Repository History: Tufts University
Anne Sauer, Director, Digital Collections and Archives
Deborah Kaplan, Digital Resources Archivist, Digital Collections and Archives

Introduction
The Tufts Digital Library (TDL) and Tufts Digital Repository (TDR) have been in development since 2003 and in production since August of 2005. A layered service, the TDR provides the underlying object management for resources that are both open access and restricted access. The Tufts Digital Library is the public access point for open content contained in the TDR, and provides basic services for discovery and retrieval of digital resources.

Organisational Context
The TDR and TDL are jointly administered by two units at the university. Digital Collections and Archives is a central university library reporting to the Provost, and is charged with providing stewardship of the university’s permanent records and collections in all formats. Academic Technology is a unit of University Information Technology and reports to the Chief Information Officer.

DCA provides leadership on collection management, metadata, and digital preservation issues. AT provides application management and development, user interface design, and systems architecture. Both groups share responsibility for assessing and interpreting user needs.

Mission Statement
In service of Tufts University's teaching and research mission, the DCA is the steward of the University's permanently valuable records and collections created in any format, ensuring their permanent preservation and accessibility. Academic Technology is a university-wide resource that provides guidance and professional advice on using technology in support of teaching and research initiatives. The TDR serves both of these missions by providing services for storage and management of digital assets of enduring value, including university records, manuscript collections, and faculty publications. The TDL provides access and discovery for those digital assets which are made accessible to the Tufts community and the wider scholarly world.

Building a Business Case
We have been successful in building a business case for our service in two areas:
• Bringing security and management to digital assets resulting from faculty research.
• Providing a place for management and preservation activities for electronic records.
We have built this business case with the administration around best practices and new developments in academia; use statistics of currently available resources; changing laws;
and communication about the services we are capable of providing. We have worked with the university to mandate a formal University Records Policy whose implementation we are in the process of integrating into the TDR.

### Current Repository Contents

Our repository currently contains more than 75,000 objects including images, xml texts, pdfs, datasets in xml, collection objects, and oral histories (comprising xml text, audio, and SMIL datastreams). The provenance of these objects is widely varied, but includes university archival photographs; EAD-encoded finding aids describing our manuscript and archival collections; graduate theses; faculty papers; university newspaper backfiles; datasets from special projects; and many others.

### Current Deposit Activity

Currently, the majority of our deposit activity comes from one of two sources:

- The constant and steady pipeline of digitizing DCA-managed materials.
- Special projects, funded to digitize particular collections held by Tufts libraries, faculty, and affiliated institutions.

We expect the NIH-mandate for faculty self deposit will drive up our deposit activity of faculty publications.

### Faculty Engagement

Faculty were initially not attuned to the need for a repository at the university. Our initial conversations with potential early adopters revealed a nearly universal lack of understanding of digital preservation needs and almost no interest in depositing publications in a repository – even if the publication agreements they had signed over the years would have made this possible. We therefore started by questioning faculty about the existing management of digital assets resulting from their research. While most were unfamiliar with the challenges posed by digital preservation, the realities of hardware failure, software obsolescence, and for some, personal knowledge of data loss due to these issues, made an excellent context for discussing the role of a repository at the university.

We began to work with a few early adopters as we scaled up workflows for preparation, ingest, and delivery from the repository. Using collections from the DCA as a testbed, and working with a handful of faculty with a variety of object types, we began to build a set of services to support our vision. We continue to do so, and anticipate making a general call for deposits in the next fiscal year.

We both expect and hope that the NIH mandate, and other open access initiatives, will drive a demand and a need for an increased integration between the TDL and faculty research.

### Policy Formulation

The DCA has been proactive on forming policies to support the TDR and TDL, and on gaining institutional support for their implementation. DCA has just received, from the
NHPRC Electronic Records Program Expansion Project grant program, funding for TAPER: the Tufts Accessioning Program for Electronic Records. TAPER is designed to implement a machine-readable format for electronic record submission agreements.

**Hosting and support**

Hosting and support for the systems used by the repository is provided by the University Systems Group. Support for Fedora and related applications is managed by AT.

**Service Sustainability**

We have been very successful in gaining institutional support for the notion of sustainable preservation services. However, we have faced challenges in clearly delineating what it means to create a sustainable service. Many of our data objects were ingested into the TDR with special projects funding, and such funding has not always accounted for all of the administrative and technological maintenance required to keep sometimes-unique object types available online in perpetuity, upgraded as necessary with the repository and its supporting hardware and software. Particularly, we are in constant negotiation between the archival notion of perpetual access and the modern technological reality of unavoidable flux.

We are constantly reevaluating what it means to create sustainable digital preservation, and future grants and projects are written with lessons learned in mind.

**Major Achievements**

While we are proud to have a service that has been in production for several years, with a steadily growing collection of wonderful digital resources, our major achievements – from a manager’s perspective – lie very much behind the scenes.

Institutional Support: We have worked very hard to articulate the significance, importance, and value of our repository for both research and administration to the university leadership. Our core program has been entirely funded by the university, with grant funding used judiciously for the creation of special digital collections, and building particular tools.

Partnership between DCA and AT: The partnership that has evolved between AT and DCA is a tremendous asset. The decision to work together was not the result of a mandate from above, but arose from allied interests in fostering development of innovative approaches to teaching and research at the university through digital resources. The respective strengths and perspectives of both groups have made our program stronger. Coordination is not always easy between two groups that have different missions, reporting structures, and skills, but the collaboration has broadened both of our perspectives, and the program is the better for it.
Important Unresolved Issues

Some of the big challenges we face are probably common ones. We struggle to secure funding for ongoing capital purchases for servers and storage. Our vision for what the repository could be must always be tempered by the limits of our staffing levels.

As with many repositories, we confront something of a timing dilemma when we consider how best to advertise our services to the university at large: we have to keep advertising to keep the TDR and TDL relevant, but if we advertise too successfully too quickly, there may be an influx of demand we are unable to meet in a timely manner.

We would like to take advantage of new technologies and generous grants to design compelling interfaces to our objects, particularly special projects; to ingest new and interesting object types into the TDR; and to write automation scripts to streamline our daily processes. However, we need to do this without creating technological dependencies we don't have the staff to maintain in-house.

We need to migrate old objects which were brought into the TDR during past grant funded projects, and which need correction, modification, or modernization to be fully supported. Staying modern and relevant without leaving behind older material is a constant dilemma.