

Evolution of Scientific and Technical Information Distribution

Sandra L. Esler, Michael L. Nelson

NASA Langley Research Center, MS 124, Hampton VA, 23681-0001. E-mail: m.l.nelson@larc.nasa.gov

World Wide Web (WWW) and related information technologies are transforming the distribution of scientific and technical information (STI). We examine 11 recent, functioning digital libraries focusing on the distribution of STI publications, including journal articles, conference papers, and technical reports. We introduce 4 main categories of digital library projects: based on the architecture (distributed vs. centralized) and the contributor (traditional publisher vs. authoring individual/organization). Many digital library prototypes merely automate existing publishing practices or focus solely on the digitization of the publishing cycle output, not sampling and capturing elements of the input. Still others do not consider for distribution the large body of “gray literature.” We address these deficiencies in the current model of STI exchange by suggesting methods for expanding the scope and target of digital libraries by focusing on a greater source of technical publications and using “buckets,” an object-oriented construct for grouping logically related information objects, to include holdings other than technical publications.

Introduction

The rising cost of traditional scientific scholarly communication (Quandt 1996; Cummings, et al. 1992), coupled with the increase of widely available Internet communications tools, such as the World Wide Web (WWW) (Berners-Lee, et al., 1992), have provided the catalyst for a revolution in the exchange of scientific and technical information (STI). This article examines digital libraries (DLs) applied to STI distribution, and the resulting impact. We review 11 DLs providing access to full-text journals or technical reports and evaluate the history, direction, and design of such services. Abstract servers are not considered. Projects that produce CD-ROMs of STI are not considered either. Our focus is on projects that provide networked delivery of the final product -- a dusty CD-ROM is just as inaccessible as a dusty journal.

We find that most DLs follow strictly along discipline or publishers' boundaries, with little interaction occurring or even possible between servers. In addition, most DLs tend to fall along artificial media boundaries, and servers that provide electronic access to traditional publications generally do not provide access to video, software, datasets, or other information.

We introduce a nomenclature for categorizing STI servers and examine methods of expanding their holdings. We also suggest that the best method for future access is not in placing traditional journals on the World Wide Web or other information servers, but rather redefining the unit of STI exchange. To this end, we introduce the concept of “buckets,” an object-oriented construct for logically grouping related STI, including non-publication STI.

Background and Shortcomings of Current STI exchange

STI includes more than just research journals. Scientific journals evolved in the 17th century to replace the system of exchanging personal letters between scientists, which evolved because of unacceptable time delays in publishing books (Odlyzko, 1995). However, journals are no longer used for rapid communication, but rather as “a medium for priority claiming, quality control and archiving scientific

work” (Bennion, 1994). In some disciplines, such as high-energy physics, the pre-print culture is well established and Ginsparg (1994) notes that “The small amount of filtering provided by refereed journals plays no effective role in our research.” While noting that not all disciplines embrace the pre-print culture equally, Odlyzko (1995) states “it is rare for experts in any mathematical subject to learn of a major new development in their area through a journal publication” and also relates a comment by Rob Pike, noted computer operating systems researcher, “that in his area journals have become irrelevant.” Others have called for change in the process of STI distribution, either in the form of a (logically) centralized archive of STI (Gardner, 1990), or in the replacement of the binary review system (reviewed/non-reviewed) with a continuum of review status (Harnard, 1990; Okerson & O’Donnel, 1995).

A journal article is often only a fraction of the available technical literature about a given subject. Theses, dissertations, conference papers, and technical reports are known as “gray literature,” and receive varying degrees of peer review. “White literature,” available through standard publications channels and processes, is often supported by a larger body of gray literature. The role of the large amount of gray literature and its relation to the smaller amount of white literature, and the issues associated with integrating the two have been present since the post-World War II U.S. Government sponsored research boom (Bennington, 1952; Scott, 1953; Gray, 1953). David Patterson (1994), co-inventor of the RISC computer chip, noted that in one of his first research projects, the output was 2 journal articles, 12 conference papers, and 20 technical reports. If we consider this pyramid of publications (Figure 1) to be typical, then a journal article actually functions as an abstract of a larger body of STI. Although this process of condensation and abstraction may be useful for reading by non-specialists, specialists may want access to the full array of available information, including the gray literature.

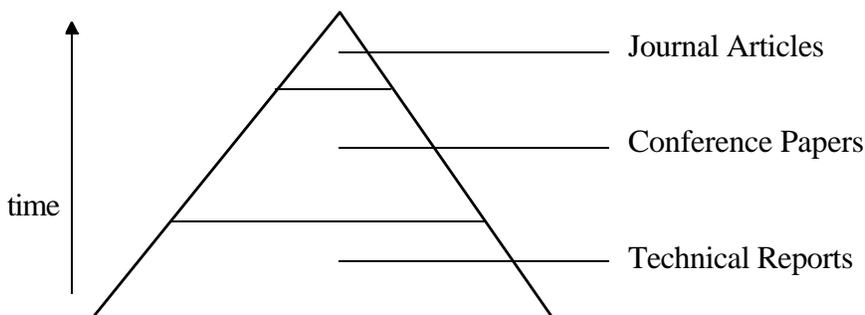


Figure 1: Pyramid of Publications for a Single Project/Concept

Using the following “back-of-the-envelope” numbers, we estimate at least 100,000 domestic, unrestricted technical reports are published annually:

600 Federal labs * 100 reports/year	=	60,000 reports
250 Research universities * 6 departments * 40 reports/year	=	60,000 reports
100 Corporate research laboratories * 30 reports/year	=	3,000 reports

		123,000 reports

600 federal laboratories is taken from the Federal Lab Consortium (FLC, 1996), 250 research or doctoral universities is taken from CAUSE (CAUSE, 1996), and 100 corporate research laboratories is taken from the American Association for the Advancement of Science (AAAS, 1996). Number of departments, and rates of publishing are based only on cursory observation.

Rahman Khan of the National Technical Information Service (NTIS) stated that the NTIS acquires approximately 100,000 reports per year (Khan, 1996). However, approximately 30% of NTIS accessions originate outside of the U.S. On the other hand, NTIS does not receive non-U.S. government-sponsored

academic or industrial research reports. In summary, 100,000 reports a year seems accurate to at least an order of magnitude.

The result is that even if there are 20,000 primary research journals (Bennion, 1994), they do not represent the entirety of STI. This is without addressing 1) confidential, secret, proprietary, and otherwise restricted reports; or 2) non-report STI, such as computer software, data sets, video, geographic data, etc. Schatz and Chen (1996) give a summary of current research projects focusing on building large digital libraries of non-report STI.

The limitations of current STI exchange mechanisms can be summarized as:

- highly focused on journal articles, despite their decreasing value to researchers and practitioners in some fields;
- inadequate acquisition of gray literature, the grist of technical exchange;
- inability to offer non-publication media, such as datasets, software, and video.

These limitations are largely side effects of the hard copy distribution paradigm. As STI exchange moves toward electronic distribution, existing mechanisms should not merely be automated, but the entire process should be revisited.

Overview of Current Projects

Below are a number of DLs that provide full text access to publications; abstract- or meta-data-only servers are not considered. Also, for purposes of classification, we consider only STI publication servers. With the exception of the University of Illinois Digital Library Initiative, DLs fall along the scientific discipline they service (Table 1).

Discipline	Digital Library Server / Technology	URL
Aerospace	Langley Technical Report Server / TRSkit	http://techreports.larc.nasa.gov/ltrs/
	NASA Technical Report Server	http://techreports.larc.nasa.gov/cgi-bin/NTRS
Astrophysics	STELAR	<defunct>
	Astrophysics Data Systems	http://ads.harvard.edu/
Chemistry	Chemistry Online Retrieval Experiment	http://www.oclc.org:5047/oclc/research/projects/core
Computer Science	Unified Computer Science TR Index	http://www.cs.indiana.edu:800/cstr/search
	CS-TR	<defunct>
	WATERS	<defunct>
	NCSTRL	http://www.ncstrl.org/
Physics	Physics e-Print Server	http://xxx.lanl.gov/
Various	Digital Library Initiative	http://www.grainger.uiuc.edu/dli/

Table 1: Digital Libraries by Discipline

Langley Technical Report Server / TRSkit - Begun as an anonymous FTP server in 1993 (Nelson & Gottlich, 1994) and graduating to WWW in 1994 (Nelson, et al., 1994) the tools used to develop the Langley Technical Report Server, TRSkit (Nelson & Esler, 1997), have been distributed to a variety of NASA and Air Force installations.

NASA Technical Report Server - The most prevalent aerospace digital library is the NASA Technical Report Server (NTRS) (Nelson, et al., 1995). NTRS is really a gateway for various servers based on TRSkit. NTRS provides access to over 12 NASA centers, institutes and programs, and contains 2,000+ papers and 3,000,000+ abstracts, mostly NASA- and NASA Contractor-authored technical reports. Access is via WWW and there are no access restrictions.

STELAR - An early digital library that predated the popularity of WWW (Van Steenberg, 1994). The public version provided only abstract access to several thousand abstracts; only registered users could access scanned journal pages. Funding of STELAR expired in 1994 and its holdings were merged into the Astrophysics Data Systems.

Astrophysics Data Systems - The Astrophysics Data Systems (ADS) began as an abstract only service, and has added scanned journal pages where available (Accomazzi, et al., 1995). ADS also absorbed the STELAR data upon its termination. ADS has created 2 related databases: Space Instrumentation and Geophysics. Access is via WWW and there are no access restrictions.

Chemistry Online Retrieval Experiment - The Chemistry Online Retrieval Experiment (CORE) began in 1991 as a method of distributing SGML pre-prints of 7 chemistry journals (Entlich, et al., 1995). Access is via an X Window System client and users must register with OCLC.

Unified Computer Science TR Index - The Unified Computer Science TR Index (UCSTRI) began in 1993 as method providing uniform access to the various computer science departmental anonymous FTP servers (Van Heyningen, 1994). An automated program would traverse a list of known FTP sites, and assemble a searchable index from information available from the FTP site. Access is via WWW and there are no access restrictions.

WATERS - The Wide Area Technical Report Service (WATERS) was a prototype of several computer science departments to offer on-line access to theses and technical reports through a WWW interface (Maly, et al., 1994). WATERS shares its heritage with LTRS, and eventually was merged into NCSTRL.

CS-TR - Begun in 1992 with Carnegie Mellon, Cornell, MIT, Stanford, and the University of California at Berkeley, this project focused on scanning 5,000 computer science technical reports from the participating institutions (Kahn, 1995). Along with WATERS, CS-TR formed the core of the Network Computer Science Technical Report Library.

Networked Computer Science Technical Report Library - The Networked Computer Science Technical Report Library (NCSTRL) has a highly defined protocol, Dienst (Lagoze, et al., 1995; Davis, et al., 1995), for coordinating over 50 cooperating technical report servers (Davis & Lagoze, 1994). There are multiple levels of NCSTRL compliance, with the former WATERS evolving into "NCSTRL-lite." Access is via WWW and there are no access restrictions.

Physics e-Print Server - The Physics e-Print server has existed in various forms since 1991, transforming from an e-mail based LaTeX/TeX server, to the current WWW service that generates PostScript on the fly (Ginsparg, 1994). The e-Print server has been especially prolific, servicing over 70,000 transactions a day (Ginsparg, 1996) and spawning over 15 servers for physics sub-disciplines. The American Mathematical Society has also duplicated the architecture to handle math e-prints. Access is via WWW and there are no access restrictions.

Digital Library Initiative - The Digital Library Initiative (DLI) project at the University of Illinois is building a prototype that will provide access to a variety of technical journals, including those from the IEEE (computers and electronics), AIAA (aerospace), APS (physics), AIP (physics), and ASCE (chemistry) (Schatz et. al, 1996). The project is still in development, so no usage information is available. Access is via a Microsoft Windows client and usage will likely be restricted.

The lineage of the various projects described above is summarized in Figure 3. A feature summary of the DLs is provided in Table 2.

	Sponsor	WWW Access?	Free Access?	# abstracts	# papers	Formats available	Contributors
LTRS / TRSkit	NASA	Yes	Yes	N/A	N/A	PS, PDF, HTML	Organizations
NTRS	NASA	Yes	Yes	3,000,000+	4,000+	PS, PDF, HTML	Organizations
STELAR	NASA	Yes	Abstracts free; Papers restricted	10,000+	10,000+	TIFF	Journals
ADS	NASA / Harvard	Yes	Yes	870,000+	20,000	GIF, JPEG	Journals
CORE	OCLC	No	No	31,000	31,000	SGML	Journals
UCSTRI	Univ.	Yes	Yes	unknown	unknown	PS	Organizations
CS-TR	Univ.	Yes	Yes	5,000	5,000	TIFF	Organizations
WATERS	Univ.	Yes	Yes	1,000+	1,000+	PS	Organizations
NCSTRL	Univ.; Labs	Yes	Yes	5,000+	5,000+	PS, TIFF, GIF	Organizations
e-Print	LANL	Yes	Yes	100,000+	100,000+	PS, LaTeX	Individuals
DLI	UIUC	No	No	4,000	4,000	SGML	Journals

Table 2: Summary of Digital Library Features

Distribution Method Categories

A natural partitioning of DLs is apparent. The various projects can be differentiated by their architecture (distributed or centralized) and by identity of the sponsor of the DL (traditional publishers or authoring individuals/groups). Figure 2 illustrates the partitioning along with the abbreviations for their taxonomy, and Figure 3 shows the lineage and progress of the various DLs.

Centralized Architecture, Traditional Publisher (CP) - Input is from traditional publishing sources such as journals and professional societies, and all input is collected in a single physical and logical location. The server is either up or down, there is no graduated level of availability.

Distributed Architecture, Traditional Publisher (DP) - Input is from traditional publishing sources such as journals and professional societies, but the input is not transmitted to a single physical location. The user interface may give the appearance of a central location, but the service is comprised of several servers.

Centralized Architecture, Authoring Individual/Organization (CO) - Input is from either individuals (a few papers at a time) or from an organization (papers transmitted in batches) and the input is transferred to a central location for indexing, processing and redistribution.

Distributed Architecture, Authoring Individual/Organization (DO) - Input could still be from individuals, but separate servers encourage clustering of publishers along organizational boundaries. Input stays at the server it was posted at and the user interface handles querying all appropriate servers and collating and presenting the results.

	distributed	centralized
traditional publisher	DLI <i>DP</i>	ADS, STELAR, CORE <i>CP</i>
authoring individual / organization	NTRS, NCSTRL, UCSTRI <i>DO</i>	WATERS, CS-TR, Physics e-Print Server, LTRS/TRSkit <i>CO</i>

Figure 2: Distribution of Digital Libraries

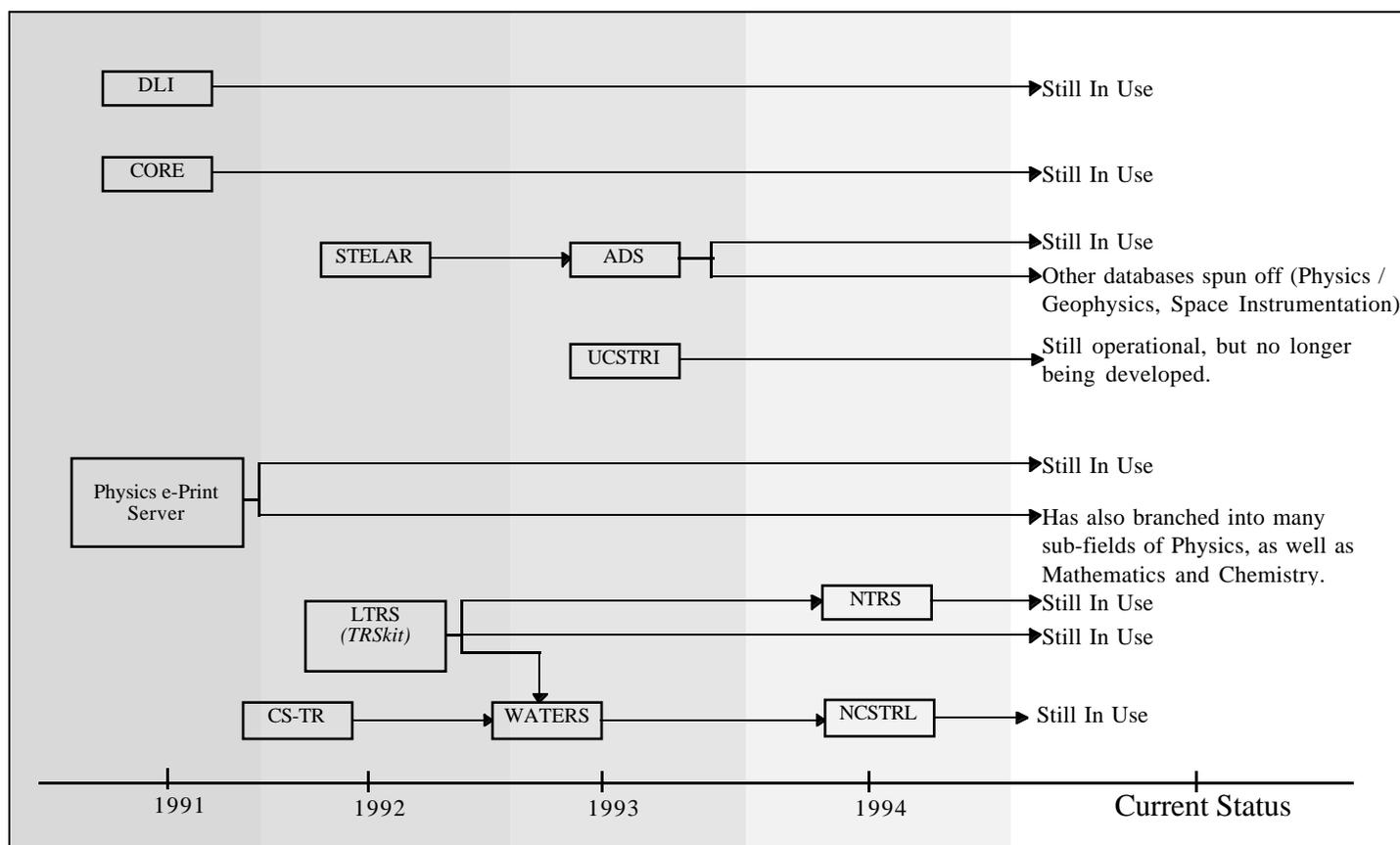


Figure 3: Digital Library System Ancestry

New Model for Distribution

We believe the distributed architecture with author/organization input is the model of distribution with the most growth potential. Of the servers examined so far, NTRS, NCSTRL, and UCSTRI come the closest to fitting this model of distribution, with NCSTRL being the best defined and most extensible. However, no DL project has fully implemented all facets:

- Institutions being their own publishers;
- Many distributed, cooperating servers;
- Increased number of accessible publications; and
- New media formats beyond traditional publications.

Redefining “ublisher” - Digital Libraries should be fed by the publishing process itself. Institutions and authors are already acting as defacto publishers of gray literature; this proposal formally defines their role. Note that there is no technical reason why professional societies and other traditional publishing entities could not participate as well.

Distributed, cooperating servers - A hierarchical model allows greater scalability than a monolithic database. Databases can be added, deleted, and upgraded in real time without impacting other databases. There is also no single point of failure for the entire system.

Accessing more technical papers - This includes access to traditionally poorly cataloged gray literature. The collection and abstracting by central organizations has another interesting side effect -- there is information loss as it is moved around. Having the institution as publisher pushes the responsibility for collection integrity to the level that has the most interest in its maintenance: the authoring organization. Additionally, by making it easier to be a publisher, organizations currently on the periphery of return on investment of publishing will be brought in.

Access more than technical papers - Taking NASA Langley Research Center customers as typical, customers want access to more than just the technical publications (Roper et al., 1994). They also want access to software, datasets, and other technical material. This architecture allows individual organizations to add different media formats at their own pace in the manner that makes sense in their domain.

Figure 4 illustrates the architectural model. The model is “logically central, physically distributed”; the user will interact with a single interface which hides the details of the various servers underneath. Organizations publish up to the protocol, and the central interface is responsible for handling transactions above that level. While it is assumed that all organizations will maintain their own servers, it is not a requirement. The central organization or some third party could offer an intermediate server for those who did not wish to maintain the operational server.

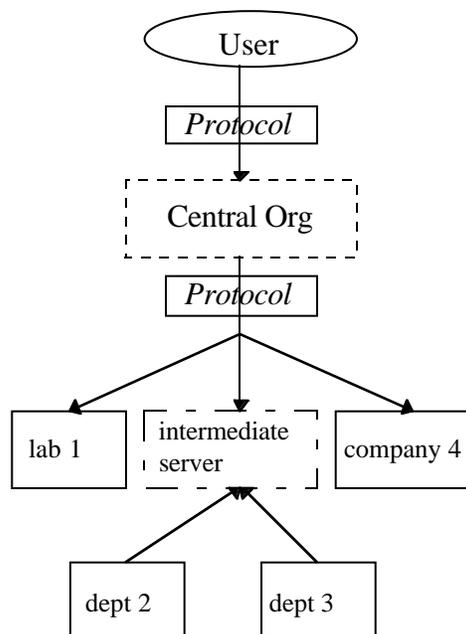


Figure 4: Distributed, Organization as Publisher Model

Organizations considering creating a digital library are often overwhelmed by the apparent magnitude of the problem. For example, NASA Langley Research Center has been publishing reports since 1917 and a substantial percentage of these reports are still in demand (Smith, 1992). However, it is helpful to think of this as two separate problems: conversion of legacy information, and capturing current and future information (Figure 5).

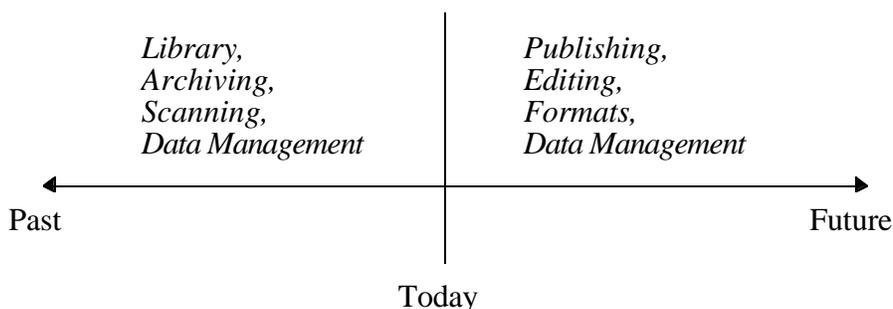


Figure 5: Existing STI and Future STI are Different Problems

Although there should be coordination between the two efforts, they are not necessarily the same. It is possible to begin ensuring today's and future publications are captured, even if there are no resources for the conversion of legacy hard copy holdings. The size of the past collection should not deter one from addressing the current and future problem.

While the focus so far has been primarily on freely available gray literature, the DO model does not preclude peer review of, or monetary exchange for, the documents. A current journal title could be a unit in the schema, with digital signatures attached to the documents attesting to their authenticity. Likewise, the content could be encrypted for security and economic concerns, and one would only be able to decrypt if they had paid a "subscription," or a dynamic micro payment mechanism could be used, or the appropriate user authentication made.

New Media, New Formats

A side effect of shifting the publishing role from traditional commercial academic publishers will be a decreased emphasis on the Standard Generalized Markup Language (SGML) (Goldfarb, 1990). SGML has received a lot of attention in the electronic publishing community, but has yet to find widespread acceptance as a transmission or display format. Some projects that receive their input directly from the publishers (DLI, CORE) are working with SGML.

However, the bulk of scientific publishing is done in LaTeX/TeX, Framemaker, and Word. Output formats have historically been DVI and PostScript, and now increasingly PDF. Until popular word processing / desktop publishing packages have a "Save as SGML" option, SGML will not establish a foothold in the new model of STI distribution. We believe that Hypertext Markup Language (HTML) will server primarily as a navigation language, useful en-route to the final product, likely to be in PDF. PDF's hypertext capability is likely to be sufficient for scholarly requirements as stated in Kintsch (1990).

SGML is the technically superior product, and may have a role as an archival format with, say, PDF generated from it on the fly. It is important to note that an archival format and the format presented to the user do not have to be the same. But the increased focus on end-user publishing products will likely relegate SGML to a niche market. A useful analogy is Beta and VHS. Beta is the technically superior product, and is used in video editing outfits, TV stations, etc. However, the rest of the U.S. uses VHS.

Universal Clients

Use of the World Wide Web as the transport mechanism allows the DL provider to focus resources on increasing the server's capacity and functionality by leveraging the client development expertise of a wider market, WWW browsers. While it is possible that specific user functionality could be needed above what a basic WWW browser can deliver, the prevalence of Java (Arnold & Gosling, 1996) and other languages provides a convenient mechanism to extend browser functionality from the server side. Creating custom client applications is now rarely justified.

The New Role of Central STI Organizations

By deferring the job of publishing to the lowest possible level, the nature of existing STI central organizations changes. Much of the functionality remains at a high level, but the processes are not bound by the transmission of hard copy.

Academic publishers, specifically journal publishers, can easily transition to the new model. The journal's main roles of priority claiming, quality control, and archiving (Bennion, 1994) all map into a digital medium. Journals can maintain their own servers, and "publish" as any other organization does. In fact, journals could serve as a review and summary of the large amount of information available across the various servers. Journals could be constructed by "author-push" or "editor-pull" into a journal. A journal's utility would be rated by its customers in its ability to present the most relevant and engaging summary, and improvement of the large quantity available information. To ensure quality control, digital signatures can be used to authoritatively establish review status. Hard copy journals will be available for many popular titles, but they will likely be considered a derivative of the canonical digital version, as opposed to the current reverse situation. The more focused, limited readership journals are the more likely candidates for all-electronic publication. The period of transition will undoubtedly involve missteps, but the path is inevitable. The quality of a journal is derived from the editors and authors, not necessarily from longevity (Odlyzko, 1995), so the transition will not undermine the institution of quality.

Some organizations exist to collect and redistribute technical reports, particularly U.S. government-sponsored research. These organizations include:

- National Technical Information Service (NTIS -- federal, state, and local government);
- Center for Aerospace Information (CASI -- NASA);
- Defense Technical Information Center (DTIC -- Department of Defense); and
- Office of Scientific and Technical Information (OSTI -- Department of Energy).

Currently, these organizations collect hard copy input and redistribute it to institutional library and individual customers on a cost recovery basis. Their future focus will not be on the logistics of receiving, cataloging, and distributing hard copy, but rather on the value added to the end-customer on information retrieval and support. and coordination for institutional publishing.

To the customer, central organizations will provide the searching interface, update services, intelligent agents for automated pre-processing, cross-disciplinary searches, and other value-added services. For the publishers, the central organization will provide the coordinating service for publication, as well as be the maintainer for the necessary tools, protocols, etc. Figure 6 illustrates the current institution, user, central organization relationship, and Figure 7 illustrates the future relationship.

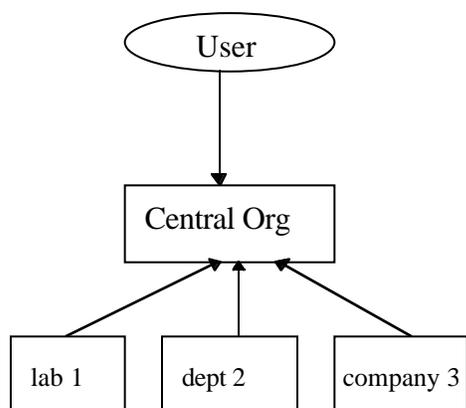


Figure 6: Current model -- central organization collects input, processes requests

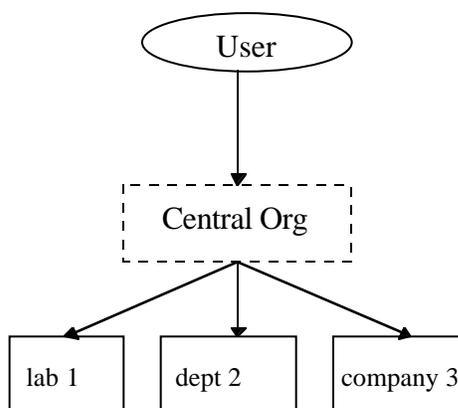


Figure 7: Future model -- central organization passes requests through to authoring organizations

Proposed Agenda

The distributed server, distributed publisher model has not been fully realized by any DL project. The groundwork is laid in many projects, notably the Dienst Protocol used in NCSTRL and the multi-disciplinary nature of DLI. However, too many projects still fall along discipline boundaries. A canonical digital library protocol, or at least a method for exchange and mapping between a set of canonical protocols, is needed. A robust protocol and toolset would allow the focus of DLs to switch to the more interesting areas of value-added services and non-traditional holdings.

An interesting area of research would be the use of expert systems to search archives of multiple disciplines for relevant information. Users could express a request in the terminology they are comfortable with, and the terminology could be translated to other fields, perhaps the underlying meaning abstracted and generalized, and applied across multiple domains (Figure 8)

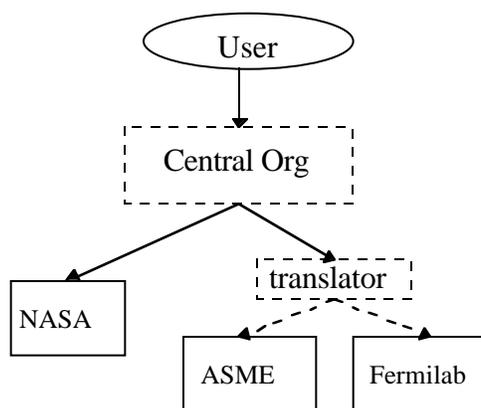


Figure 8: A User's Aeronautics Request

In addition, digital libraries could gateway to objects beyond traditional publications. There are many digital library projects dealing with visual, geo-spatial, music, and other information formats, but these tend to arise as separate stand-alone databases. For example, NASA research projects often produce tuples of a formal report and a research software code (Sobieski, 1994). The report element of the tuple has a well defined review, distribution, and archival path to follow. However, the software does not. Some digital libraries are in use for software, the most well-known example being Netlib (Browne et al., 1993). However, the best solution is not to have a logically separate software digital library (or for video, datasets, etc.), but to build a digital library that can accommodate the tuple as a single logical entity (Figure 9).

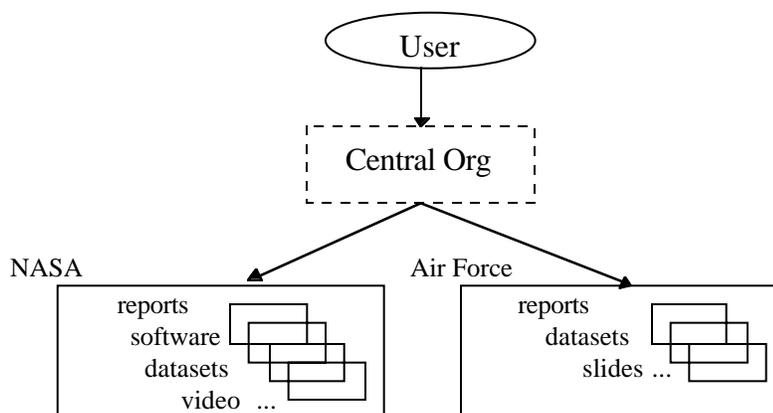


Figure 9: Digital Libraries With Expanded Information Formats

It is possible to stretch the definition of “report” so that it includes formats such as video and software, but overloading the term “report” becomes especially confusing when distinguishing between the “special” on-line version, with its added functionality, and the traditional hard-copy version, without the additional formats. It is best to use a new term for a publishable unit that bounds the set of a traditional hard-copy publication, as well as software, datasets, etc. We suggest the term “bucket” (Figure 10) because terms such as “package,” “container,” and “object” are overloaded, and bucket provides a clear visual metaphor for its functionality. The key feature of a bucket is not only does it provide for the logical grouping of related information items, but a bucket is also designed to be a customizable, intelligent agent. That is, a bucket can exist inside a traditional archive, or exist as a stand-alone object. Buckets can perform actions by

themselves, communicate with other buckets, learn new access or display methods, provide logging and annotation support, and provide object-level, customizable access control. Buckets also have customized control over how their contents are unveiled to the user. Bucket prototypes are discussed in more detail in Nelson, Maly & Shen (1997).

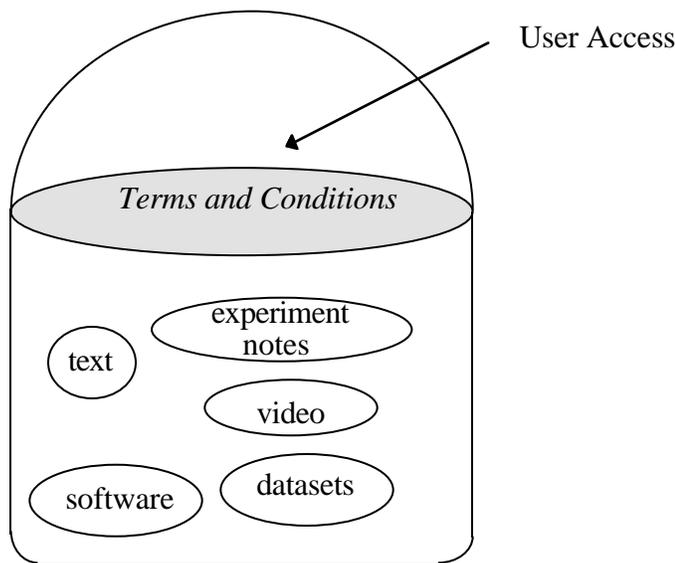


Figure 10: A Bucket -- The set of all publishable information units

Conclusions

An examination of 11 digital library projects reveals a low level of inter-operability between various servers. Digital library services tend to be specific to the discipline they serve, with no knowledge of the status or existence of digital library projects in other fields. Many digital library projects are focused narrowly on simply providing electronic access to journal articles, despite evidence suggesting the decreased usefulness of journals in some fields.

There are 4 significant architectural classes of digital libraries: distributed versus centralized servers crossed with traditional publishers versus authoring individual/organization. The class of *distributed server and authoring individual/organization* is best suited for scaling to future requirements as well as expanding the definition of an STI digital library. An expanded digital library will include more non-traditional publication sources such as gray literature, and other non-publication objects such as software codes, datasets, video, etc. We feel that it is best to group publications with related non-publication STI in "buckets," rather than create separate digital libraries for each STI format.

While no current digital library project successfully realizes this goal, NCSTRL and the Dienst protocol have the most well-defined architecture to serve as the foundation for future progress in digital libraries.

References

- Accomazzi, Alberto, Grant, Carolyn Stern, Eichhorn, Guenther, Kurtz, Michael, & Murray, Stephen (1995). ADS Abstract Service Enhancements. *Proceedings of the Fourth Annual Conference on Astronomical Data Analysis Software and Systems*, H. Payne, D. Shaw, J. Hayes (ed). Publications of the Astronomical Society of the Pacific Series.
- American Association for the Advancement of Science (AAAS). (1996). Corporate and Industrial Research. <http://sci.aaas.org/sources/corp.html>
- Arnold, Ken, & Gosling, James. (1996). *The Java Programming Language*. Addison-Wesley: Reading, MA.
- Bennington, J. (1952). The Integration of Report Literature and Journals. *American Documentation*, 3(3), pp. 149-152.
- Bennion, Bruce C. (1994). Why the Science Journal Crisis? *ASIS Bulletin*, February/March 1994, pp. 25-26.
- Berners-Lee, Tim, Cailliau, Robert, Groff, Jean-Francois, & Pollermann Bernd (1992). World-Wide Web: The Information Universe. *Electronic Networking: Research, Applications and Policy*, 2(1), pp. 52-58.
- Browne, S., Dongarra, J., Grosse, E., Green, S., Moore, K., Rowan, T., and Wade, R. (1993). Netlib Services and Resources. University of Tennessee Technical Report UT-CS-93-222.
- CAUSE (1996). Institutions by Carnegie Classification. http://cause-www.colorado.edu/member-dir/institutions_by_carnegie.html
- Cummings, A. M., Witte, M. L., Bowen, W. G., Lazarus, L. O., & Ekman, R. H. (1992). University Libraries and Scholarly Communication. Washinton D. C.: Association of Research Libraries.
- Davis, James R., Krafft, Dean B., & Lagoze, Carl (1995). Dienst: Building a Production Technical Report Server. *Advances in Digital Libraries*, Springer-Verlag, pp. 211-222
- Davis, James R. & Lagoze, Carl (1994). A Protocol and Server for a Distributed Digital Technical Report Library. Cornell University Technical Report TR94-1418.
- Entlich, R., Garson, L., Lesk, M., Normore, L., Olsen, J., & Weibel, S. (1995). Making a Digital Library: The Chemistry Online Retrieval Experiment. *Communications of the ACM*, 38(4), p. 54.
- Federal Lab Consoritum (1996). The Federal Laboratory Consortium For Technology Transfer. <http://www.zyn.com/flc/>
- Gardner, William (1990). The Electronic Arcive: Scientific Publishing for the 1990s. *Psychological Science*, 1(6), pp. 333-341.
- Ginsparg, P. (1994). First Steps Towards Electronic Research Communication. *Computers in Physics*, 8, pp. 390-396.
- Ginsparg, P. (1996). Winners and Losers in the Global Research Village. *Proceedings of Electronic Publishing in Science*, UNESCO HQ, Paris, Feb. 19-23, 1996.
- Goldfarb, Charles F. (1990). *The SGML Handbook*. Oxford University Press.

- Gray, D. E. (1953). Organizing and Servicing Unpublished Reports. *American Documentation* 4(3), pp. 103-115.
- Harnard, Stevan (1990). Scholarly Skywriting and the Prepublication Continuum of Scientific Inquiry. *Psychological Science*, 1(6), pp. 342-344.
- Kahn, Robert E. (1995). An Introduction to the CS-TR Project.
<<http://www.cnri.reston.va.us/home/describe.html>>
- Khan, Rahman (1996). Personal Communication.
- Kintsch, Walter (1990). The Personalized Journal. *Psychological Science*, 1(6), p. 345.
- Lagoze, Carl, Shaw, Erin, Davis, James, R., Krafft, Dean B. (1995). Dienst: Implementation Reference Manual. Cornell University Technical Report TR95-1514.
- Maly, K., French, J., Selman, A., & Fox, E. (1994). Wide Area Technical Report Service. *Proceedings of the Second International World Wide Web Conference*, Chicago, IL, October 19-21, 1994, pp. 523-533.
- Nelson, Michael L. & Gottlich, Gretchen L. (1994). Electronic Document Distribution: Design of the Anonymous FTP Langley Technical Report Server. NASA TM-4567.
- Nelson, Michael L., Gottlich, Gretchen L., & Bianco, David J. (1994). World Wide Web Implementation of the Langley Technical Report Server. NASA TM-109162.
- Nelson, M. L., Gottlich, G. L., Bianco, D. J., Paulson, S .S., Binkley, R. L., Kellogg, Y. D., Beaumont, C. J., Schmunk, R. B., Kurtz, M. J., Accomazzi, A. & Syed, O. (1995). The NASA Technical Report Server. *Internet Research: Electronic Network Applications and Policy*, 5(2), pp. 25-36.
- Nelson, Michael L. & Esler, Sandra L. (1997). TRSkit: A Simple Digital Library Toolkit. *Journal of Internet Cataloging*, 1(2), pp. 41-55.
- Nelson, Michael L., Maly, Kurt, Shen, Stewart N. T. (1997). Buckets, Clusters, and Dienst. Old Dominion University Computer Science Technical Report 97-30.
- Odlyzko, A. M. (1995). Tragic loss or good riddance? The impending demise of traditional scholarly journals. *International Journal of Human-Computer Studies*, 42, pp. 71-122.
- Okerson, Ann S. & O'Donnel, James J. (1995). Scholarly Journals at the Crossroads: A Subversive Proposal for Electronic Publishing. Washinton D. C.: Association of Research Libraries.
- Patterson, David A. (1994). How to Have a Bad Career in Research/Academia. Keynote Address at the First Symposium on Operating System Design and Implementation, Monterey, CA, November 14-17, 1994.
- Quandt, Richard E. (1996). Simulation Model for Journal Subscriptions by Libraries. *Journal of the American Society for Information Science*, 47(8), pp. 610-617.
- Roper, D. G., McCaskill, M. K., Holland, S. D., Walsh, J. L., Nelson, M. L., Adkins, S. L., Ambur M. Y., & Campbell, B. A. (1994). A Strategy for Electronic Dissemination of NASA Langley Technical Publications, NASA TM-109172.

- Schatz, B., Mischo, W. H., Cole, T. W., Hardin, J. B., Bishop, A. P., & Chen, H. (1996). Federating Diverse Collections of Scientific Literature. *IEEE Computer*, 29(5), pp. 28-36.
- Schatz, B., & Chen, H. (1996). Building Large-Scale Digital Libraries. *IEEE Computer*, 29(5), pp. 22-26.
- Scott, E. W. (1953). New Patterns in Scientific Research and Publication. *American Documentation*, 4(3), pp. 90-95.
- Smith, Ruth S. (1992). The NACA Collection: When Old is New Again -- A Status Report on the NACA Documents Project. *Sci-Tech News*, 46(2), pp. 15-20.
- Sobieski, Jaroslaw (1994). A Proposal: How to Improve NASA-Developed Computer Programs. NASA CP-10159, pp. 58-61.
- Van Heyningen, Marc (1994). The Unified Computer Science Technical Report Index: Lessons in indexing diverse resources. *Proceedings of the Second International World Wide Web Conference*, Chicago, IL, October 19-21, 1994, pp. 535-543.
- Van Steenberg, Michael E. (1994). NASA STELAR Experiment. *The Serials Librarian*, 24, pp. 135-151.