

**Digital Production and Integration Program (DPIP)
Production and Content Integration Working Group**

Final Report

**Yale University Library
August 2006**

Frederick Martz, Chair
Katie Bauer (Assessment/Usability)
Matthew Beacom (Metadata)
Tracy Bergstrom (Visual Resources Collection)
John Gallagher (Medical Library)
Rebekah Irwin (Liaison to Metadata and Digital Preservation Committees)
Brian Kupiec (Beinecke Digital Studio)
Derek Merleaux (MSSA)
Tobin Nellhaus (RSC)
George Ouellette (ILTS)
Karen Reardon (ILTS)
Martha Smalley (Divinity)
Lisa Thomas (Access Services) October-November 2005 Only
David Walls (Preservation)
Jen Weintraub (E-Collections and DCF)
Meg Bellinger (Sponsor)

Table of Contents

Executive Summary

1.0 Needs Assessment and Environmental Scan

- 1.1 Trends in Academic Teaching and Research and in Libraries**
- 1.2 Peer Institution Programs**
- 1.3 Yale University Needs Assessment**

2.0 The Digital Production and Integration Program

- 2.1 DPIP Will Consist of Three Interdependent Services**
- 2.2 Audience for Services**
- 2.3 Organizational Structure**
- 2.4 DPIP Staff**
- 2.5 Communication Plan**
- 2.6 Library / ITS Division of Labor**
- 2.7 Technical Infrastructure and Support**

3.0 DPIP Initiatives

- 3.1 Digital Projects**
- 3.2 User Studies and Assessment**
- 3.3 Services to Students and Faculty**
- 3.4 Recommendations for Future Initiatives**

4.0 Digitization Resources at Yale University Library and Yale University

5.0 Implementation Timetable

See separate Excel spreadsheet for a detailed DPIP Task List.

6.0 Budget

6.1 Budget Utilizing Existing Resources

See separate Excel spreadsheet for the DPIP Budget Utilizing Existing Resources.

6.2 Budget Necessary to Achieve an Ambitious Goal

See separate Excel spreadsheet for the DPIP Budget Necessary to Achieve a Digital Collections Strategy Appropriate for the Yale Library.

Appendix A: DPIP Virtual Organization (Diagram)

Committee Web Site: <http://www.library.yale.edu/iac/dpip/>

Executive Summary

There is widespread recognition that Yale University Library must take more aggressive action to build and deliver digital collections for our local and global community of users. Many staff believe that we are not meeting digital expectations for the second largest academic library in the country and that instead of viewing the library as the place to go for authoritative research and information management assistance, our readers will ultimately go elsewhere (See section 1.0 for elaboration). In order to develop recommendations for action, a plan was established in 2005 to review service needs and gaps for digitization services in the library, and to investigate existing services on campus. This report is the result of that analysis and it presents plans and priorities for development, coordination or implementation of needed services, as well as associated policies and guidelines for best practices.

The Digital Production and Integration Program (DPIP) concept is an Integrated Access Council (IAC) initiative and has gradually evolved since late 2004 to include three major components:

1. Market / User Research Services (including assessment and usability efforts).
2. Digital Production and Content Integration Services.
3. Consultation, Advisory, Referral and Management Services.

In October 2005, with a primary focus on the second of these three components and with LMT overview, the DPIP Production and Content Integration Working Group was charged to analyze needs for digitization services in the library (e.g. scanning, text markup, metadata creation), develop plans and priorities for implementation of targeted services, and in the course of the coming year introduce selected services along with associated policies and guidelines for best practices. The primary goal was to identify important unmet needs, with a focus on practical, realistic solutions.

The working group's activities included:

- Widespread consultation with stakeholders.
- Compilation of Digitization Needs and Challenges.
- Site visits to Harvard and Cornell digital production / service facilities.
- Roundtable discussion of issues and concerns among working group members.
- Prioritization of issues and determination that some issues were out of DPIP's scope.
- 1-2 page briefs on the seven highest-priority topics giving a statement of needs, challenges, and possible solutions, including the logical role for DPIP. The seven topics were: digital collection development policies, support for non-arts images and full text, digitization workflow, metadata production / standards, division of labor between the library and ITS, cross-collection searching / global integration of resources, and e-reserves.
- Iterative review of draft versions of the final report.
- Discussion of and planning for the several proposed initial projects, especially the e-reserves prototype.

The documents mentioned above, including the charge, the needs compilation, and the seven preliminary reports, along with a glossary of terms, are available on the committee web site at: <http://www.library.yale.edu/iac/dpip/>.

In the course of its review, the working group identified six key problems that must be resolved in order to achieve a coherent, well-funded and large-scale digitization program in the Yale University Library:

- 1) Numerous unrelated digital initiatives.
- 2) A need for well-communicated standards and recommended practices.
- 3) No clear and comprehensive pathway for Yale faculty members who want to create digital teaching resources using YUL material or services including stewardship, or for staff seeking to create digital library collections.
- 4) Limited mechanisms for understanding the requirements of faculty and responding effectively to expressed desires.
- 5) Inadequate repository infrastructure to support a wide range of digital content and a versatile set of discovery and delivery services.
- 6) Ad hoc selection mechanisms for digitization.

(See section 1.3.1 for a full explanation of these challenges).

After consultation with stakeholders throughout the library organization, examination of input from users and units across the university, and a thorough review of possible actions, the working group concluded that a DPIP organization consisting of a small core staff, a larger set of closely allied virtual participants drawn from existing units, and partnerships with other well established campus initiatives can achieve substantial progress at Yale even in an environment of limited financial and staff resources (See Appendix A for a diagram of the DPIP Virtual Organization).

DPIP can offer a rational and efficient approach to digital collection building for units and individuals who want guidance, by providing a consistent point of contact and service referral for staff and faculty beginning digitization and metadata production activities, and a reliable source for the information needed to implement, optimize, and manage digital projects, along with continuity in production services and consistent application of current standards and best practices.

This report operates at two distinct levels. At one level, in order to establish context, the report documents views of stakeholders regarding the state of YUL's digital library program and makes several recommendations about what the library as a whole needs to accomplish over the next five to ten years. These views were elicited through a series of interviews and working group discussions. At another level of detail, the report presents specific recommendations that YUL should undertake in order to develop a sustainable, coherent digital production program and defines the contribution that a DPIP organization can make in transforming the current processes through which several units of the library address the many challenges of digitization. While the report discusses a very broad range of library digital initiatives, it is important to recognize that DPIP is designed to address only a subset of these and that many other units will deliver critical components of a comprehensive digital library program at Yale.

1.0 Needs Assessment and Environmental Scan

1.1 Trends in Academic Teaching and Research and in Libraries

The digitization of scholarly primary and secondary research materials has made it “possible for individual scholars to do things not possible—not even fully imagined—as little as ten years ago” (D'Arms, 2000).

The change to a digital environment was presaged by the emergence of the dominance of electronic formats in journals, which happened over the span of only a decade (OCLC Reports, 2003; Smith, 2003; Tenopir, King, & Boyce, 2003). A 1999 study by Lenares of faculty at Association of Research Library institutions found extremely rapid adoption of electronic journals. From 1998 to 1999, faculty who identified themselves as users of e-journals jumped from 46% to 61%. Growth cut across disciplines, with the greatest increases occurring in physical sciences and humanities.

Increasingly academics are demanding the ease of electronic access to all library materials, not just journals. In a 2001-2002 survey of 3234 faculty and students sponsored by the Digital Library Federation (DLF) and the Council on Library and Information Resources (CLIR), the vast majority of respondents reported using electronic resources some or most of the time; this was true in research (41.7% all of the time and 51.5% some of the time, total 93.2%) and to a lesser extent in teaching (21.9% all of the time and 67.1% some of the time, total 80.0%). While adopting digital access, these faculty and students were also using print resources to do their teaching and research. In an environment where both types of resources (traditional and digital) were needed, libraries adopted a hybrid approach where older printed material was provided along with the newer digital material. Due to the perceived advantages of digital material for discovery and access, a predominantly digital model is quickly supplanting the hybrid model, leading to more projects where older printed material is converted to electronic format, to the point where “eventually, printed information will become more or less invisible because most users will regard the network as their one and only source of information” (Owen, 2002).

Adapting to the predominantly digital environment requires that a library be able to deal with the challenges that have been identified in the electronic scholarly environment. These include the cost of digitization, the atomized nature of digitization projects to date, which has prevented coherent collection policies, and insuring that ongoing digitization meets the most important needs of scholars (D'Arms, 2000). If libraries in general, and Yale in particular, fail to meet these challenges, current research suggests that they will be left behind by a scholarly population increasingly dependent on digital resources which they can easily find and access on the Internet. In the 2001 DLF / CLIR survey 85.7% of faculty and 83.8% of graduate students either disagreed or strongly disagreed with the statement, “The Internet has not changed the way I use the Library.” This attitude crossed subject lines: 84.1% of faculty and students in biological sciences, 77.3% in the humanities, and 83.7% in the social sciences. In a disturbing trend, almost a majority (48.7% of faculty and 53.4% of graduate students) indicated that they “use the library significantly less than I did two years ago.” Maintaining relevance with scholars in a digital world is the primary challenge facing academic libraries. Making sure Yale’s collections can easily be digitized for teaching and research purposes is one of the tasks Yale University

Library must take up in the 21st century in order to continue to be a trusted, authoritative source for scholarship.

“In the past, libraries constantly updated their collections to remain relevant and to meet the changing expectations of their users. Now, they must update their delivery methods, as well, by offering new means of access to information. One of the next major trends in library modernization is the creation of digital collections based on a library’s own distinctive resources. To remain relevant and useful in an increasingly digital world, where often the same information is available from a variety of sources, each library needs to focus on the resources that make it unique. Libraries should make those resources easily accessible to those who may benefit from them by pursuing well-planned digitization projects” (“Digitization of Library Collections: The Future is Now,” <http://www.ncsconline.org/WC/Publications/Trends/2005/LibCtsDigitizationTrends2005.pdf>). The argument is persuasive that YUL must keep up with the changes demanded by researchers, faculty and students, and by technology.

Yale University Library staff have a critical role to fulfill as colleagues, advisors, and suppliers of vital services to our valued clientele, and a finite window of opportunity to become the heart of the future digital university.

1.2 Peer Institution Programs

Impressive digital library programs are in place at numerous institutions of higher learning, public and private, large and small, generously endowed and modestly funded. The sample below provides a taste of what others have achieved.

Harvard

Harvard’s Library Digital Initiative (LDI), launched in 1998, elegantly demonstrates what can be accomplished when a digital information program is funded as a high priority at the university level with resources on the order of \$20 million (<http://hul.harvard.edu/ldi/>). DPIP’s site visit to Harvard revealed a comprehensive technical infrastructure and support organization for digital content and delivery. Much of the basement level of Widener Library is dedicated to a large-scale digital production facility. A program of internal grants distributes \$500,000 twice each year to digital projects across the campus, encouraging the creation of diverse digital collections that adhere to accepted digitization guidelines and metadata standards and reside in centrally supported repositories.

Cornell

Cornell’s experience with digital library development is particularly instructive for Yale because Cornell has made substantial progress in reorienting its priorities toward digital initiatives without a large infusion of new funding (<http://www.library.cornell.edu/about/digital.html>). For this reason the design of the DPIP program is strongly influenced by Cornell’s three-year-old Digital Consulting and Production Services (DCAPS). During our site visit to Cornell, library staff explained that they are continually redirecting staff in units throughout the organization from traditional work to digital initiatives, thereby gradually increasing capacity to build new infrastructure and collections. It was clear that a strong mandate from senior library

management guides their focus on this transition and helps overcome any natural resistance to change or coordination. DCAPS operates on a partial cost recovery model, based on the assumption that if metadata and digitization services develop a reputation for quality, reliability, and efficiency, many units planning a project will choose to contract work to DCAPS rather than create their own capacity to do the work independently.

University of Virginia

Through its library Electronic Text Center, its Institute for Advanced Technology in the Humanities, and other innovative ventures, the University of Virginia has been a leader in the creation of electronic texts since 1992. A key characteristic of UVa's efforts has been the close collaboration between their faculty, librarians, and technologists in the pursuit of digital collections focused on teaching and research, including the widely acclaimed Valley of the Shadow project sponsored by UVa's Virginia Center for Digital History. Supported by funding from the Mellon foundation, the library's Digital Library Research & Development group, headed by Thorny Staples, has been instrumental in developing the Fedora digital repository architecture into a functional system for general library use. The beta installation at UVa now supports selected collections of full text and images, including the Tibetan and Himalayan Digital Library, and is scheduled for production release in the fall of this year.

University of Michigan

The University of Michigan has invested heavily in digital collections and infrastructure over the past 10 years through its Library Information Technology division, now employing a total of 45 FTE staff. The Digital Library Production Service (DLPS) was formed in 1996 to provide infrastructure for campus digital library collections, including both access systems and digitization services. The DLXS (Digital Library eXtension Service) software developed by Michigan staff provides a stable repository for many of their 248 collections and has been adopted by many other institutions. Michigan's pioneering collaborative work with the Open Archives Initiative has resulted in the widely respected OAIster service which now provides access to 7,437,466 records from 636 institutions (<http://oaister.umdl.umich.edu/o/oaister/description.html>).

University of Oregon

Yale is not alone in confronting the digital challenges outlined below and many other institutions have pursued a similar entrepreneurial approach to digital library development, at least in the early years of their efforts. After several years of such activity, the University of Oregon library recently concluded that "just doing it isn't enough." A mature approach to digital library initiatives requires enhanced coordination and substantial planning. "The Oregon model of 'just do it' is about to do us in.... Clearly, the next step must be to develop business plans and impose project planning on our digital collections work" (Hixson, 2005).

1.3 Yale University Needs Assessment

The development of a next-stage, mature, integrated library with integrated access to collections is a stated major strategic goal of the Yale University Library in response to users' needs and expectations. To provide true integrated access the library must have the strategy and the means to systematically create and provide access to a core collection of digitized material according to

an understood and accepted set of criteria. In 2006, as several important technologies in higher education have converged (network capacity, flexible course management systems, distributed repository environments, cross-search capabilities, and associated tools), a strong digital infrastructure supporting comprehensive digitized collections is increasingly part of expected library infrastructure within the research university. Digital provision of materials enables greater access to library materials in the user's environment of choice, fosters creative and new uses of multi-media materials, enhances pedagogical choices, and not only saves on-campus users time and effort by facilitating access to materials from their personal computers, but also allows the library to break physical barriers to its collections for researchers who cannot come to New Haven.

At her fall 2005 information sessions, the Yale University Librarian expressed the goal that the YUL should be among the top digital libraries in the country. In order to catch up with many of our peer institutions in digital collection building, Yale University Library must invest significant time and resources in developing infrastructure now. As the second largest academic library in the country, we too must serve our readers equally well in the electronic environment as in the physical one, as the source for authoritative research and information management.

1.3.1 The Challenges

While the Yale University Library is still far from the vision of a mature integrated library, important progress has been made over the past five years. The work of the ELI Program and efforts to enhance digital services to Arts faculty have led to improved collaboration and an understanding of who does what and who provides what service—especially with regard to the VRC, the Beinecke, and ILTS. The work of the Integrated Access Council (IAC) includes the introduction of metadata standards such as the Yale Element Set (<http://www.library.yale.edu/cataloging/metadata/YES/guidelinesVersion0-5-July2005.doc>), which is increasing the consistency of metadata across collections. The development of the DLXC cross-collection search for six image collections based on the locally developed DL software (<http://images.library.yale.edu/dlxc/>) represents a significant step toward the ultimate goal of an integrated interface to diverse resources, and the introduction of MetaLib federated searching provides an avenue for integration of both local and remote resources. The creation of the Rescue Repository now provides an interim safe haven for digital masters generated from a variety of Yale digital initiatives and the Digital Preservation Committee is formulating plans for a permanent preservation archive. The VITAL/Fedora research has developed deeper understanding of YUL repository requirements.

In order to leverage the digital initiatives that have been achieved and to continue coordinated development, several challenges remain to be addressed.

Numerous unrelated digital initiatives: Yale's digital library remains a dispersed set of discrete digitizing projects conducted by staff with a wide variety of other duties. Selectors, curators, managers, and faculty, working with various digital groups in the library and at ITS, often forge ahead on their own out of necessity and begin scanning projects without the necessary infrastructure to track library materials from the shelf to the production unit, to perform quality control, to catalog the finished product, and to make it readily available for use.

There is too little sharing of information and too much duplication of effort among library groups who are conducting digitization programs. The outcome for the reader is an array of disconnected projects created with inconsistent standards for digitization, cataloging, and archiving, many idiosyncratic repositories in which to store digital content, and limited means of accessing the valuable digital resources across all collections through a unified interface. The outcome for the library is an inefficient use of resources and staff time.

Need for standards and recommended practices: Up until now, most library digitization projects have developed their own scanning, metadata creation, and presentation practices, expending considerable effort in the process, and delivering products that vary widely. This situation works against the strategic goals of digital preservation and convenient reader access. Long-term preservation of digital objects depends heavily upon the use of reliable, well-documented content formats and upon the consistent application of metadata. Likewise, the success of cross-collection searching (whether federated or harvested) also depends heavily on predictable metadata that adheres to national standards and well-documented practices. In addition, meaningful integration of library collections with learning management systems like Sakai (yaleclasses*v2: <https://classesv2.yale.edu/portal/>) cannot occur without adherence to shared standards.

No clear pathway to success for Yale faculty members who want to create digital teaching resources or for staff seeking to create digital library collections: A wide variety of digital production support services exist within the library and across the campus but they are not well-connected or integrated. As a result, there are redundancies, gaps in service, and inconsistent practices. This situation may lead to confusion and frustration on the part of faculty and staff when they seek basic digitization and metadata production services as well as the information they need to implement, optimize, and manage digital projects. It is clear from limited surveys and individual conversations that some faculty feel strongly that the library has not provided support for digital teaching and research equivalent to its traditional services in the print world—for example, in collection building or course reserves—and has not been able to meet the demand for new services such as widespread assistance in the use of technology in teaching or convenient integrated access to digital resources.

Limited mechanisms for understanding the requirements of our clientele and responding effectively to expressed desires: The rapidly changing needs of library users are not adequately understood and the effectiveness of existing or newly implemented services is not easily measured. Stakeholders throughout the library need tools and mechanisms that will enable them to assess user needs and systematically evaluate web-based services from the users' perspective in order to achieve continuous improvement of services and support.

Inadequate repository infrastructure to support a wide range of digital content and a versatile set of discovery and delivery services: While the library maintains substantial infrastructure for storage and delivery of digital images and provides cross-collection searching of collections housed in the Luna Insight and locally developed DL system, there is no robust library repository capable of accepting material in other formats (full text, audio, video, data sets, etc.) and no commonly accepted practices for achieving global integration of available resources.

Ad Hoc selection mechanisms for digitization: Currently, Yale University Library does not coordinate digitization projects well. Individual units choose material based on grant opportunities, project funding, or user requests. YUL does not yet have an overall plan and long-range objectives for the library's production of digital resources. Ultimately, in order for the library to coordinate digitization and delivery, selectors need a structure that will enable them to prioritize collections for digitization and provide guidance and assistance in funding the most important collections, while taking advantage of opportunities for other digitization projects as they arise. The focus of this effort should be on units and selectors who do not already have a digital collection development program in place. In many ways we have a chicken-and-egg problem: do we rely on funding opportunities to tell us what collections we should digitize, or do we find the collections we want to digitize and aggressively seek funding for them? The answer is that we need to do both.

1.3.2 Ambitious Library Goals

To achieve the goal of establishing and maintaining enduring, accessible digital collections, YUL must create a clear plan with quantifiable goals. This report advocates that the development of a formal DPIP organization become a top YUL strategic priority and that the following objectives be accomplished by June 2008:

The Yale University Library will have

1. Identified a list of the most important library collections to digitize.
2. Completely digitized and provided access to all material in 5 major collections, in the process producing digital formats for:
 - 1,000,000 text pages
 - 500,000 images
 - 10,000 hours of video and sound recordings
3. Established an on-demand system of digitizing library material requested by Yale faculty for classroom instruction or research.
4. Established a strategy to guide digital library development at Yale until 2010.
5. Implemented a unified search which will allow a Yale searcher to find material from any of the library-created, digitized, or born-digital collections.
6. Ensured lasting access through the creation of a robust digital repository.

DPIP will provide leadership in objectives 2 and 3, while other library units or groups will take the lead in the remaining initiatives, with DPIP contributing assistance and defining demand for fulfilling the objectives. The successful completion of these goals is critical to the success of multiple units throughout the library.

Accomplishing this vision by the end of fiscal year 2008 will require that we make the most efficient use of our limited resources, align our work with the needs of students, faculty, and researchers, work cooperatively with other groups on campus, and pursue multiple creative funding strategies, including the quest for outside sources of funding from donors and granting agencies.

Resources must be reallocated from traditional services and a variety of existing activities must be relinquished in order to secure the digital future of the YUL. Digital services, programs, and

collections must command a higher priority than they now do if the Yale Library wishes to achieve a leadership role in the future as information provider for faculty, students and the global research community. For an increasing number of staff, digital initiatives and fundamental production services must have the highest priority and other work must be deferred when the load of digital work is heavy—exactly the reverse of traditional behavior and a key element of a successful digital program. In emphasizing this crucial shift of priorities, we are talking about the full range of digital activities, without a particular focus on DPIP.

To make best use of limited resources and to judge correctly which initiatives will contribute the most benefit to the institution as a whole, a program of assessment must be widely implemented so that the library keeps its goals in line with the needs and expectations of faculty, students and researchers. Additionally, data about the costs of digital library projects should be gathered and analyzed so that refinements can be made to the procedures for carrying out such projects. This data-driven approach will enable the library to realize efficiencies and maximize benefits from a unified digital library service.

2.0 The Digital Production and Integration Program

In order to achieve these goals, a formalized Digital Production and Integration Program will play a critical role in the following ways:

1. DPIP will provide a core staff where knowledge will be developed in the most efficient and cost-effective ways to provide digitization services.
2. To aid in minimizing costs, DPIP staff will track workflow so that they can continuously look for ways to streamline their work as they move from project to project.
3. DPIP will establish a program of assessment of trends in scholarly teaching and research at Yale and beyond its walls to be sure that YUL is responding to the needs of faculty, students, and researchers who use Yale's collections.
4. The DPIP initiative will also provide a mechanism for reaching out to other Yale service providers, such as ITS, so that we can be explicit about the digitization services we provide, while understanding and promoting the services that others may provide in better more efficient ways.

The Digital Production and Integration Program will introduce an opportunity for coordination and planning of library digital initiatives through a three-pronged approach to many of the challenges identified above, and through a virtual organization that will include many of the established players in Yale's digital arena. DPIP's overarching goal will be to provide an intelligible, rational, comprehensive approach to the pursuit of the digital initiatives that it undertakes.

DPIP's activities respond to the recommendations of two important Integrated Access Council reports completed in 2005. The IAC's Portal Opportunities Group Report (<http://www.library.yale.edu/~lso/pog/>) includes among its policy recommendations three topics directly addressed by DPIP:

- "All library-created digital resources should support established integration standards for access, search, metadata practices, and download support (see Digital Library Content

and Course Management Systems: Issues of Interoperation:
<http://www.diglib.org/pubs/dlf100/cmsdl0407.htm>.”

- “The library must engage in usability studies and establish metrics using survey / feedback tools that support user centered design of all new integrated library services. The library should establish a permanent usability lab to support this and other new library service activities.”
- “The library should engage in support and advocacy for the development and implementation of digital repositories that enable the repurposing of content in portal and course management systems.”

IAC’s Review of Digital Repositories (<http://www.library.yale.edu/iac/>) advocates several activities central to DPIP’s services:

- “A cross-unit team should develop policies, tools, workflows, and good practices for digital collection building in order to promote interoperability and sustainability by multiple creators and compilers across the university, including faculty projects and research center-based digital productions.”
- “Articulate standards and best practices for digital assets, repository interfaces, and applications. Articulate the benefits of and build tools that support the life-cycle view of digital assets. Develop best practice guidelines for digital production processes for digital assets. Explore ways of optimally partnering with faculty from the beginning of their data collection and digital production efforts.”
- “Develop deep understanding of user requirements based upon use case studies and other assessment methods.”
- “Develop a multi-dimensional model of requirements based upon the expected life span of digital resources. The model should extend from personal collections without long-term preservation requirements to fully archived resources, and should address stewardship requirements and support costs.”

2.1 DPIP Will Consist of Three Interdependent Services

1. Market / User Research Services

- Systematically gather and analyze user feedback in order to understand user requirements.
- Utilize the Library’s newly initiated Usability Lab to evaluate services such as MetaLib, the EGCDL, and VITAL.

2. Production and Content Integration Services

- Digitize instructional and research materials from and for library collections to fit specific needs.
 - Digital conversion, file reformatting, creation of digital collections, creation of archival and web versions of digital assets.
 - Text imaging, OCR, markup, and presentation.
 - Descriptive, structural, administrative, and preservation metadata creation.
 - Content integration with course management systems (primarily Classes*v2).
 - Quality Assurance for digital production.
- Investigate new production technologies and opportunities, such as outsourcing and partnerships, and bring these into the DPIP workflow.

- Offer Content Management Services.
 - Integration of DPIP content into repositories as appropriate (DL, Insight, Rescue Repository, VITAL).
3. Consultation, Advisory, Referral and Management Services
- Direct staff, faculty, research projects, and students to the services best suited to their needs, whether in the library, at ITS, elsewhere on campus, or outsourced.
 - Publicize best practices—metadata, production, preservation, lifecycle management.
 - Offer project management services.

While the working group focused its attention on digital production services (Service #2 above), it has always been clear to the group that the three prongs of the DPIP plan are closely intertwined. Production services cannot be designed without an assessment of user needs and the services cannot be further refined without user feedback through usability studies or surveys. Digitization, metadata, and content management services can only be delivered effectively to library staff and faculty if there are knowledgeable consultants available to provide consistent, expert, cost effective, and thoroughly practical advice.

2.2 Audience for Services

DPIP's primary goal is to satisfy important unmet needs and to focus attention on those departments, units, groups, or individuals who want assistance in pursuing digital initiatives (especially resources and expertise), both in the library and elsewhere on campus. Services will therefore be targeted primarily at library staff and faculty. A prominent priority is to develop and disseminate the skills and capabilities necessary to build enduring digital collections both from library material and from external collections of research value. For more detail on the nature of services DPIP can provide, see examples in sections 2.6, 3.1, 3.4.1, and 3.4.4.

2.3 Organizational Structure

DPIP will be constructed as a “virtual” organization consisting of core staff in a central group as well as staff distributed throughout the library who report through other library units (e.g., the Metadata Services Group). DPIP activities will not duplicate existing services or pursue initiatives that are being addressed elsewhere on campus. An important characteristic of the virtual structure is that participants must have real authority and designated obligations—a more formal arrangement than the library's committee structure, which is largely based upon voluntary contributions beyond normal work assignments (See diagram in Appendix A).

- Proposed DPIP *Core* elements:
 - Digital consulting services.
 - Digital production services.
 - Assessment and Usability.
- Proposed DPIP *Virtual* elements
 - Metadata Services, including catalogers from SML Catalog Department, Beinecke, VRC, Medical, Walpole, etc.

- Course Support, including staff from ELI program, SQI efforts, VRC, IDIR, Beinecke, Medical, POG, etc.
- Collection Development (including the Collection Development Council) and Preservation Reformatting.
- Copyright Guidance.
- Systems Support for DPIP Services (from ILTS and others).
- Library Digital Production Facilities (DCF, Beinecke, VRC, Maps, MSSA, etc.)
- Proposed DPIP *Partnerships*
 - Library
 - Digital Preservation Committee
 - Metadata Committee
 - Repository Development Initiatives
 - Global Integration / Architecture Initiatives (yet to be established)
 - Development Office
 - External
 - ITS Instructional Technology Group
 - ITS Media Services & RIS
 - Course Management (Sakai) and Portal initiatives

A unified and coordinated program such as DPIP will provide two critically important benefits:

- A consistent point of contact for people beginning digitization and metadata production activities and a reliable source for the information needed to implement, optimize, and manage digital projects. All participants in the DPIP virtual structure will consistently refer inquiries to the most appropriate provider of information or service, whether in the library, at ITS, or elsewhere.
- Continuity in production services. Staff will learn from each project they take on and build expertise that can be applied to the next project, as opposed to starting from scratch with every new independent project. Consistent application of standards and best practices simplifies and enhances lifecycle management of digital objects. This approach can achieve economies of scale and minimize the fragmentation of people's jobs.

A virtual DPIP organization will enable the library to achieve coordinated services within a limited budget by maximizing the value of existing services and filling in gaps where necessary. Strategic planning groups in the library advocated a "federated approach" to Yale's integrated library and argued that a separate digital library unit would create a digital elite and discourage mainstreaming of digital activities. On the other hand, while it is still desirable that elements of digitizing, cataloging, and creating access to digital works should ultimately be integrated into the same Acquisitions, Preservation, Catalog, and Access Services Departments that handle print material, this transformation will not happen easily or quickly without a catalyst. The DPIP virtual structure provides a way to kick-start the process, establish precedents for production workflows, and then involve an ever-widening circle of staff in the support of digital collections. Established programs such as Cornell's DCAPS program have proven that such an approach can succeed.

It is important to emphasize that DPIP is *not* an instrument for exerting unwelcome centralized control over any individual or group—a concern frequently encountered in the deliberations of the working group. The virtual organization assumes mutually beneficial commitments on the part of participants. Core staff will dedicate a major proportion of their time to DPIP responsibilities. Virtual participants will continue to observe their current reporting lines but will make a commitment to share their expertise with DPIP members and will benefit from the services and expertise provided by other DPIP members. Assuming that future detailed negotiations among managers, department heads, and AULs are successful, virtual participants will contribute an average of 30% FTE for each of the six virtual elements (with variations between 10% and 100%). Some virtual participants such as Metadata Services and ILTS may elect to perform substantial work on DPIP projects while others may not. DPIP Partners will have a less formal relationship with DPIP but are regarded by DPIP core and virtual participants as essential to the success of the overall integrated library program, especially in providing fundamental infrastructure or as expert resources to whom DPIP clients may be referred for particular services. Genuinely collaborative work on the E-Reserves prototype (section 3.3.1) over the past two months has proven that this organizational structure can work exceedingly well in the YUL.

2.4 DPIP Staff

DPIP recommends an immediate modest investment in staffing, with the understanding that when projects grow, further staff will be necessary and paid for with project funds.

- Management: DPIP Consulting Services will be managed by Karen Reardon, WW&DCS, ILTS. DPIP Production Services will be managed by Jen Weintraub through an expansion and re-conception of the existing Digital Conversion Facility. DPIP Assessment and Usability Services will be managed by Katie Bauer, the Usability and Assessment Librarian. At the outset, establishment of the virtual DPIP team will require significant coordination and negotiation.
- Backfill: Enable existing staff to devote time to DPIP services by hiring new staff to backfill for them in their present units (to be hired on 2-year term appointments, with likelihood of continuation, funded initially by DPIP startup money and later through reallocated permanent positions or cost recovery income). Jen Weintraub is currently devoting 30% of her time to DPIP through December 2006. Extend this arrangement for an additional year, increasing the time allocation to 40% if possible. In addition, arrange for Karen Reardon to devote 30% of her time to DPIP by hiring a half-time person in WW&DCS at the M1 level beginning in August 2006.
- DPIP will consist of staff at various levels (Student, C&T, M&P, Librarians) from various sources including contributions from virtual members, and with varying time commitments.
- DPIP staff must be capable of providing consistent practical advice to those who wish to start digitization projects at Yale and must develop expertise that will allow projects to be planned and executed efficiently. They must also be able to refer projects that do not fit into DPIP or the library's scope to services or partners at Yale or beyond when necessary.
- DPIP should establish a physical presence in order to achieve a visible profile and provide clients with a concrete place to work with DPIP staff in a professionally suitable environment. This office space will enable DPIP core staff to reside together and meet with staff or faculty who need assistance. A physical space is especially important for

conducting usability studies. Cornell's DCAPS service, for example, has an easily accessible office on the first floor of the main library.

2.5 Communication Plan

DPIP will write and post (on the DPIP web site) quarterly reports on its activities. They will include an account of tasks completed, progress in initial projects, material purchased or services contracted, and adherence to or deviation from the timeline.

In addition, the DPIP working group will remain active for the life of DPIP's first iteration (through June 2008). This group will meet monthly for briefings on DPIP activities and to serve as a sounding board for problems, ideas, and projects.

DPIP core and working group members are on many other committees and groups in the library. They will act as liaisons to these other committees, keeping the library up-to-date on DPIP accomplishments. DPIP will enhance communication on digital initiatives as co-sponsor of forums, presentations, or workshops as opportunities arise.

DPIP projects will originate with requests for services or assistance from faculty or library staff, often through referrals from AULs and through a public web site that DPIP will create. Priorities will be set by DPIP managers, by the DPIP working group, or by LMT, depending on the magnitude of the commitments that need to be made. Preference will be given to services that have the widest impact and digitization projects that contribute most significantly to the library's permanent collections.

2.6 Library / ITS Division of Labor

DPIP will establish an effective working relationship with university units involved in digital production such as ITS Media Services and RIS. This relationship will be non-redundant in that DPIP will not replicate services provided elsewhere.

- DPIP attention will focus on conversion of material held by the library and external collections with enduring research value, especially those promoted by faculty.
- The library should refer to ITS ITG, Media Services, or RIS most requests for conversion of personal faculty collections, short-term course-related content, and quick turn-around tasks. (August 2006 update: At a July DPIP/ITS meeting, ITS staff urged the library to place more emphasis on helping faculty create personal or course-related digital collections regardless of their long-term value to the library. ITS staff stressed repeatedly that digital collection building is not their mission. The recently conducted faculty cyber-infrastructure survey provides further evidence that faculty do indeed want the library to satisfy their most pressing needs for immediate support of digital teaching and research collections. DPIP could appropriately coordinate these activities, but additional funding would be required to support this significant expansion of scope.)
- DPIP will work closely with interested parties to establish one clear pathway for Yale faculty members who want to create digital resources. The ideal solution will consist of many doorways leading to the same answer. If a faculty member's digital resources

would be a valuable addition to the university's permanent collections, then the library's role in the situation becomes highly pertinent.

2.7 Technical Infrastructure and Support (Led Primarily by ILTS)

Several infrastructure investments are necessary to support DPIP's digitization capacity. DPIP will participate in these efforts but DPIP cannot undertake these major tasks on its own and will rely on other library units to assume primary responsibility for them. It must be recognized that DPIP will generate demand for this infrastructure. The library must:

- Establish a repository capable of supporting storage and delivery of electronic text by December 2006—Fedora/VITAL if practical, otherwise DLXS, DigiTool, or Greenstone as a functional alternative. This text repository is an essential pre-requisite for several proposed projects, including the World War I pamphlets, EAD finding aids, and other full text ventures in various units.
- Provide long-term, large-scale, centrally managed, reasonably priced storage for reader-accessible digital files (the use copies), available to all units and systems supporting digital collections. Note that this storage may support multiple discovery and delivery systems and must not be confused with storage required by a preservation repository (for archival master copies).
- Implement a persistent naming system.

DPIP technical support:

- The library should allocate ILTS resources to DPIP initiatives in advance based on projects planned for the coming year, including adequate programming resources for metadata and digital production units.

DPIP facilities:

- At the outset DPIP should not expend limited funds on hardware, software, equipment, or facilities, but should instead make use of existing resources throughout the library and identify outsourcing options, wherever possible.
- As the DPIP services develop, it will become clearer what hardware, software, equipment, and facilities are required. When there is a demonstrable need, purchases should be made. If for example the library decides to engage in a large-scale book digitization project, it might be appropriate to purchase a Kirtas robotic scanner and create a production facility to house the operation.

(See Section 4.0 for more information on hardware, software, equipment, and facilities).

3.0 DPIP Initiatives

Several initiatives have developed with DPIP overview as a result of working group recommendations and also because there have been quickly emerging demands for services such as a DPIP organization would deliver. Not only are these activities good examples of the types of services DPIP will provide, but they also indicate that an unfulfilled need has been quickly satisfied by DPIP.

3.1 Digital Projects

In the spring of 2006, DPIP should begin several substantial, relatively straightforward projects that will have a recognizable impact and a high likelihood of success. A key goal of the initial projects is to build digital production expertise in the library.

- Digitization of an external collection with enduring research value suitable for inclusion as a permanent library resource, such as the Classics Department slides (~10,000 lantern and 40,000 35mm slides), could have significant positive impact with faculty in a particular discipline. (August 2006 status: Classics project in progress.)
- Cataloging and ingest of a Hellenic Studies Byzantine collection into Luna Insight (6,000 images)—a second example of a valuable external collection, this one supplied by a newly appointed faculty member. (August 2006 status: Done.)
- A substantial subject-based collection of pamphlets (an example of “unlocking collections”). The CDC Pamphlets Task Force has won a grant from the Collections Collaborative for a project revolving around World War I pamphlets and other related non-standard material in various formats. This project transcends any one department or Special Collection and intends to identify, describe, digitize, and provide access to the content. DPIP support will be highly beneficial since the Collections Collaborative has not been able to grant enough money to complete the entire project. (August 2006 status: In progress.)
- Collaboration with other grant-funded initiatives, such as the Yale Daily News (YDN) conversion or the Davis project, in order to build shared infrastructure and expertise. (August 2006 status: Assisting Preservation in reviewing vendor responses to YDN RFP.)
- Conversion of the VRC slide collection to digital form in anticipation of the opening of the new Arts Library in the fall of 2008 (371,000 images). A cost estimate for this project was submitted to the provost on January 18, 2006. (August 2006 status: Awaiting response from provost.)
- Provision of access to digitized content generated through YUL’s ongoing preservation reformatting process. Specifications for this task will be developed in partnership with the Preservation Department. (August 2006 status: Initial discussions with Preservation have taken place; samples requested.)

3.2 User Studies and Assessment

- DPIP assisted in the development of an important survey of the Yale faculty. The Yale University Provost's Advisory Committee on the Digital Landscape requested a faculty survey to be completed in draft form by April 2006 for review by the committee. The survey was sent to all faculty members in May. Small focus groups will be formed in the summer and fall to follow up on the survey results. This survey is intended to reveal the particular roadblocks that faculty now encounter in using digital resources at Yale. For example, do faculty find that they have questions about intellectual property law when they create digital materials for teaching? Do they find it easy to search across diverse library collections? This survey targets faculty, and in the future the library will need to look at the particular needs of students as well. Undoubtedly the habits of the Yale community will evolve along with the changing digital world, and assessment will need to be an ongoing activity.

- The Usability and Assessment Librarian is already engaged in several important studies including work with the Integrated Digital Image Resources (IDIR) program and with the MetaLib federated searching implementation.

3.3 Services to Students and Faculty

Some DPIP activities will focus on delivering new or enhanced services to students and faculty, as opposed to the building of digital collections. One example is described below.

3.3.1 E-Reserves (Collaboration between DPIP and Public Service Units)

DPIP will lead the development of a prototype e-reserves service that will define workflows and build basic infrastructure in preparation for a later widespread implementation of e-reserves. In its broadest definition the term E-Reserves refers to providing digital access to materials used for educational purposes; it does not refer to, and is not meant to imply, any specific system or process designed solely to manage e-reserves. Materials include articles, book chapters, film and audio clips, images, etc. The preferred solution for students would be one that provides convenient 24/7 access to the full text of assigned readings and multimedia material within online courseware (i.e. Classes*v2 (Sakai) or Blackboard). Faculty need a resource that will help them solve the complications surrounding the use of copyrighted material in course management systems, and they need a service that will make the entire process of getting both copyrighted and other materials into their online courses easier and more manageable.

DPIP can provide leadership in the investigation of a prototype service and support to e-reserve efforts in various areas including metadata creation, production of digital content from library or library-worthy material, and in usability studies, but cannot (at least initially) take an instrumental role in establishing a comprehensive e-reserve service.

We propose that DPIP work in collaboration with selected school and department libraries, the Portal Opportunities Group, the ELI program, Access Services, RIS, and the ITS AM&T Instructional Technology Group to develop the prototype with a scope including locating citations, scanning material, obtaining copyright permissions where necessary, and making the content available through Orbis and Classes*v2. Access Services does expect that by summer 2006 a newly hired reserves supervisor will be able to participate in a limited e-reserves initiative. The Medical, Social Science, and Divinity libraries are logical potential participants in this effort. ITS has offered assistance in the development of tools that will enable the integration of reserve material into the Sakai collaborative learning environment.

There is some urgency in moving forward with an e-reserves program because the moment is opportune and the need appears to be great. The implementation of Sakai at ITS provides a special opportunity for the library to integrate its resources and services into the environment where faculty and students will conduct much of their teaching and learning activity. In their evaluation of digitization needs and challenges incorporating input from many stakeholders, members of the DPIP working group rated e-reserves as one of the highest priority issues the library needs to address. Limited survey results and conversations with individuals reveal that an increasing number of faculty and students want e-reserves. (A larger-scale needs assessment

could determine more precisely how widespread this sentiment is.) For example, the Working Group on User Requirements for the Teaching and Learning Portal reported that “Faculty had several negative comments about the library resources (5). These ranged from the lack of an online reserve system, to difficult search interfaces, to a lack of subject-specific information. It was also mentioned by two faculty members that it is difficult to integrate library material into Classes Web pages.” At the Art History Faculty Digital Day session on February 21, several faculty expressed serious frustration over the lack of this service. There is anecdotal evidence that some faculty even obtain electronic course material through colleagues using e-reserve services at other universities.

The goal of this prototype effort will be to implement a use case included in the report of the Portal Opportunities Group, where faculty submit requests for e-reserves through Classes*v2 and staff at the library and at ITS respond promptly by locating and delivering the e-content regardless of format or source. Ideally the use case will be broadened to include non-print content and to support the Blackboard CMS in use at the Medical Library. Costs, staff resource requirements, logistics, and levels of user satisfaction will be tracked in order to assess the viability of a wider implementation.

The role of DPIP in this prototype e-reserves service should be to offer project management, conduct needs assessment and usability studies, coordinate the identification of tasks that need to be completed, digitize library material requested or direct work to appropriate units at ITS, and provide funds for software, technical development, obtaining permissions from copyright holders, and temporary operational staffing during the two-year development period. Permanent implementation of the service and expansion throughout all library units including Access Services will require a commitment from senior administrators (LMT and LMC) to reallocate resources for this critical function. (August 2006 status: Processing of material for the e-reserves prototype has begun with four libraries participating (SSL, Divinity, Medical, Arts) and workflow details have largely been determined. Three staff members from Access Services are contributing to the planning. RIS has agreed to do the scanning and copyright clearance work for items that are not already available via links to licensed library e-resources.)

3.4 Recommendations for Future Initiatives

3.4.1 Collection Development (Led by Selectors)

In order for the library to coordinate digitization and delivery, selectors need a structure that will enable them to prioritize collections for digitization and provide guidance and assistance in funding the most important collections, while taking advantage of opportunities for other digitization projects as they arise.

- DPIP, CDC, and the AUL for Collection Development will discuss forming a high-level collection development digital projects committee to administer digital collection development policy and set priorities, supported by a technical group (DPIP participants?) that assesses the feasibility of proposed projects (analogous to the Harvard model). The committee should review and prioritize the list from the forum proposed below and from other projects and opportunities that may develop based on current library initiatives, practices and programs. Prioritization should be based both on content

of the projects and on the technical goals and abilities of Yale University Library's digital library program.

- The collection development digital projects committee will create a prioritized list of collections to be digitized as part of a coherent digital collection development policy. A forum will be held in the coming year to begin the process of creating the list. The focus of this effort will be on units and selectors who do not already have a digital collection development program in place. Selectors will be invited to talk about their collections, and can ask for help in preparing a list. DPIP and CDC members could organize this forum and collect the results into a list.
- The collection development digital projects committee will establish a relationship with the Library Development Office to pursue grant and donor funding for the top priority projects and collections.

(August 2006 update: Ann Okerson, the AUL for Collection Development, has appointed a working group to discuss digital selection policy and develop a set of criteria that subject specialists should consider when selecting collections for digitization. A subsequent working group will then identify and prioritize initial target collections to be digitized. Both groups will report directly to Ann. A recent gift of \$100,000 for digitization of a library collection has demonstrated that such a list of candidate projects could indeed prepare the library to respond quickly to such funding opportunities.)

3.4.2 Metadata (Led by Metadata Specialists in Cataloging Units)

Since metadata is a critical part of the digital information infrastructure necessary to support teaching, research, and publishing at Yale, the library needs to create and sustain a coherent, agile, and scalable metadata production and development program for digital resources, collections, and services. This program must provide administrative, descriptive, and structural metadata to members of the Yale Community to support them as they manage, discover, and use the digital resources, collections, and services available at Yale.

- DPIP will support the development of capabilities of the Metadata Services Team in the Catalog Department so that it can function as a preferred source for assigning metadata elements to digital production projects for Yale University Library. Plans must include increased staffing in the unit in response to demand for services.
- With the Metadata Committee and the Metadata Services Team, DPIP will define recommended metadata production methods for YUL digital projects including element sets, content standards, and best practices. Policies will address descriptive, administrative, structural, and preservation metadata and will adhere to established national standards wherever possible.

DPIP will also assist in:

- Establishing a tiered system for quality of metadata and images (based on factors such as need and expected lifecycle).
- Implementing a set of tools for metadata creation that streamline the work of metadata staff.

The Metadata Services Team will:

- Define a *virtual* Metadata Services Group that includes metadata specialists from other units including Beinecke, VRC, Medical, Walpole, etc. The virtual metadata services group will coordinate metadata production and share management and development

responsibilities for metadata production across existing units and administrative structures.

ILTS should:

- Provide technical support for automated manipulation of data and for the acquisition or development of necessary productivity tools. The Metadata Services Team, working closely with the Metadata Committee and ILTS, should develop the specific requirements for such technical support.

(August 2006 update: The Preservation Metadata Task Force has distributed a draft proposal for a PREMIS-based metadata element set and usage guidelines designed to support the preservation of digital assets in repositories at Yale. After the IAC Metadata Committee and the IAC Digital Preservation Committee have reviewed the report, DPIP will participate in the refinement and implementation of this important set of recommendations.)

3.4.3 Workflow

The experience of the various digital production units in the library should be shared and coordinated and a set of best practices and methods for scanning should be developed. Starting from that basis, we recommend:

- The library needs a system for tracking the digitization process. The library should purchase software (or develop software in-house) to track the movement of an object through digitization. Such software will enable us to know where an object is and to track the time it takes to complete the discrete parts of a task.
- DPIP will thoroughly explore outsourcing opportunities for digital scanning production to determine if and when this approach will be most cost effective and will identify, analyze, and establish relationships with a small number of trusted external vendors for digitization services.
- DPIP will establish an in-house DPIP quality assurance capacity.
- DPIP will work with other groups in the library, the Preservation Department for example, to formulate best practices for different formats of material, taking into account preservation, access, workflow, and storage considerations.

3.4.4 Course Support / Digital Consulting (Often Led by Agents beyond DPIP)

DPIP will play an important role in the library's emerging Course Support Service initiatives. Although DPIP does not have a role in direct course support or in developing new teaching technologies, it will produce, or coordinate production of, content used in teaching and learning at Yale. In this role DPIP will provide services for course support providers such as RSC, the Visual Resources Collection, IDIR, the ELI Program, grant-funded projects like the Davis Educational Foundation Grant, and external organizations such as ITS Academic Media and Technology. DPIP will view these groups as virtual partners in this effort and will work closely with them to achieve a logical distribution of responsibilities in potentially overlapping activities. DPIP will:

- Document service expectations and guidelines, establish strong two-way communication channels, and hold regularly scheduled meetings with its virtual partners in order to achieve timely and responsive course support at Yale.

- Negotiate the creation of unified, permanent repositories for course-related digital content (images, audio, video, text, datasets, etc.) which will house collections created both by the library, by the ITS ITG, and by any other unit on campus generating content of ongoing value. This effort could begin with a focus on arts-oriented images and discussions including VRC, Arts Faculty, ITS ITG, and ILTS.
- Develop and communicate minimal standards for metadata and image quality as well as provide centralized support for faculty wishing to move their collections into an institutional environment.
- Provide consultation resources or metadata services (those things the library does well) to faculty digitization projects done through ITG.

3.4.5 Education / Outreach

DPIP will assist in the creation and advertisement of a matrix of available digitization services in the library and on campus, so that library staff know what facilities can be used, which kind of digitization projects each location can support, whom to contact, and who can use the service.

3.4.6 Copyright

DPIP recommends the creation of a Library Copyright Guidance Group to serve as a clearinghouse for copyright questions and issues related to digitization for any purpose.

- This group would meet to examine the library's documents concerning copyright for digital course support created by the ELI project (at <http://www.library.yale.edu/eli/copyright/>), to update these documents for DPIP's needs, and to publicize these documents to the library. It would also establish mechanisms for helping librarians with their copyright questions. The group would then cease meeting regularly, but would be called upon to help with specific projects or questions and to provide guidance to DPIP consultation services. It may also update the documents as the need arises (ideally on a yearly basis).
- Members of the DPIP working group see the need for a larger Yale University "copyright advisory committee" which would include representatives from ITS, the General Counsel's office, and library administration. This high-level body should define the copyright and IP policies which DPIP staff can then implement and interpret for clients. However, previous library experience in asking for copyright guidance from the Yale administration leads us to believe that the University is not ready for this step. In the meantime, a Library Copyright Guidance Group can address questions about copyright from the library staff until such time as a university-wide group can be formed.

3.4.7 Digital Preservation (Led by Digital Preservation Committee)

Preservation is a crucial component in the lifecycle of digital material. DPIP will assist in YUL's current digital preservation efforts.

- The Digital Preservation Committee (DPC) is primarily responsible for digital preservation policy.
- DPIP should consciously address lifecycle and preservation issues in all aspects of its operations in support of the objectives of the DPC.

- As DPIP increases production of library digital content, its dependence on the Rescue Repository (RR) and its successor will be critical to its success. The DPC and ILTS will be responsible for enhancing the Rescue Repository as needed and for designing and implementing the preservation repository that will serve as the ultimate replacement for the RR. DPIP can contribute to these efforts by assisting with an assessment of current RR services, especially the ingest functions.

3.4.8 Global Integration / Cross-collection Search (Led by a Newly Formed Group)

- Cross collection searching and global integration of content through technologies such as OAI metadata harvesting or federated search are highly important objectives, but the issue is a universal challenge affecting the library's entire web presence and is well beyond the scope of DPIP, especially since it has heavy technical and architectural components. There is no existing library group with a suitably broad charge to address this issue. The topic should be approached through a newly-formed, broad-based avenue, such as a service-oriented digital architecture committee.
- DPIP should support people working on this problem. For instance, DPIP can encourage staff working on specific projects to ensure that their metadata is reusable by higher-level systems such as harvesters and compatible with external projects such as the DLF Aquifer initiative.

4.0 Digitization Resources at Yale University Library and Yale University

Significant digital production resources already exist within the library and across the university. In addition, several library units have extensive experience with issues surrounding the outsourcing of materials.

4.1 Digital Production Facilities within the Library:

(Note: Many of these resources are currently dedicated to digitization of the collections in which they reside. Use of such facilities for DPIP tasks may be possible under certain conditions through negotiations at upper management levels, i.e. AULs and Department Heads.)

Library Digital Conversion Facility (available for booking by library staff):

Open book grey-scale scanner
Large color flatbed scanner (also scans slides and other transparencies)
Large format scanner (for posters and maps)
Small digital camera (consumer grade)
Many different OCR and image-editing software packages

Beinecke Rare Book and Manuscript Library (digitization of Beinecke materials):

Medium format camera with scanning back
Large flatbed scanner
Film scanner
2 printers for duplication
Also testing software for audio/video digitization

Map Library (priority given to digitization of Map Collection materials):

Large format scanner
Large flatbed scanner
Large format printer for duplication
Many different software packages

Medical Library (digitization of Medical Library materials):

Open book grey-scale scanner
2 large color flatbed scanners

Divinity Library (digitization of Divinity Library materials):

Large color flatbed scanner

4.2 Additional Non-hardware-related Resources within the Library:

Preservation Department: Current relationship with Kirtas and Bridgeport bindery for scanning; experience with preservation and reformatting of all media.

Social Science Library: Experience with project management and RFP development and scanning data for EGCDL.

Manuscripts and Archives: Extensive experience with video and sound digitization.

4.3 Digitization Production Facilities across the University:

Media Services (focus is graphic images):

Media Services will do large format scanning / photography as well as their many other services. They currently scan slides for the VRC and have worked with other library units on audio and video digitization.

Prices: http://www.yale.edu/mediaservices/photography/services_fees.html

RIS (focus is text reproduction and document management):

In addition to document scanning, RIS will perform the following digital document processing services:

- Format conversion (PDF, TIFF, etc.)
- Optical Character Recognition (OCR)
- Page Editing
- Page Indexing
- Copyright Clearance

RIS has a production flatbed scanner that can scan text that has been disbound.

Prices: <http://gdwbsrv.its.yale.edu/ris/pricelist.asp?dept=docsvcs>

Media Services Medical:

Media Services Medical also does large format scanning / photography and has worked with the Medical Historical Library. They can provide on-site services if necessary as well.

4.4 Digitization Production Elements Needed for the DPIP Pilot Projects

- Ability to perform full color scanning of bound volumes and oversized materials. For this we need to purchase either a color open-book scanner or a digital camera and obtain training in this kind of digitization. Such a system costs at least \$25,000 not including staffing required to use the camera. If we do not purchase such equipment, we need to establish a budget line item for full color digitization by external vendors.
- A color book scanner for the digitization needs of the multiple in-house projects. (Estimated cost: \$25,000, some funding already budgeted, purchase anticipated in fall 2006).
- A color correction software / monitor / workstation (\$6000).
- Quality Control software and procedures. (It will require more research and experience managing digital projects before we can provide a cost estimate. Initially, we should establish a workflow and after understanding the extent to which we will need to do QC, look for funds for more development or for specific software packages).
- Project management (tracking) software (see above).
- Metadata tools for establishing structural metadata for text.
- Preservation / administrative metadata standards and editing capacities.
- Coordination of library expertise with other multimedia formats (audio / video) to establish what resources would be helpful to acquire on a central level for all collections.

5.0 Implementation Timetable

DPIP has already begun its work. We have started four projects and continue to meet regularly. Our timeline is detailed in a separate spreadsheet. Below is a brief description of the timeline for the various projects.

- **Classics Department Slides:** This project began in April 2006 and was scheduled to finish in June. It is ongoing due to changes in staffing plans in the Classics Department.
- **Hellenic Studies Collection:** This project began in April and was successfully completed in July.
- **Collections Collaborative WWI Grant:** This project was funded and the first phase has begun. It runs until December 2007.
- **E-Reserves:** Planning for this project has begun; a pilot will occur in fall 2006.
- **Electronic Text Repository:** This project has not yet begun; we are awaiting decisions from other library committees, especially regarding plans for a possible VITAL/Fedora implementation.

In addition to these projects, we anticipate examining metadata tools in fall 2006 and beginning work with selectors on the collection development initiatives in summer / fall 2006.

See separate Excel spreadsheet for a detailed DPIP Task List.

6.0 Budget

Attached are two budget proposals to support DPIP's initiatives. The first proposal allocates the funds currently available to DPIP and focuses on the minimum infrastructure and staffing requirements to meet DPIP's current and projected projects. This proposal also outlines funds gained through DPIP's 'virtual' staffing implementation. Allocations of funds include backfill staffing for managerial support of DPIP's digital production and consulting services, funds for student hours, hardware and software purchases and an allocation to support the e-reserves prototype project. It is anticipated that funds will be requested via the equipment request process and it is hoped that additional staffing can be gained through the library's implementation of a 'floating' staffing pool. This budget totals \$890,050 over two years. Of this amount, \$363,105 must be supported by DPIP funding and \$526,945 will come from other sources such as equipment requests and contributions of staffing by virtual participants. All but \$35,000 are ongoing expenses. The budget does not include grant funding or cost recovery income that will help to support some DPIP projects.

The second proposal is the 'strawman' document that describes the funding necessary to build a robust and technically sound digital production capability and its supporting infrastructure for the next 5+ years. This proposal is designed to stimulate discussion and raise an awareness of the funding required for such an initiative. It should be noted that the funds requested are not dissimilar to those being allocated at Yale's peer institutions. The proposal incorporates funding for the digitization of 500,000 images, including the conversion of the VRC's slide collection (funding has already been requested from the Provost's office for this initiative), funding for a large-scale text digitization program, and funding for the technical infrastructure to support these projects—including the design and installation of a robust Digital Repository capable of supporting multiple library services. This budget totals \$7,649,440 over two years.

The DPIP organization is designed to be *scalable* so that it can achieve practical results both on a small scale right away within existing resources and on a much larger scale later on with substantially greater funding. DPIP, along with its virtual participants and partners, can grow in a cumulative fashion over time as the library increases investments in its digital infrastructure.

6.1 Budget Utilizing Existing Resources

See separate Excel spreadsheet for the DPIP Budget Utilizing Existing Resources.

6.2 Budget Necessary to Achieve an Ambitious Goal

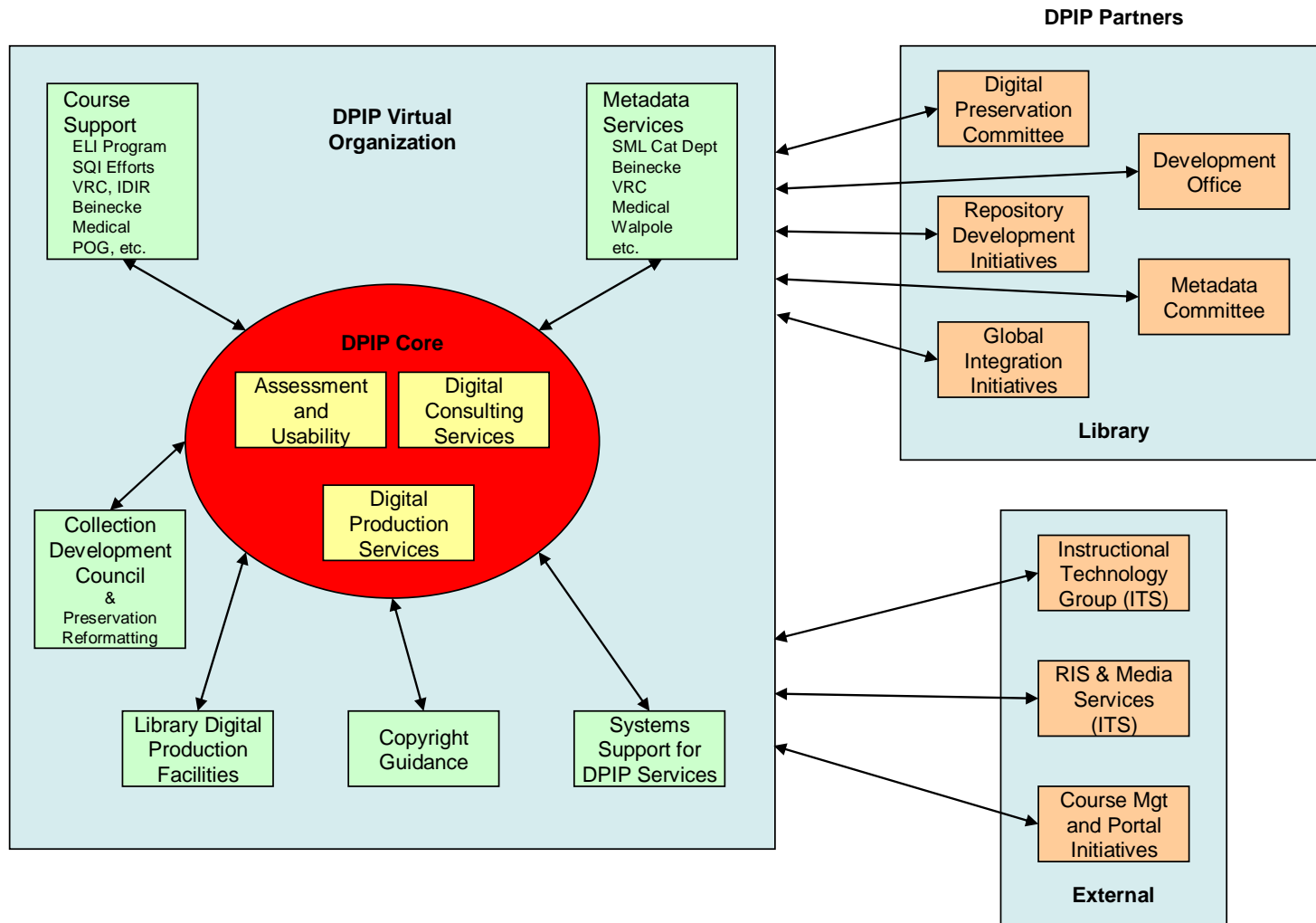
See separate Excel spreadsheet for the DPIP Budget Necessary to Achieve a Digital Collections Strategy Appropriate for the Yale Library.

References

- D'Arms, J.H. (2000). Scholarly publishing in the 21st century, paper presented at the *Annual Meeting of the American Council of Learned Societies*, Chicago. Retrieved May 23, 2006, from <http://www.acls.org/jhd-aha.htm>
- Digital Library Federation. (2002). *Dimensions and use of the scholarly information environment*. Retrieved October 7, 2005 from <http://www.diglib.org/pubs/scholinfor/>
- Hixson, C. (2005). When just doing it isn't enough: the University of Oregon takes stock. *RLG DigiNews*, 9(6). Retrieved May 24, 2006 from http://www.rlg.org/en/page.php?Page_ID=20865#article1
- Lenares, D. (1999). Faculty use of electronic journals at research institutions. *Racing Toward Tomorrow: Proceedings of the Ninth National Conference of the Association of College and Research Libraries*, Detroit, MI.
- OCLC Reports. (2003). *Five-year information format trends*. Retrieved October 31, 2005 from <http://www5.oclc.org/downloads/community/informationtrends.pdf>
- Owen, J. (Nov 2002). The new dissemination of knowledge: Digital libraries and institutional roles in scholarly publishing. *Journal of Economic Methodology*, 9(3), 275-288.
- Smith, E. T. (2003). Changes in faculty reading behaviors: The impact of electronic journals on the University of Georgia. *Journal of Academic Librarianship*, 29(3), 162-168.
- Tenopir, C., King, D. W., & Boyce, P. (2003). Patterns of journal use by scientists through three evolutionary phases. *D-Lib Magazine* 9(5). Retrieved May 22, 2006 from <http://www.dlib.org/dlib/may03/king/05king.html>

Digital Production and Integration Program (DPIP)
 Report of the Production and Content Integration Working Group

Appendix A



Digital Production and Integration Program