1] Fan Jinshi and Wu Jian. 2004. *Dunhuang Grottoes*. Beijing: China Travel and Tourism Press.
- [2] Fan Jinshi and Liu Yongzeng. 2007. *Appreciation of Dunhuang Grottoes*. Nanjing: Jiangsu Fine Arts Publishing House.
- [3] Xudong Wang, Peter Zhou, Eugene Wang, and J. Stephen Downie. 2015. "Digital Dunhuang: Enhancing Virtual Exploration of the Real Dunhuang." Unpublished panel proposal for DH2015 conference, Sydney.
- [4] Peter Zhou. 2011. "The Development of a Repository System of Dunhuang Murals and Digital Resources for Dunhuang Studies." *Proceedings of the International Conference on Cultural Heritage and Digitization*, 317–23. Dunhuang: Dunhuang Academy.
- [5] Linda Tadic and Peter Zhou. 2012. *Digital Dunhuang: A Repository System of Dunhuang Murals and Digital Resources: Functional Requirements*. Dunhuang: Dunhuang Academy.