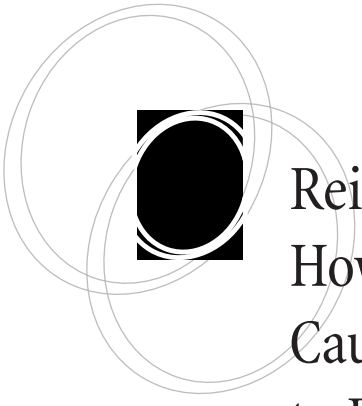
Project
MUSE®
Scholarly journals online



Reinventing the Library—How Repositories Are Causing Librarians to Rethink Their Professional Roles

Tyler O. Walters

abstract: The rise of digital repositories is helping libraries reinvent themselves. The benefits to libraries and universities creating institutional repositories (IRs) are great as libraries restructure, pursue collaborations, and re-position themselves to become major digital publishers and broadcasters in the scholarly world. They will no longer be passive receivers of information but active disseminators of intellectual output for entire universities. This article explores the IR's role in overall library transformation and examines the organizational changes and internal partnerships necessary to strengthen IR programs. Specifically, the creation of digital library initiatives units as well as changes to other library units are discussed.

Introduction

In many ways, when libraries create institutional repositories (IRs), they are reinventing themselves. Traditionally, libraries have managed information produced by organizations—namely publishers—outside of their parent institutions. They select, acquire, organize, make accessible, promote, preserve, and instruct people about how to use these information resources. However, IR developers are primarily concerned with content generated internally—that is, with the intellectual output (usually in digital form) of

In many ways, when libraries create institutional repositories (IRs), they are reinventing themselves.

portal: Libraries and the Academy, Vol. 7, No. 2 (2007), pp. 213–225.
Copyright © 2007 by The Johns Hopkins University Press, Baltimore, MD 21218.

their university communities. As they establish IRs to collect and disseminate this information, their libraries become active producers, publishers, and broadcasters.[1] The steps to develop and manage IRs may be similar or even the same—select, acquire, organize, access, instruct, promote, and preserve—but the procedures required to accomplish them are beginning to have a profound impact on libraries, changing their work processes, redefining their on-campus roles and relationships, and changing their collective philosophical mindsets.

This article will explore the IR's role in overall library transformation and examine the organizational changes and collaborative partnerships necessary to strengthen IR programs. As an example, the author provides the Georgia Institute of Technology Library and Information Center's experiences in building its own IR, SMARTech, and reviews the restructuring the library underwent, including the ascendance of its new IR-hosting unit, the Digital Initiatives Department.

Institutional Repository Developers—"Good Stewards" of Digital Resources

The rise of digital institutional repositories is extremely significant to scholarly communications. Many universities are gathering content produced by research faculty, making it searchable and maintaining it within digital repositories. These institution-based repositories may one day feed into larger "meta-repositories," providing the ability to search timely research information across universities within a discipline or subject area.[2] A large national and international network of repositories may arise, presenting new technical challenges and underscoring the need for standards (such as the Open Archives Initiative [OAI] and beyond[3]) to maintain interoperability.

According to Clifford Lynch, "A university-based institutional repository is a set of services that a university offers to the members of its community for the management and dissemination of digital materials created by the institution and its community members. It is most essentially an organizational commitment to the stewardship of these digital materials, including long-term preservation where appropriate, as well as organization and access or distribution."[4] A university's digital materials include a diverse body of items, such as:

- Annual reports
- Computer programs
- Conference papers
- Data sets
- Learning/complex objects (digitally captured courses, multimedia simulations/visualizations, textual documents, and captured notes of faculty and students)
- Models
- Lecture series materials
- Pre-prints/post-prints
- Proceedings
- Research reports
- Simulations
- Technical reports and working papers

- Web pages
- White papers

As librarians become committed stewards of their universities' digital resources, they are organizing, preserving, providing access to, and creating rights management systems for these kinds of institutional resources as part of their daily responsibilities.

Digital Library Initiatives Units—"Champions of the IR Cause"

In many cases, libraries will find it necessary to establish a digital library initiatives unit (DLI) to act as a conduit through which collecting, disseminating, preserving, and collaborating with other organizations occur. Such a unit can develop the single voice and vision needed to articulate the myriad possibilities for scholarly communications—promoting new services and developing and explaining new processes. They also may act as conduits for other departments within the library, exploring the contributions these areas can make to IR development.

Composed of librarians, archivists, technologists, and other staff who make the case for overall changes in scholarly communications, DLI units become the "champions of the IR cause." While taking the leadership role to establish the IR-specific agenda, DLI units also develop robust programs to manage and broadcast intellectual content, provide the services to manage that output, and become advocates for faculty intellectual property rights. Thus, the unit contributes to a new scholarly communications paradigm and becomes an integral part of the infrastructure to support it.[5]

At the Georgia Institute of Technology Library and Information Center, we conceive of our IR and Digital Initiatives Department in the following way, as expressed through the department's mission statement:

> The GT Library Digital Initiatives Department is charged with building effective, dynamic knowledge and research management systems to preserve and provide access to the intellectual output of Georgia Tech. Its responsibilities include identifying, assessing, collecting, preserving, providing access, and making this output more valuable through digital information technologies, whether "born-digital" or convertible to digital formats. The Department provides these resources and services in support of the research and educational endeavors of the Georgia Tech community and to scholars around the world.[6]

In the past, DLI units largely worked on discrete projects created from the holdings in archives departments and special collections. Although these units continue such projects as a part of their mission, increasingly they are shifting their focus to address changes in scholarly communications and the impact those changes are having upon their universities. To guide the universities through these transitions successfully, DLI units must face a great challenge. How do they build a robust technical architecture and service provision model? To meet this challenge, the units are developing strong partnerships with other library departments, engaging them and making use of their expertise in a wide variety of areas, ranging from technology development to marketing and outreach, instruction, and other areas like copyright administration. Library departments such as cataloging, systems, information services/reference, and special

collections/archives must become active, integral partners in the programs responsible for managing university-based digital intellectual output. The design of this architecture and model must involve a university's colleges and departments in order to meet the needs of students, faculty, and researchers. Technical requirements also need to be met to deliver worldwide "on demand" desktop access and address associated preservation issues, ensuring user access over time and across technologies.

Activities of the Digital Library Initiatives Units

DLI units engage in a variety of activities, with a growing central responsibility for IR development, maintenance, and growth. Many unit activities support university functions and processes to create the digital objects themselves. These serve the IR's ultimate purpose of gathering, distributing, and preserving intellectual content. Some new IR-related digital services, as well as more traditional ones, include:

- Conference production processes that result in intellectual output
- Electronic journal production
- Digital collections building
- Audio/video capture of lecture series speakers, symposia, and instructor lectures
- Technical support of virtual communities, online scholarly forums, and other scholarly-related social software content that produce scholarly conversations capable of being captured
- IR technology development for managing, accessing, and preserving digital collections
- Scanning and conversion of intellectual resources into digital form and converting from one digital form to another
- Technical support of future processes that produce scholarly communications
- Digital format migration and emulation processes for preservation and access purposes

University libraries at Cornell and Tennessee have developed such services. For example, Cornell has been experimenting with format and media migration services for faculty.[7] Whereas these services initially focus on helping faculty keep their digital objects current for instructional use, they also help maintain the digital objects in the latest formats and media, thereby making them accessible and operable once deposited in an IR. Tennessee has accomplished much the same results with its Digital Media Service.[8] Cornell also has developed its Digital Consulting and Production Services (DCAPS) to assist faculty with Web and e-journal publishing, database building, copyright issues, and other technology services for generating and disseminating digital learning and information objects.[9] Georgia Tech (GT) is experimenting with a scholarly conference output service to assist faculty in conference hosting by producing, capturing, and disseminating their intellectual output via the IR. These and other services will play a growing role in assimilating IRs into university-based information services.

Organization of the Digital Library Initiatives Units

Library administrators embracing these new trends are beginning to wonder how they can organize their internal library functions to support such services. At Georgia Tech, this very question shaped our library's goals, organization, and IR program. From 2003 to 2004, we diligently pursued pilot project status to become a regularly funded program of services as quickly as possible. This was an imperative in the library's five-year (2002–2007) strategic plan.[10] Figure 1 illustrates the program functions for managing digital intellectual output.

Attempting to address all of these program elements can be daunting. DLI units and IR developers must rely upon the existing expertise resident in various quadrants of the library. To gain the diversity of expertise needed, Georgia Tech established its own DLI unit, called the Digital Initiatives Department (DI), and created the Digital Initiatives Work Group (DIWG). The former is currently comprised of a librarian/department manager with two librarians and two skilled professional/exempt personnel. The latter is comprised of DI Department members, as well as members of Library Systems (an application developer for digital repositories and collections and a senior systems analyst IV), the Archives Department (a digital collections archivist), and the Cataloging Department (a catalog and metadata librarian). The following chart illustrates the

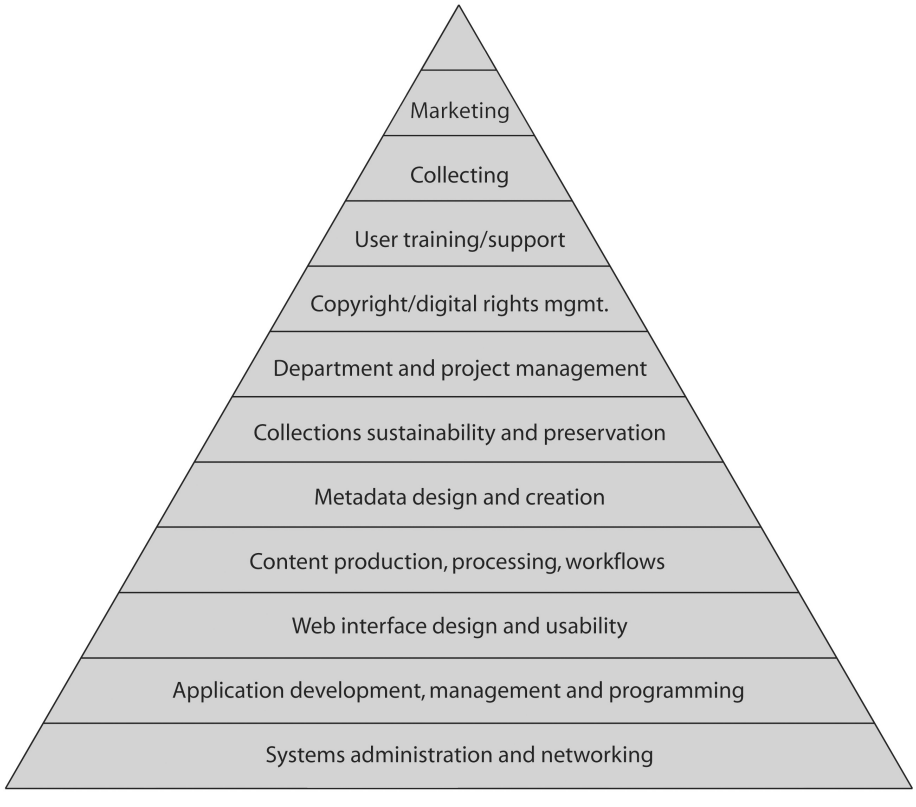


Figure 1. Program Elements to Manage a University's Digital Intellectual Output

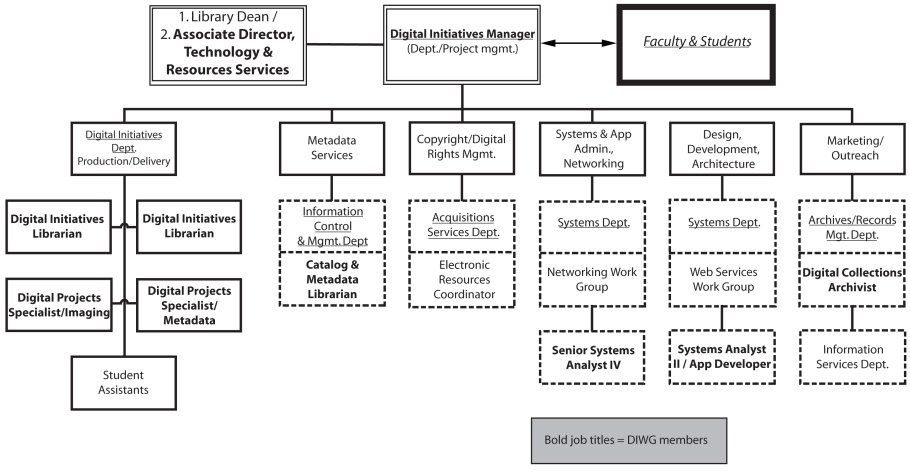


Figure 2. Functional Assignments for the Digital Initiatives Work Group (DIWG)

functional areas and the non-DLI library departments involved in each of them, with DI Department staff involved to varying degrees in all.

Although DI Department personnel are actively engaged in each area, their exact role depends upon existing staff knowledge and the production capabilities in each department. For instance, new metadata needs are typically identified and addressed by a DI librarian, who involves the catalog and metadata librarian (GT Library's Information Control and Management [ICM] Department is responsible for cataloging) early in the process and who, in turn, takes the new metadata practices to the ICM Department and trains and supervises cataloging staff in metadata creation and input. In other areas, the DI Department is not as involved, such as with the campus e-publications program in SMARTech, which is managed and implemented by the Archives Department. Thus, staffs from other library departments apply their expertise to sustain production workflows (as is the case with metadata production) and to maintain a collection development program (as is the case with campus e-publications). Currently, we are looking at ways to more actively engage the library's Information Services Department in marketing and outreach and the Acquisitions Services Department in intellectual property issues. This approach allows DI staff to focus on acquiring new content, addressing related access and preservation issues, and continuously developing the IR itself.

Cataloging Departments—Inputting "Armies" Ready to Serve

Library catalogers are making the transition from metadata creators for resource discovery alone to metadata managers for resource discovery, technical/administrative, preservation, and possibly even structural metadata as well (for example, TEI encoding of documents). Someone must handle all of these aspects of metadata—their creation, maintenance, and migration—to manage successfully the digital objects contained in the IR. Cataloging and metadata specialists can fill this role. The management of digital

objects involving metadata has grown far beyond MARC and Dublin Core—they include applying schema such as PREMIS, SCORM, MPEG-7, METS, OAIS, DIDL, EAD, and more. As catalogers learn about the metadata needs of digital objects, the systems that contain them, and the users who need to find the objects, they are also seeing practical and important ways in which they can contribute to IR content management.

Cataloging departments, with their finely honed workflows and materials cataloging procedures, can become the much-needed soldiers who volunteer for such IR-related duties as ingesting content and generating metadata. Whether the content is self-archived, library-collected, or transferred automatically between information systems, staff with a background in metadata should review submitted items and augment the metadata supplied. Although many universities have operated IRs for at least the past two years, their experiences suggest that self-submission has not yet been adopted widely. Libraries need to adjust for this and offer to carry out the submittal process on behalf of faculty and students. This phenomenon may change over time as intellectual output producers such as faculty become familiar and comfortable with IR systems. However, this has yet to gain much popularity in the United States. Many faculty may view content submission as a library task. Libraries have never asked faculty to label and barcode books, catalog them, and place them on the shelves, so some librarians question why we would ask submitters to perform similar tasks in IRs. In any case, it seems that libraries need to take on the content ingestion/metadata creation routines for some time. As a result, cataloging departments can play this vital role in IR development.[11]

Although many universities have operated IRs for at least the past two years, their experiences suggest that self-submission has not yet been adopted widely.

Because self-submission is not a universally accepted practice, some libraries have adopted creative approaches to ingest content. By collecting content from the university's public Web pages and setting up automatic transfers from other campus information systems, libraries are expanding their IR holdings. Some of this content includes research outputs from sponsored research programs, managed by offices of sponsored programs; e-theses and e-dissertations, collected by the graduate school thesis offices; and student portfolio systems, operated by student services offices. This type of automated transfer helps to populate the IR with rich faculty- and student-produced intellectual content.

At Georgia Tech, the library's Information Control and Management (ICM) Department is in charge of transferring the research reports into the IR, SMARTech, from the GT Office of Sponsored Program's document management system called WebWISE. With each output type mentioned above, the author (such as student, faculty, or principal investigator) usually provides some basic metadata with the content. Cataloging departments can assist IT staff with designing the transfer process, taking into account content verification, metadata crosswalks to map metadata into the IR, and metadata quality reviews for correct word spellings and titles, appropriate keywords from applicable thesaurus terms, and related quality checks. Because many externally published resources can be purchased with MARC or DC records that are easily downloaded

into a library's online catalog, cataloging departments can now devote time to internal digital resources that are worthy of broader dissemination. By managing the automatic transfer of intellectual content into the IR as well as reviewing self-submitted content and metadata, cataloging departments are core contributors, increasing holdings and advancing IR services.

Systems Departments—"Engineers" of the IR World

Library systems departments also serve a crucial role in developing IRs. They are the "engineers," building bridges between applications and maintaining both systems and software to make the whole IR function. Many systems departments have staff, such as systems administrators, application developers, programmers, information architects, and Web designers. Working closely with DLI units, these staff members integrate as much of the IR technology support as needed by the unit into the routine responsibilities of the library systems department. It is counterproductive to separate the daily management of IR systems from the same routines that library systems departments perform to maintain and develop other library information systems. These "routine activities" range from maintaining the server on which the IR runs—with its operating systems and applications software—to administering security and network access protocols, designing the user interface and associated Web pages, and developing plug-in applications that provide value-added content services. If the library systems department is already providing services like these for other library applications, then the department can perform the same services for IR technology development and maintenance. In addition, because much of the IR software is open source, the systems and DLI units can both contribute to software development communities and user groups. This also creates buy-in among systems department staff as collaborative efforts allow the DLI units to gain their input during IR tool development.

The GT Library approach to systems department integration is as follows:

- The systems administrator manages the server hosting SMARTech.
- The Web developer/designer designs the site, completing all HTML work.
- The senior systems analyst IV performs application-specific activities, including upgrades and applying plug-in tools, and handles metadata and content ingest issues, EZProxy integration, and so on.
- The systems analyst II (a new application developer position) develops repository- and digital collections-related applications such as value-added services and integrates them with existing SMARTech functions and content as well as with other campus information systems.

As library systems departments perform IT core functions and collaborate with IR specialists to develop software for institutional repositories, additional useful IR services will be possible on individual campuses and globally.

Reference/Information Services Departments—The "Sales, Marketing, and PR Staff" for IR Development

Reference and information services departments are in powerful positions to shape and define successful IR services. Public service librarians, especially those acting as subject liaisons to academic units, can play several important roles. Primarily, they can observe faculty and students who produce and use intellectual content. This is vital because librarians must understand thoroughly how members of the academic community utilize their own, as well as their colleagues', intellectual output. This will influence repository design, especially searching capabilities, content submission routines, and IR-based, value-added services such as personalized services, statistical assessments, and document outputs. Many public service librarians possess these observation and assessment skills already; they use them when observing and learning about their patrons' information seeking and use habits with externally published information. Part of this observer role can include gathering feedback on IR usefulness and the need for possible improvements and sharing this information with the library staff responsible for technical and functional development. Public services, in general, are taking an ethnographic approach to understand the communities they serve and the role information plays in academic endeavors. As they continue to gain this knowledge, these "market analyzers" will strengthen the usefulness of IRs and broaden campus support for their development.[12]

Public service librarians can foster the IR in other ways as well. They can act as "sales" and "public relations" representatives, promoting IR services and training users and contributors how to interact with the system. These activities are vital to gain IR users and establish the repository as an indispensable hub for intellectual output. Public service librarians can also help train users of other "allied" information systems, such as the university's digital asset management system(s), student portfolios, sponsored research reporting systems, and others. They can educate faculty and students by explaining how these systems link to the IR and can emphasize the IR's role as a central collection and access point for university intellectual content. Indeed, they can be "information problem solvers" when it comes to using such a diversity of information systems on a modern campus. These librarians can also keep faculty abreast of changes in scholarly communication, due the rise of the open access movement, making them aware of their publishing options. This is critical spadework if libraries are to carry out the Scholarly Publishing and Academic Resources Coalition (SPARC) agenda of open access.[13] For libraries to set new scholarly communication trends and support the SPARC agenda, well-developed IRs will be necessary. Public service librarians are well qualified to promote the IR and tie it to this larger agenda.

For libraries to set new scholarly communication trends and support the SPARC agenda, well-developed IRs will be necessary.

University Archives and Special Collections Departments—Committed Partners and Mentors

University archives and special collections departments are natural partners with DLI units. Archivists tend to build strong campus alliances to promote archival services and develop their collections. They also approach building their departments and services in a way very similar to librarians building an IR. Historically, archives have collected campus-generated print material that we refer to as "grey literature," such as technical and working papers series, pre-prints and post-prints, research reports, annual reports, and campus unit and student-produced publications. These materials may form the basis of an active IR. The network established by archives to collect these and similar materials can be tapped by IR managers trying to gather digital versions. In fact, the DLI and archives departments would work most effectively together to collect these materials in the IR as opposed to operating independently.[14]

At Georgia Tech, the GT Archives manages SMARTech's campus ePublications program.[15] The Archives Department is responsible for collecting and ingesting university units' publications and has recently sought out, collected, and assisted in producing one generated by the student body. The *Technique*, GT's student newspaper, is hosted on a GT Office of Information Technology server. The library archives the Web version of the newspaper and its related XML, PDF, and image files in SMARTech.[16] The Archives Department oversees this project while working closely with the library's digital initiatives manager.

Archivists can act not only as partners but also as mentors. For instance, their experience and methods would be invaluable to IR personnel as they select digital learning objects for retention. These objects, which are proliferating on many campuses, are some of the most challenging types of digital output to select and appraise for long-term value. They are complex, multiple file-based objects further defined as being "modular digital resources, uniquely identified and metatagged that can be used to support learning."[17] Instructors design learning objects generally in support of class-related instruction for the purposes of illustrating, demonstrating, or simulating principles or conditions conveyed to students through educational programs. IR managers could apply archival appraisal methods to determine the continuing intellectual, educational, and historical value of these objects (such as documenting the state of instruction in a given subject area).

Similarly, IR managers could benefit from archivists' experience when setting information access restrictions. Although the role of an academic librarian is typically to provide public access to information, there are many reasons to incorporate access/restriction management tools into IRs. Some materials that may require restricted access include:

- Faculty-authored research reports on corporate-sponsored or confidential government projects
- Student-authored dissertations containing patent-pending ideas
- Research team-generated data sets that may help produce future commercial technologies
- Instructor-developed modules not ready for disciplinary community review, which are used to expose students to the latest research in a field

These are a few of many instances when access to digital intellectual output must be restricted. Academic librarians can draw from the experiences of archivists to develop their IR services for a clientele that needs to manage both public and restricted information and learning resources.

Archives and records management programs also establish records retention schedules. Archivists use these schedules to determine the life expectancy of a specific body of records when managing the overall life cycle of records. The schedules set both the flow of these materials from one information system location to another and the retention time at each location. Applying this approach to digital objects in an IR can be very effective. The GT Archives' Records Management Program staff developed such an approach when working with the GT Office of Sponsored Programs (OSP). Retention schedules resulting from this collaboration ensure the transfer of all publicly accessible final reports of research projects into SMARTech. OSP transfers all other materials to the GT Archives-managed records center; these include the financial and administrative records and the research progress/status reports for all public and confidential research projects and the final research reports of confidential research projects. The library's Digital Initiatives Department was involved in these retention decisions. There are myriad areas in IR service management in which archives can contribute significantly. Aligning the programs of the archives/special collections and DLI departments to collaborate and complement one another will be an essential strategy in building successful IR services.

Conclusion

IRs are helping libraries reinvent themselves. They are no longer passive receivers of information but active disseminators of intellectual output for entire universities. The benefits to these libraries and universities are great because they are positioning themselves as major digital publishers in the scholarly world. As the experiences at the Georgia Tech Library attest, library departments are restructuring and diligently pursuing collaborative partners to manage their university's digital intellectual output. Librarians now realize that to gain support for their IRs they must assertively pursue new content as well as new customers and continuously provide new services and the technologies to support them. This involves many skills, ranging from the ability to create metadata and integrate information systems to the ability to promote the IR as an indispensable information management tool. Libraries must use every human and technical resource available to lead the design of a new technical infrastructure for modern scholarly communications and research. Only then will they become the hub for campus scholarly communications of all kinds and attain a leading position in the world of Web-based information dissemination.

Tyler O. Walters is associate director for technology and resource services, Georgia Institute of Technology Library and Information Center, Atlanta, GA; he may be contacted via e-mail at: tyler.walters@library.gatech.edu.

Notes

1. Tyler O. Walters, "Strategies and Frameworks for Institutional Repositories and the New Support Infrastructure for Scholarly Communications," *D-Lib Magazine* 12, 10 (October 2006), http://www.dlib.org/dlib/october06/walters/10walters.html (accessed January 26, 2007).
2. Cornell University Library, "arXiv.org Homepage," Cornell University Library, http://www.arxiv.org (accessed January 9, 2007).
3. Open Archives Initiative, "Open Archives Initiative—Home," Open Archives Initiative, http://www.openarchives.org (accessed January 10, 2007).
4. Clifford A. Lynch, "Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age," *ARL Bimonthly Report* 226 (February 2003), http://www.arl.org/newsltr/226/ir.html (accessed January 10, 2007).
5. Leslie Chan, "Supporting and Enhancing Scholarship in the Digital Age: The Role of Open Access Institutional Repositories," *Canadian Journal of Communication* 29, 4 (2004): 277–300, https://tspace.library.utoronto.ca/bitstream/1807/2786/3/Chan_CJC_IR.pdf (accessed January 10, 2007).
6. Catherine M. Jannik and Tyler O. Walters, "Mission and Collecting Policy" (Digital Initiatives Department, Georgia Tech Library and Information Center, February 22, 2006).
7. Cornell University Library, "Format and Media Migration Service," Cornell Research and Assessment Services, http://www.library.cornell.edu/iris/migration/ (accessed January 10, 2007).
8. The University of Tennessee, "Digital Media Service," The University of Tennessee (2003), http://digitalmedia.utk.edu/ (accessed January 10, 2007).
9. Cornell University Library. "Digital Consulting & Production Services," Cornell University Library (2005), http://dcaps.library.cornell.edu/ (accessed January 10, 2007).
10. Office of the Dean, The Georgia Institute of Technology Library, "Five-Year Strategic Plan, 2002–2007" (September 2003), https://intranet.library.gatech.edu/info/documents/stratplan03.pdf (accessed January 26, 2007).
11. For background on the intent of self-archiving, see: Stevan Harnad, "Self-Archive Unto Others as Ye Would Have Them Self-Archive Unto You," *International Journal on Science Communication* 6 (September 2003), http://jekyll.comm.sissa.it/commenti/foc06_02_eng.htm (accessed January 26, 2007); Harnad, "Open Access to Peer-Reviewed Research Through Author/Institution Self-Archiving: Maximizing Research Impact by Maximizing Online Access" in *Digital Libraries: Policy Planning and Practice*, ed. Derek Law and Judith Andrews (Ashgate Publishing, 2003), http://www.ecs.soton.ac.uk/~harnad/Temp/digital-libraries.htm (accessed January 10, 2007).
12. Barbara Jenkins, Elizabeth Breakstone, Carol Hixson, "Content In, Content Out: The Dual Roles of the Reference Librarian in Institutional Repositories," *Reference Services Review* 33 (September 2005): 312–24.
13. Heather Joseph, "The Scholarly Publishing and Academic Resources Coalition: An Evolving Agenda," *College & Research Libraries News* 67 (February 2006): 84–6.
14. For some other strategies to collect IR content, see: Morag Mackie, "Filling Institutional Repositories, Practical Strategies from the DAEDALUS Project," *Ariadne* 39 (April 2004), http://www.ariadne.ac.uk/issue39/mackie/ (accessed January 26, 2007); see also, Walters for details of other collecting strategies.
15. The Georgia Tech Library and Information Center, "Georgia Tech ePublications: Community Home Page," Georgia Tech Library and Information Center (2002–2006), http://smartech.gatech.edu/handle/1853/7354 (accessed January 10, 2007).
16. Georgia Tech Student Publications, *Technique* 91, 11, (October 7, 2005), Georgia Tech Library and Information Center, http://smartech.gatech.edu/handle/1853/7781 (accessed January 10, 2007).

17. David A. Wiley, "Connecting Learning Objects to Instructional Design Theory: A Definition, a Metaphor, and a Taxonomy," in *The Instructional Use of Learning Objects: Online Version*, ed. David A. Wiley (2000), http://reusability.org/read/chapters/wiley.doc (accessed January 10, 2007); David Beck, "Learning Objects: What Are Learning Objects?" (2007), http://www.uwm.edu/Dept/CIE/AOP/LO_what.html (accessed January 10, 2007).