

CDL Digital Preservation Program Curation Services

About the Digital Preservation Program

The CDL Digital Preservation Program was established in 2002 to ensure long-term access to the digital information that supports and results from research, teaching, and learning at the University of California. This digital information represents the intellectual capital of the University and needs to be managed carefully to ensure it will remain available for use by future scholars.

The Problem: One-Size-Fits-All Doesn't Fit

To support this mission, the Digital Preservation Repository (DPR) was created to provide a shared solution for the UC community to curate – that is, preserve, manage, and disseminate – their digital collections efficiently and effectively. More than 50 terabytes of content have been contributed to the repository by campus partners, including important holdings of digitized newspapers and texts, cultural heritage and scientific imagery, sound and moving image recordings, and web resources. Ongoing projects will add significant new content in existing collection categories as well as new areas such as scientific data. Facing this diversity of information types and uses, the “one-size-fits-all” approach of the current-generation repository falls short in fully meeting the wide range of preservation needs of campus partners and scholars.

The Solution: Digital Curation Services

Rather than providing curation function through a monolithic system like the current DPR, we envision an evolving suite of granular, decentralized components providing highly available, comprehensive, and sustainable curation services for access to and use of authentic digital assets over time. These curation services will better enable us to remain responsive to the ever expanding needs of our University stakeholders, and to collaborate more effectively with the broader curation community.

Overview of the Digital Curation Infrastructure

In order to accommodate an increasing diversity of content types and uses, and to respond appropriately to rapid technological change, the digital curation infrastructure is designed to be inherently flexible, responsive, and deployable in contexts beyond traditional library workflows.

Our approach to the curation infrastructure is to create a set of independent but interoperable services that will be applicable throughout the full digital lifecycle and will not depend on sophisticated computing environments and skills. Since the individual services are small and self-contained, they are easier to develop, maintain, and, when necessary, replace. Although the functional scope of any particular service is narrow, the services can be freely combined to support complex curation functions.

Digital Curation Services

The range of digital curation services provides the following:

- ***Safety through distributed replication***, ensuring the reliability and accessibility of curated content and the availability of services built around that content.
- ***Meaning through interpretive description***, allowing content to be exposed to user communities in appropriate contexts.
- ***Usefulness through services***, facilitating the widespread integration of curated content into the discourse of the University.
- ***Value through use***, enriching curated content through the multiplier effect of its creative use and re-use.

How Digital Curation Services Can Be Used

This approach to digital curation services promotes the idea that curation is an outcome, not a place; the services are therefore applicable for non-traditional uses and contexts. For example, high value digital assets on scholars' desktop computers would benefit from such things as persistent identification or regular audits to discover and repair bit-level damage – functions usually available only in the context of a “preservation system” but now easily applied to content where it most usefully resides.

Other ways the curation services can be used:

- A ***faculty member*** may create course materials (such as audio, video, and text) that are posted online. The faculty member could capture these materials using the Web Archiving Service (WAS) to preserve them for future re-use or sharing with departmental colleagues. Comprehensive training materials for the CDL-managed WAS are available.
- A ***library, museum, or archive*** may have a collection of unique content (such as specimens, oral history recordings, or photographs) that they would like to preserve and make available online. IT staff could deploy local instances of the ingest, indexing, and search services to expose this content to a new, worldwide audience. All curation services are packaged in a self-contained form that can be installed and operated with minimal technical assistance.
- A ***researcher*** who is the recipient of a grant may have a variety of curation needs. For example, the funding agency may require that the researcher's data (such as observational data, experimental results, and code books) be kept for a certain number of years. To preserve the data, the researcher could use the identity, storage, and replication services. Or, the researcher may want to collaborate with different institutions in the course of his or her work, making use of the publication and annotation services to share data. These services can be deployed in the ORU's local computational environment, or the researcher can make use of centrally-managed services operated by the CDL.

Contact

Patricia Cruse <patricia.cruse@ucop.edu>
Director, Digital Preservation Program