



2004 Information Format Trends: Content, Not Containers

In March 2003, OCLC released the *Five-Year Information Format Trends* report, a snapshot of trends in format purchase and usage. In January 2004, we published the *OCLC Environmental Scan*, a high-level view of the information landscape, designed to inform and stimulate discussion about future strategic directions. The trends identified and analyzed in the *Scan* made it clear that the focus on formats the library community has traditionally used as measures of collection fit and strength overlooks interesting new patterns in content distribution and use.

In the 18 months since we wrote the previous *Format* report, the rapid “unbundling” of content from traditional containers such as books, journals and CDs has had a significant impact on the self-search/find/obtain process. Digital content is often syndicated instead of being prepackaged and distributed, and access is provided on an as-needed basis to the information consumer by providers outside the library space.

This follow-up report to the 2003 version updates our predictions of format trends for material collected by libraries. But first, we look at the growing phenomenon of content being created, published and shared outside of the traditional structure of the library.

Contents

2 Freedom of choice

A sudden shift
McLuhan saw it coming
Content explosions
Smaller pieces, smaller payments
New voices all around
The new vocabulary

9 Format update 2003 to 2004

Popular materials
Scholarly materials

13 Context is King

16 Notes

Written and compiled by
OCLC Marketing staff

More than ever, content consumers are “**format agnostic**”¹ in that they do not care much what sort of container—such as a book, journal, blog or a Web page—the content comes from. Commercial content deployers are increasingly catering to these more experimental information consumers, providing content in a variety of formats, often with different cost structures related as much to its consumption as to the content itself.

For example, part of an interview with an author may be available in HTML and PDF format on a magazine’s Web site, but the full interview—“**premium content**”—may require registration and/or payment. Amazon’s “search inside the book” is another example of this tiered access structure. We may read a couple of pages of the content—to make sure we like it and that it’s relevant—but in order to read the entire item, the content must be purchased. And all of this happens not by navigating the containers of content in a physical space but by riding the surfboard of telecommunications over the sea of the Web to find the **perfect content wave**. Helping in the discovery process are increasingly sophisticated search engines and **data mining tools**.

More and more, the “format” is a communication device that moves from creator to consumer in channels completely outside of the traditional ones such as the library. Content is no longer **format-dependent** and users are not dependent on traditional distribution channels for access to content. This is true both in the realms of scholarly communication and popular materials. For libraries and content sellers, this means the processes of acquisition, organization and delivery of content need to change to accommodate the expectations of our communities.

A sudden shift

When we released the *Information Format Trends* report in 2003 many of these new content consumption trends were not yet clear. Traditional publishing, both scholarly and popular, showed signs of slowing, but **e-books** weren’t being adopted as fast as predicted in 2000, and the rate of Web growth, although immense, was slowing.

In 2004, the changes these movements foreshadowed are becoming clearer. Traditional **print publishing** is indeed slowing. E-books have entered the adoption phase. And good quality content is leaking out of its containers and making its way to the open Web, facilitated by self-publishing tools and communication devices and technologies.

For example, medical faculty at a large east coast university are **self-publishing** their research articles on personal Web sites and blogs so that Google will expose their **intellectual property** in the marketplace of ideas, in addition to publishing in more traditional print and digital journals. Stories such as this one are becoming more common. There’s a transformation underway in the world of content and its deployment.

56 million American adults are “wireless ready.” 28 percent of Americans within the past month have used a laptop that can connect wirelessly to the Internet or a cell phone that lets them send and read e-mail.

*Pew Internet Project Data Memo,
May 12, 2004*

Libraries have always faced a shifting world of materials, new models alongside old, the increasingly complex matrix of commodity and open-source publications—published, self-published and unpublished, paper and digital. Now they must also manage content that is unbound from any sort of identifiable container.

The **major trends** in the content space are not just technological. The major trends—and challenges—are social and are profoundly changing how content is created, collected, used, shared and preserved. As with other trends in the “infosphere,” many of the most disruptive changes are taking place outside of the arena of traditional information management. They are being driven by the manufacture of phones and entertainment devices as well as by consumers of content whose interests and tastes are supported by the technology.

To meet these challenges, contextual guides must be built into the “**search, find and obtain**” event and librarians will need to pay attention to how content is created, found and used by the increasingly self-sufficient, but also increasingly demanding and discerning information consumer/producer. The advent and huge adoption rates of **smartphones**, for example, is an indicator of a new era in ambient connectivity, where people can increasingly interact with other people, and information, and content providers from wherever they are, whenever they want.

McLuhan saw it coming

In his book *Understanding Media*, published in 1964, Marshall McLuhan declared, “**The medium is the message.**” The meaning of the phrase has been debated ever since. According to Mark Federman of the McLuhan Program in Culture and Technology at the University of Toronto, the “message” of any medium or technology is the **change of scale** or pace or pattern that it introduces into human affairs.² This tells us that changes in how content is made available, accessed and consumed indicate a new message, an effect of the new medium of ever-present Internet access.

Federman goes on to say, “With this early warning, we can set out to characterize and identify the new medium before it becomes obvious to everyone—a process that often takes years or even decades. And if we discover that the new medium brings along effects that might be detrimental to our society or culture, we have the opportunity to influence the development and evolution of the **new innovation** before the effects becomes pervasive.”³

“Text is the Internet’s **über-medium**,” says Clay Shirky, professor in New York University’s Interactive Telecommunications Program, “and with e-mail still the undisputed killer app, and portable devices like PDAs and cell phones relying heavily or exclusively on text interfaces, text is a leading indicator for other kinds of media. Books are not **sacred objects**, and neither are radios, VCRs, telephones or televisions.”⁴

The use of communication devices and networks to move **multimedia content**

To be young is to be cognitively welded to a mobile. “You always want it near you,” somebody says. “You take the phone out of your purse and leave your purse behind. You take your phone even when you don’t take your purse or your keys. It’s like a little person.”

Washington Post, July 31, 2002, Page C01

“It’s the experience of the content, not its mass-manufactured container, that I value most.”

John Blossom, Shore Communications, a research and advisory service for the content business

4 2004 Information Format Trends: Content, Not Containers

around worldwide is huge, and increasing rapidly. E-mail and handheld devices such as smartphones are being used by many in addition to or as a replacement for the traditional channels of content transportation and delivery—postal services and interlibrary loan for example—and much of this content is delivered at little or no apparent cost to either the sender or the receiver.

Content consumers will tolerate some costs for content they value but that value is increasingly related to control over the content: delivery options, filtering, personalization and convenience. New communication channels, then, become a **disruptive technology** in the content world, offering alternatives to established ways of obtaining content. One of the results is that the format of the content becomes less important than its ability to be delivered via a low-cost, convenient channel.

“Expand access to postal services by doing business when and where our customers prefer.”

A directive stated in the executive summary of *United States Postal Service Transformation Plan*, April 2002

Content explosions

E-mail to a friend

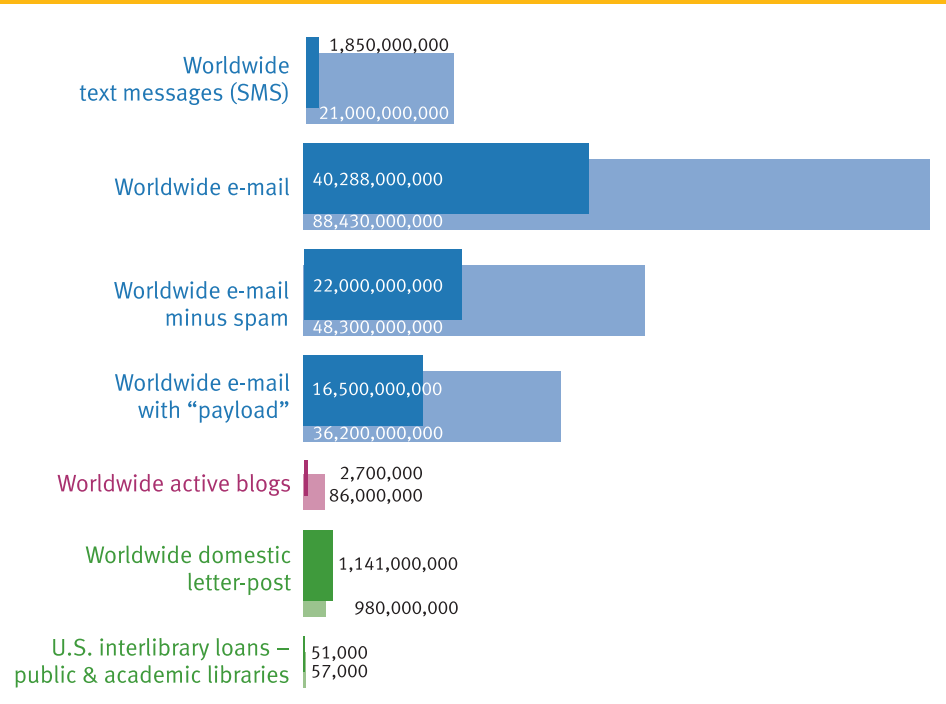
In 2002, the research group IDC estimated that 31 billion e-mails traveled the Internet daily. By 2006, IDC predicts 60 billion e-mails will be sent each day.⁵ Current estimates put e-mail spam as high as 40 or 50 percent of that traffic but if we exclude possible spam, that still leaves about 22 billion e-mail messages delivered to e-mail accounts that have content meaningful to the recipients.

It comes to you.

Projected shifts in average daily volume of content

2004–2009

Click on the bars in the graph for related research data. The darker bars represent activity in 2003–2004. The lighter bars represent our projections for activity for 2009. References for sources may be found in Notes.



Bar graph values for each row in this chart are scaled only within respective rows, not across the entire chart.

Gartner, a U.S.-based provider of research and analysis on the global IT business, has estimated that as much as 75 percent⁶ of U.S. enterprise e-mail is “**payload**”⁷—that is, the message is part of a high level knowledge exchange, often carrying attachments or URLs to content. This could suggest the number of e-mails carrying content might be around 16 billion per day. In contrast, content moving through the more traditional **distribution channel** of ILL accounts for far less activity. In 2000, the U.S. Dept of Education reported that U.S. ILL traffic for both books and journal articles loaned in public and academic libraries, per day, was about 51,000 items.⁸

Libraries do send content by e-mail in the form of attached articles; however, e-mail is not tracked like ILL transactions in libraries, and so statistics related to this content distribution channel do not exist. Libraries miss an opportunity to report on significant content traffic that may be flowing to and from the library in new containers.

By 2007, **mobile phone** service adoption is predicted to be 82 percent in the U.S., 92 percent in the U.K., 82 percent in Australia and 75 percent in New Zealand.⁹ Smartphones can be used as media players, text readers, digital wallets, remote controls and PDA organizers.¹⁰ An estimated 57 million camera phones were sold worldwide in 2003.¹¹ Worldwide camera phone sales for 2004 are predicted to be a quarter of all mobile phone sales.¹²

Worldwide, almost 2 billion SMS (Short Message Service) text messages are sent daily using mobile devices. **Young people** for whom the Internet is more than a source of direct entertainment send many of those messages; they also make heavy use of the Web to find local information such as club listings, movies and local restaurants, and to keep in touch with friends.

The huge significance of this **convergence of technologies** is that people are not tied to a computer for the delivery of content and a major social change is underway as that content is inextricably woven into the context of people’s lives. Another significant consequence is that mobile communication devices are much more affordable than computers in developing countries and so will bring in new content consumers far more rapidly than in the past.

Joi Ito, a venture capitalist and well-known blogger, wrote in an essay: “Users are shifting their attention away from packaged content to **social information** about location, presence and community...Mobile communication tools are shifting away from a 1-to-1 model, allowing for increased **many-to-many interactions**... communication carriers, hardware manufacturers and content providers must understand and build models that focus less on the content and more on the relationships.”¹³

Smaller pieces, smaller payments

The technologies of communication that move content around the Internet increase users’ expectations for “delivery at the item, document or even fact level, just in time, in the desired format for the need at hand.”¹⁴ Content companies react by reducing content to smaller and smaller consumable units. Cost is often “**downsized**” as well. Micropayment for microcontent is increasingly common.

“**What I want—what I need—is a camera phone that would let me take a picture of an article in a newspaper and translate it into a machine-readable form that I could send to a colleague, or, indeed, myself.**”

Michael Schrage, *Technology Review*,
Feb. 2004

“**India’s wireless market is undergoing a boom...with current wireless subscribers totaling more than 34 million, the country is expected to see that number grow to more than 150 million in the next five years.**”

“Boom Times for India’s Wireless Market”
In-Stat MDR white paper

Reading “an encyclopedia in a public library, selling a geometry textbook to a friend, copying a song...will be rerouted through a **system of micropayments** in return for which the rights to ever smaller pieces of our culture are doled out. ‘Sooner or later,’ predicts Miriam Nisbet, the legislative counsel for the American Library Association, ‘you’ll get to the point where you say, “Well, I guess that 25 cents isn’t too much to pay for this sentence,” and then there’s no...going back.’”¹⁵

Just a few years ago, no one would have thought of cell phone ring tones or cell phone “wallpaper” or single tracks of music as **big business**. But they’re now part of a \$3.2 billion microcontent industry, with innovative transactional purchase and subscription models according to Outsell, the content research and analysis company.¹⁶ According to Gartner, in Western Europe during 2003, ring tones, logos and screen savers (RTLS) downloaded to wireless devices generated 570 million euros. Other mobile applications such as games accounted for 330 million euros.¹⁷ Music providers, from the Smithsonian to Sony to Wal-Mart, are making single titles available for 99 cents or less. Many of the downloaded songs end up on Apple iPods that are being used for more than music. “It’s the limousine of the spoken word,” says Audible CEO Don Katz in a July 18, 2004 *PR News* wire report.

Pieces of microcontent do not currently figure large in the collections of libraries but such clear interest from consumers suggests there is an appetite for microcontent from a variety of content sources, and a willingness to pay for the convenience of having it delivered to personal devices—iPods, laptops, PDAs and phones. In a sense, the many “ask-a” services, including Google Answers, are delivering microcontent in the form of **answers to questions** for payment. Libraries need to find ways to deliver quality content to mobile devices.

New voices all around

One intriguing aspect of consumer interest in smaller pieces of content is described by Outsell in a February 2004 report, *Power to the People—The Rise of Social Publishing*, as “a growing atomization of content interests and the resulting publications that serve them” that is “a function of several colliding forces: an increasing demand for ways for individuals to create and distribute content, the increasing acceptance of information from ‘nonauthoritative’ sources and the resulting **commoditization of content** and the growing array of technology that underpins the personal and social publishing phenomenon.”¹⁸

Social publishing is essentially open content unfettered by licenses or firewalls or passwords, widely disseminated and available. It needs no other distribution channel other than the Internet. The individual is at the center of social publishing, not institutions.

As forms of social publishing, **wikis** and **blogs** are indicators of further change in the information landscape that could lead to a new **publishing paradigm**. Together with mechanisms to syndicate content—such as RSS and Atom—wikis, moblogs and blogs have enabled ordinary people to gain the power of publishing.¹⁹ Blogs and wikis could be a natural way for librarians and libraries to reach out to their communities—and perhaps more importantly, hear back from them.²⁰

“iPod therefore iAm”

Cover headline,
Newsweek July 26, 2004 issue

“I use weblogs to find information more than I use Google.”

Rael Dornfest, in “View from the Alpha Geek,” *Technology Review*,
Feb. 13, 2004

Tim Berners-Lee, father of the Web, might have been the first blogger in 1992 when he posted a “What’s New” page that had hot links to other pages.²¹ Since then, over 4 million blogs have been created, according to the Perseus Blog Survey (<http://www.perseus.com/blogsurvey/thebloggingiceberg.html>). “Rather than spawning a million **micropublishing empires**, weblogs are becoming a vast and diffuse cocktail party, where most address not ‘the masses’ but a small circle of readers, usually friends and colleagues. This is mass amateurization, and it points to a world where participating in the conversation is its own reward.”²²

Blogads, a Web ad network, recently surveyed 17,159 blog readers (<http://www.emarketer.com/Article.aspx?1002860>). “Far from being young kids with little money in their pockets and lots of time on their hands, the survey found that blog readers are older and richer than many people suppose. Exactly 61 percent of the blog readers that responded to the survey are over the age of 30, and 75 percent make more than \$45,000 a year. In fact, nearly 30 percent of the respondents are between the ages of 31 and 40, and over 37 percent spanned the ages of 41 to 60. And nearly 40 percent have a household income of \$90,000 or higher.”

As Henry Copeland, author of the report and CEO of Blogads, summed up: “86 percent say that blogs are either useful or **extremely useful** as sources of news or opinion. 80 percent say they read blogs for news they can’t find elsewhere. 78 percent read because the perspective is better. 66% value the faster news. 61 percent say that blogs are **more honest**.”

As information is increasingly integrated into workflow, the distinction between the medium and the message becomes even more blurred. “The future of libraries is being shaped today by **emerging technology** that is transforming the way information is created and disseminated...More and more, users want granular pieces of information and data, at the moment of need, in the right format...The mantra will be: ‘Everything, everywhere, when I want it, **the way I want it**.’”²³

It is clear that librarians, as experts in providing context for storehouses of traditional content, need to find ways to fit into a world where content and the channels to distribute it are **ubiquitous**.

The new vocabulary

Blog: “A Web page consisting of frequently updated, chronological entries on a particular topic.” The word is the shortened version of “weblog.” (<http://www.wordspy.com/>)

Blogosphere: “The collection of all bloggers, blog sites, blog readers and blog text.” (<http://www.wordspy.com/>)

Blogroll: A listing of Web sites that often appear as links on weblogs. This list of links is used to relate the site owner’s interest in or affiliation with other bloggers. Blogrolling is a brand name converted to everyday usage, much like Google or Kleenex. (<http://www.blogrolling.com/>)

Dayparting: The practice of creating different newspaper content for different parts of the day, for different audiences. Some online versions of print newspapers

“Blogging will play an increasing role in knowledge sharing...”

“InfoAboutInfo Briefing 6:32,”

Outlook 2004:

Issues in the Information Marketplace,
Outsell, Dec. 19, 2003. p.8

“61 percent of the blog readers that responded to the survey are over the age of 30, and 75 percent make more than \$45,000 a year.”

8 2004 Information Format Trends: Content, Not Containers

are diverging from their print counterparts as a way to better reflect the interests of their readers. For example, the online version of the *Seattle Post-Intelligencer* publishes news and traffic alerts in the morning; and consumer reports, media and restaurant reviews, and club listings in the evening. Which can lead to the next “buzz.”

Digital Swarming: The use of text messaging and mobile phones/PDAs to notify people of a hot event. Search Google for this term and you’ll find a story about poor Prince William being digitally swarmed every time he tries to go out. Young women track him and send messages to friends. By the time he arrives at a pub, for example, a horde of Willie Watchers is waiting.

Fleshmet: An old but revived term. Fleshmet is the past tense of fleshmeet, which means to meet a person in the flesh rather than virtually. As in “Have we fleshmet?” Gartner predicts that by 2012, “70 percent of the populations of developed countries will spend 10 times longer per day interacting with people in the e-world than in the physical one.” (“New Technologies Will Change the Way We Manage Information,” *Research Note*, 12 December 2003)

Moblogs: “Moblog” is a contraction of mobile and weblog. A mobile weblog, or moblog, consists of content posted to the Internet from a mobile device, such as a cellular phone or PDA. Much of the earliest development of moblogs occurred in Japan, among the first countries in the world to have mobile phones with built-in cameras.” (<http://en.wikipedia.org/wiki/Moblog>)

Nanopublishing: “An online publishing model that uses a scaled-down, inexpensive operation to reach a targeted audience, especially by using blogging techniques.” (<http://www.wordspy.com/>)

Payload: This is not a new term. The use is new. Payload refers to things carried as part of a task or mission. No doubt you’ve heard NASA use the term to refer to “payload” for space flights. Payload also refers to content carried by digital tasks—order forms, for example. But, what seemed new to us was the use of the term to refer to attachments to e-mail.

Smartphone: An all-in-one device that combines the flexibility and functionality of a hand-held computer with the communications ability of a mobile phone. Functionality may include Web browsing, wireless e-mail, fax, camera, scheduling software and the ability to read PDF and Word documents. An example is the Handspring Treo™.

Snam and Spim: We know that Spam is unwanted e-mail. Snam and Spim are mutated spam. Snam is unwanted e-mail generated by social networking (e.g., Friendster, Orkut) activities. A simple, older example of snam is the e-mail you get from a friend of the “send this to 12 people you know or bad luck will follow” variety. Spim is spam and snam received via instant messaging.

Wikis: A wiki is a “group blog” that allows a group of people to build, edit and modify a Web site with no programming or HTML knowledge. Because it doesn’t require technical expertise, all users of the wiki have equal ability to maintain and edit the site. Wikis are a collaborative space for projects involving participants who need not be geographically close.

Popular materials

While there have been no major trend changes since our 2003 format report with regard to production, consumer consumption continues to shift from print to all things digital. *Reading At Risk*, an NEA report released in July 2004, merely “documents and quantifies a huge cultural transformation...a massive shift toward electronic media for entertainment and information.” (Dana Gioia, NEA Chairman, quoted in *The Washington Post* July 9, 2004 p. C02) But with publication of 175,000 new print titles in the U.S. in 2003,²⁴ it’s clear that print won’t disappear anytime soon.

In the U.S., 23 million fewer **new print books** were sold in 2003 than in 2002. According to a Book Industry Study Group report, the 2003 figures show a continued trend of increasing production and declining demand.²⁵ The **magazine industry** finished 2003 with a loss. Unit sales dropped from about 2.2 billion copies in 1992 to 1.5 billion in 2002, according to Harrington Associates, a

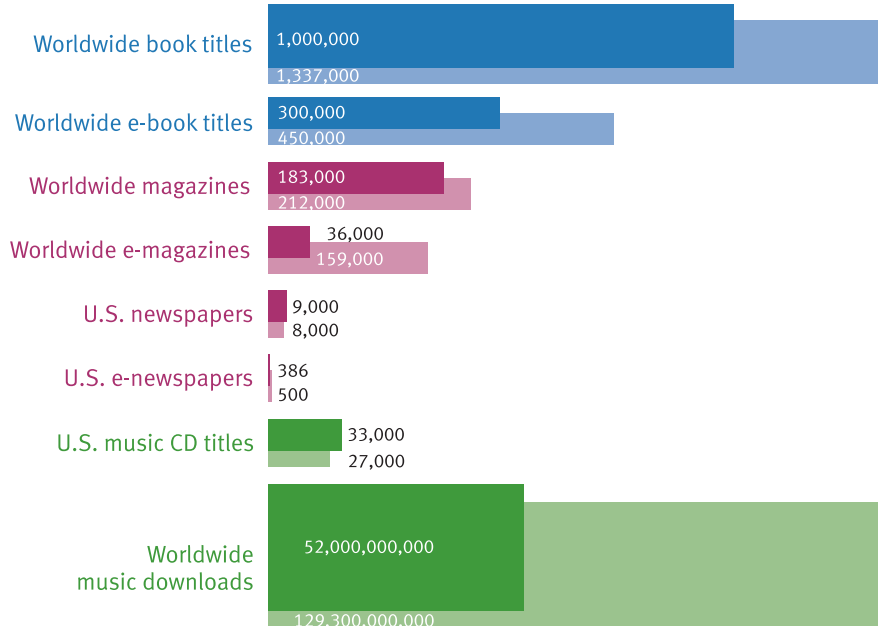
“The really broad implication for e-books is the path into the classroom.”

Arthur Klebanoff, Rosetta Books LLC in *Wall Street Journal*, Jan. 12, 2004

Projected shifts in annual production volume

2004–2009

Click on the bars in the graph for related research data. The darker bars represent activity in 2003–2004. The lighter bars represent our projections of activity for 2009. References for sources may be found in Notes.



Bar graph values for each row in this chart are scaled only within respective rows, not across the entire chart.

newsstand consulting business.²⁶

The good news is “that **eBooks** are the fastest growing segment of the publishing industry. In the first quarter of 2004, eBook sales increased 46 percent to 421,955 units compared with 288,400 in the first quarter of 2003.”²⁷ “Thanks to the spread of MP3 players, **digital audio books** are beginning to sell.”²⁸

E-book adoption is on the rise in Asia. China and Japan are “rolling out e-book products to their huge base of mobile phone users...”²⁹ unhampered by copyright limitations.

A January 2004 search of *Ulrich's* revealed that there are about 183,000 active **serial titles**. Just under 20 percent of those magazine titles are available digitally, but a recent study found that among the 25 large public libraries surveyed, over 64 percent of their titles were available electronically.³⁰ This suggests public libraries are selecting journals that can be delivered digitally.

A study by the Newspaper Association of America found the teens surveyed thought the teens depicted reading print newspapers were “nerdy.”³¹ The study goes on to suggest that newspapers need to develop digital versions that are not mirrors of the print versions.

Online video companies like Netflix, Warner and HBO are making **DVDs by mail and pay-per-view** very attractive.³² One industry expert predicts that 20 million people will have video on demand by the end of 2004. “On-demand services are the future of entertainment delivery. CDs, DVDs and any other forms of physical media will become obsolete.”³³

In a BBC interview a *Billboard Europe* executive predicts “...in three years from now, there will be [no] physical singles available in the U.S. [...] The Internet allows consumers just to pick the one song they want.”³⁴

Scholarly materials

In the North American academic library environment, where budgets remain a major concern, some of the trends are less trendy than they are facts of life, having been challenges for many years.

ARL libraries' monographic expenditures declined slightly in 2002, while spending on electronic resources accounted for 20 percent of acquisitions dollars.³⁵ A recent study shows that almost 41 percent of the academic libraries sampled plan to “aggressively” reduce spending for print and increase expenditures for electronic resources.³⁶

University presses in the U.S. continue to struggle. In the spring of 2004 the University of Idaho and Northeastern University announced their presses would close due to financial pressures.³⁷ As of late summer 2004, neither press has closed although operations have been suspended at the University of Idaho Press and Northeastern University Press is looking for ways to reduce costs.³⁸ These challenges are symptomatic of the sea change in scholarly publishing.³⁹ Although monographs remain an important venue for scholars in some disciplines, print-on-

“Coming Soon to Your Cell Phone: Just About Everything”

Billboard article, January 10, 2004, p. 37

“The traditional scholarly communications model is not sustainable any longer for libraries.”

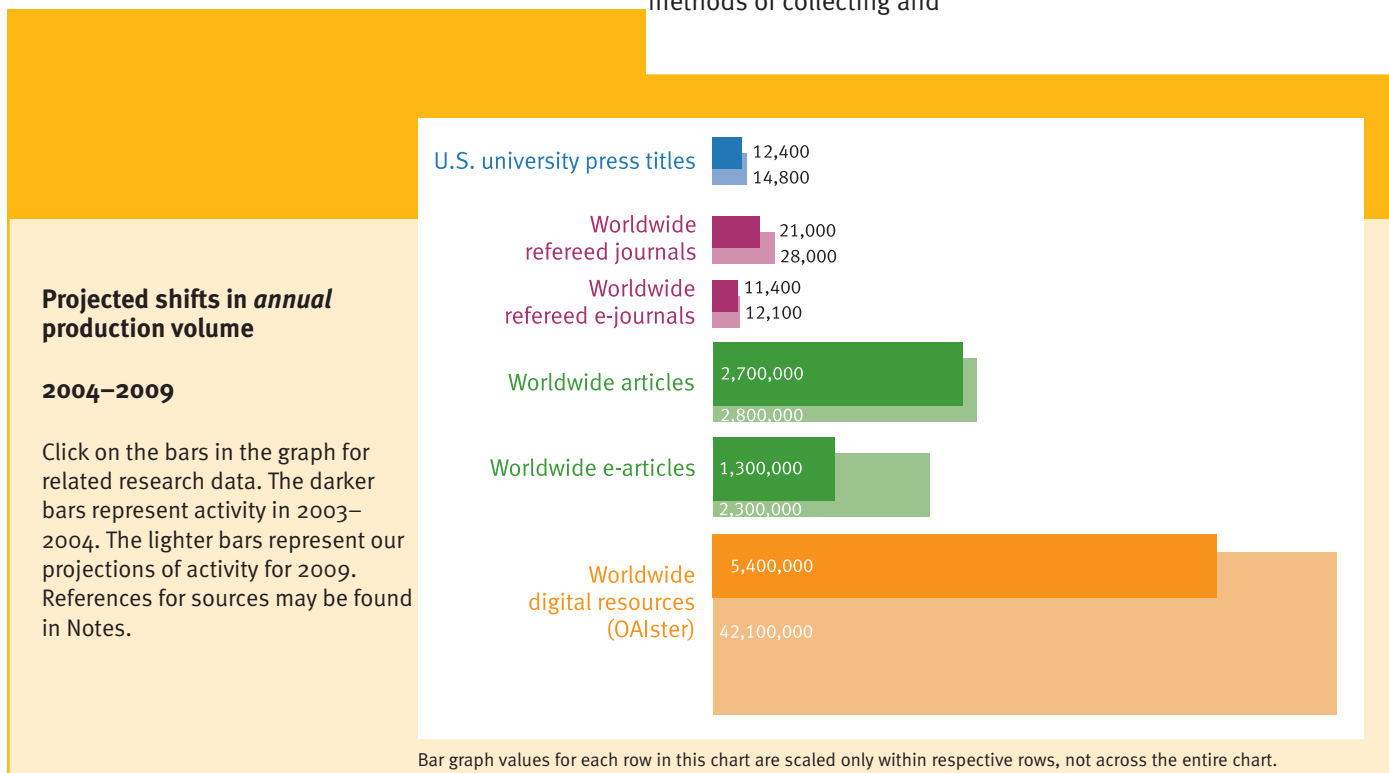
Doris Helfer, “Leading Libraries”
Searcher, March 2004, p.27

demand and e-books may prove to be the salvation of publication. In 2003, 12,000 titles were published by U.S. university presses. Given the very modest increases in production between 1993 and 2003⁴⁰ it's unlikely that traditional print university press publications will increase much beyond the 2003 total. As more academics publish on the open Web, in open access repositories and under Creative Commons licenses, the infrastructure of scholarly publishing will continue to change, and scholarship will find new distribution channels.

A search of *Ulrich's* in April 2004 yielded a count of about 44,000 active scholarly journals, with about 21,000 active refereed scholarly journals, 15,000 active scholarly e-journals and 11,000 refereed, scholarly e-journals. Carol Tenopir, professor at the **School of Information Studies at the University of Tennessee**, notes that although "these numbers are a moving target," she can say "with confidence that as of the end of 2003 there [were] just under 50,000 scholarly journals and somewhere between one-third and just over one-half of them are in digital form."⁴¹

While it's not known how much total scholarly output will increase in the next few years, a recent National Science Foundation report, *Knowledge Lost in Information*, argues that "digital resources" could easily double research output in the next decade.⁴² **OAlster**, University of Michigan's OAI harvester that is currently harvesting from about 80 e-print repositories, shows that text items (mostly e-prints) increased 41 percent from 2002 to 2003.⁴³

A Johns Hopkins Scholarly Communications Group study (<http://openaccess.jhu.edu/threats.cfm>) estimates that by 2015 ARL libraries' costs to support their current subscriptions could total as much as \$1.96 billion per year, with "individual libraries paying nearly \$16 million each year just to maintain journal collections at current levels."⁴⁴ One thing is certain: libraries' methods of collecting and



disseminating scholarly output will be different in five years.

EDUCAUSE reports that “more than 70 percent of institutions expect to increase the number of e-learning course offerings in the next year.”⁴⁵ However, there’s stiff competition for students. A for-profit e-learning company, **Corinthian Colleges’** total student population rose to 66,239 by March 31, 2004, an increase of 52.7 percent compared with 43,387 students on March 31, 2003.⁴⁶

There is an increase in activities supporting a coordinated management and disclosure of **digital assets** of institutions—learning objects, data sets, e-prints, theses, dissertations and so on in digital, institutional repositories. Over 50,000 electronic theses and dissertations (ETDs) have been gathered into the **Networked Digital Library of Theses and Dissertations** (NDLTD)(www.ndltd.org/), more than double the count last year, and probably over 100,000 ETDs now exist worldwide.⁴⁷

The **content “behaviors”** of young people—both students and faculty—have changed a great deal, and the institutions supporting their research and learning for the most part have not changed to accommodate the newer members of this community. Some have.

Cornell University’s Internet-First University Press will use DSpace to let scholars download a book or article for free. For a fee, the press will print and bind the book or article on demand. The press will also offer videos—free, low-quality streaming or DVDs for purchase.⁴⁸

Yale, Duke, Wake Forest and the University of Colorado at Boulder are testing a pilot service from Cflix, that allows students to download locally stored film clips. According to Charles Powell, Yale’s director of academic media and technology, students can “bypass trips to screening rooms or library reserves.”⁴⁹


Finally, one of the most interesting developments was the July 2004 announcement that **Duke University** will provide Apple iPods to each of its 1,650 incoming freshmen. According to **Wired News**, the school plans to preload the digital music players with orientation information, academic schedules—even the Duke fight song—and will make lectures, audio books and other class-related content available from a Web site modeled after Apple’s iTunes music store. The school expects additional ideas for the technology integration to come from faculty, students and the library.

“Without a presence on a course site, the library has a diminished role in the teaching and learning process.”

netconnect, Winter 2004

“If OA (open access) works, we can expect library budgets to be smaller, not larger...”

David Goodman, *liblicense-I*, March 2, 2004



Historically, libraries have been the unparalleled **collectors of content**, and for many reasons: their mandate to protect collections that reflect local communities; the necessity of a single place to find and obtain information; and because, frankly, no one did it better. Today, however, none of these statements is exclusively true. The “just-in-case” community collection is no longer adequate and consumers of content expect a great deal more personalization and dynamism in their content experiences.

Library content was, and still is, the **gold standard**: the best content money can buy on behalf of an identifiable audience. But it is no longer enough to present a warehouse of content and expect community members to create their own personalized **meaningful context**, post hoc, out of the raw materials. Others in the content market have read the oracle’s tea leaves and so provide syndicated and scoped content with personalization features that make perhaps inferior content very attractive to an ever more demanding, format-agnostic information producer and consumer.

What seems clear is that libraries should move beyond the role of collector and organizer of content, print and digital, to one that establishes the **authenticity and provenance** of content and provides the imprimatur of quality in an information-rich but context-poor world. The challenge is how to do this. The best way to adapt is to understand what’s forcing the change.

Research suggests that end users see the most important role for their libraries as making content available in the user’s **digital workspace**, regardless of what devices are in that space.⁵⁰ The networked **ambient environment** will support “tasks...on the appropriate computing devices and will be available anywhere, anytime. The sources of information and tools will be abstracted, much as the power plants that provide electricity and the reservoirs that provide water are invisible to the consumer. Web Services, XML and WiFi and other such technologies form the foundation for this virtualized environment. While it is not yet clear how this **marriage of technology and content** will play out, it is clear that those that have not moved to XML and Web Services will be locked out of a key channel of distribution. XML and Web Services are not options—they are imperative.”⁵¹

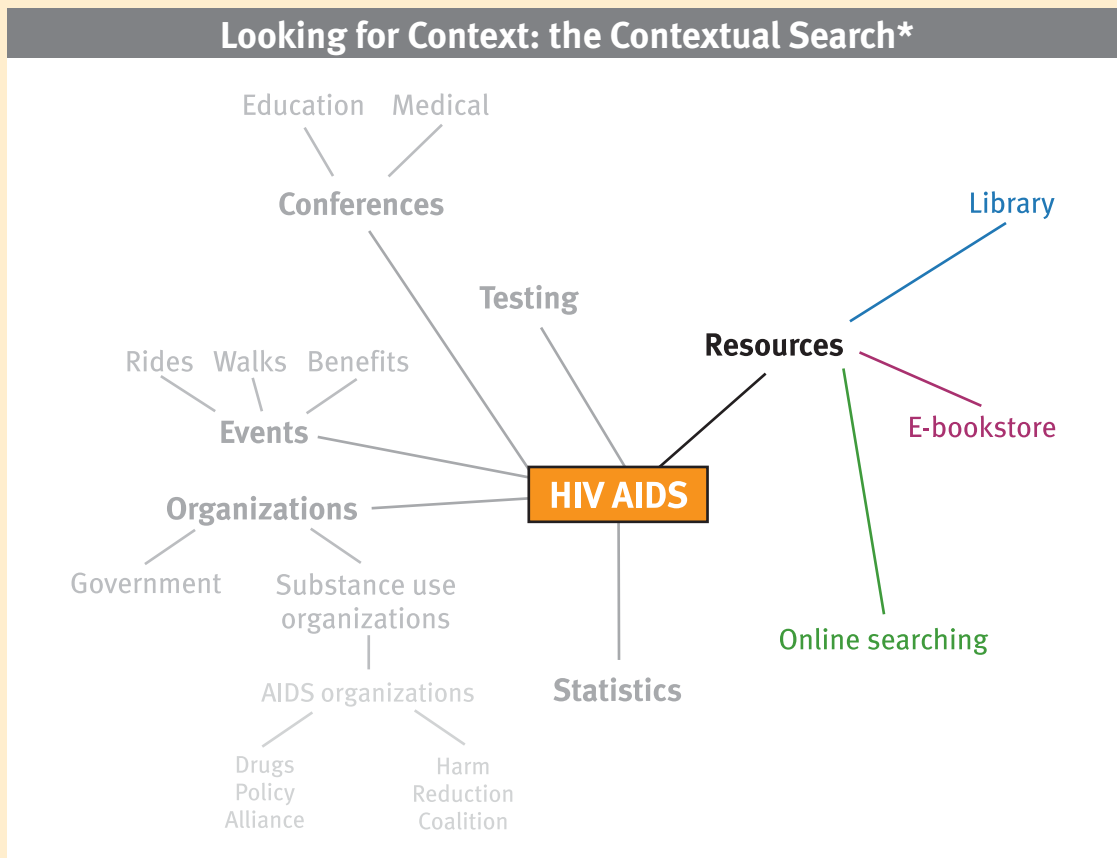
The “library’s role as archive or steward of information goods is being transformed as a collaborator and, potentially, a catalyst within **interest-based communities**.”⁵² We are at a crossroads. Technology and culture have come together to foster a transformation in the world of content. This new world is abundant and unstructured, but **contextual mechanisms** for navigating and synthesizing the information commons are scarce, even in—perhaps especially in—libraries. “We are drowning in information but are starving for knowledge. Information is only useful when it can be located and **synthesized into knowledge**.”⁵³

Content in context: synthesizing content into knowledge

In the past, much formal information seeking was done within the confines of the library building, and librarians had the opportunity to assist in the location and synthesizing of information into knowledge. In a society where information and content are wherever there is an Internet-accessible device, the information seeker is self-sufficient, navigating and making choices with no expert guides. The search engines that dominate the Web now—Google, Yahoo! Search—seem miraculous in that they can locate thousands of relevant items out of millions, but each searcher must determine relevance by examining each retrieved Web entity. It's not unlike asking a library user to examine each book, each article, each

government publication that contains the words “HIV AIDS” to find the few that satisfy the knowledge need.

What is needed is context. We no longer lack content but we do lack context. There are plenty of efforts being made to address this gap by companies that see the need and, quite rightly, we predict, see the context challenge as the next ‘big thing.’ A contextual search tool might work like an invisible librarian, searching the Web and bringing back, in real time, news, book and article links, relevant Web sites, and any documents and e-mails related to the topic on the searcher’s computing devices.



“WebBrain lets you search the Web visually, so you can explore a dynamic picture of related information.”

WebBrain™ is one of several companies presenting content in context, grouped by topic or category.

* Representation of a search for “HIV AIDS” similar to results from WebBrain™, www.webbrain.com

Content in context: synthesizing content into knowledge

The static depiction of a dynamic search page below illustrates how a visual search tool contextualizes Web documents and links a user to related information and content. Here, just as a Google™ window for search engine results might link to

relevant Web pages and an Amazon.com® window could lead a user to retail books and videos, so might a WorldCat® window link to an array of library-owned content on the desired subject.

View the interactive version at:
www.oclc.com/reports/2004contextdemo.htm

The central window, titled "Content in Context", displays a mind map centered on "HIV AIDS". The map branches into several categories:

- Education**
- Medical**
- Conferences**
- Testing**
- Resources**
- Statistics**
- Events** (sub-categories: Rides, Walks, Benefits)
- Organizations** (sub-categories: Government, Substance use organizations, AIDS organizations, Drugs Policy Alliance, Harm Reduction Coalition)

Three smaller windows are shown, each linked to a category in the mind map:

- WorldCat®** (linked to "Resources"):
 - Fiction books—11
 - Non-fiction books—8,430
 - Biographies—10
 - Dissertations—1,434
 - Music sound recordings—4
 - Sound recordings—59
 - Visual materials—432
 - Journal articles—13,500
- amazon.com.** (linked to "E-bookstore"):
 - Visual materials—100
 - Books—17,500
 - Kids books—163
- Google** (linked to "Online searching"):
 - Web images—36,800
 - Web pages—7,000,000

At the bottom, three screenshots illustrate search results:

- musicroom.com**: Search results for "John Cage: Symphony No. 1 (Full Score)".
- HIV/AIDS Statistics in Hong Kong 1984 - March 2004 (N=2311/676)**: A line and bar chart showing trends over time.
- amazon.com**: Search results for "John Cage: Symphony No. 1 (World Premier Recording) - David Barron".

* From Amazon and Google searches for HIV AIDS resources, conducted July 19, 2004.

1. Stephen Abram and Judy Luther, "Born with the Chip," *Library Journal*, May 1, 2004, <http://www.libraryjournal.com/article/CA411572> (viewed July 18, 2004).
2. Mark Federman, "What is the Meaning of the Medium is the Message?" n.d., http://www.mcluhan.utoronto.ca/article_mediumisthemessage.htm (viewed July 18, 2004).
3. Ibid.
4. Clay Shirky, "We (Still) Have a Long Way to Go" [blog], June 2000, http://www.shirky.com/writings/loc_book.html, (viewed July 18, 2004).
5. As reported by Nick Farrell, *Vnuned News*, September 30, 2002, <http://www.vnuned.com/news/1135485> (viewed July 18, 2004).
6. "Does E-Mail Get the Respect it Deserves?" *Research Note*, Gartner, May 2, 2000.
7. "In cargo transport, payload is the valuable contents of the vehicle. In communications and information science, the payload...is the data being processed or transported." <http://en.wikipedia.org/wiki/Payload> (viewed July 19, 2004).
8. United States Department of Education, National Center for Education Statistics (NCES). "Total circulation, document delivery and interlibrary loan transactions in academic libraries, by state: 2000," *Academic Libraries: 2000*, Table 1A, p. 19, <http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2004317> (viewed July 2004); and "Number of public library services and library services per capita or per 1,000 population, by type of service and by state: Fiscal year 2001," *Public Libraries in the United States: Fiscal Year 2001*, Table 4, p. 22, <http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2003399> (viewed July 19, 2004).
9. *Library Futures Quarterly* (4:1), Spring 2004, p. 12. Gartner is indicated as the source although no reference is given.
10. *Library Futures Quarterly* (4:1), Spring 2004, p. 12.
11. Michael Schrage, "VIPs: Virally Interactive Pixels," *Technology Review* (107:1), February 2004, p. 20, <http://www.technologyreview.com/articles/schrage0204.asp> (subscription required).
12. "Worldwide Camera Phone Sales to Reach Nearly 150 million in 2004, Capturing 29 Billion Digital Images," New release, Infotrends Research Group, <http://www.infotrends-rgi.com/home/press/itpress/2004/3.11.04.html> (viewed August 19, 2004).
13. Joi Ito, Essay about trends [blog], April 24, 2004, http://joi.ito.com/archives/2004/04/24/essay_about_trends.html (viewed July 18, 2004).
14. "Trend Alert: The Future Of Libraries," *InfoAboutInfo Briefing*, Outsell, Vol. 7, January 9, 2004, p. 6.
15. "Tyranny of Copyright," *The New York Times Magazine*, January 25, 2004, p. 43.
16. "Content Vendor Best Practices: Busting Up Fee Vs. Free," *InfoAboutInfo Briefing*, Outsell, Vol. 7, February 13, 2004, p. 25.
17. "Attitudes to Mobile Applications in Western Europe, 2003" (Executive Summary), Gartner, March 30, 2004.

18. "Trend Alert: Power to the People—The Rise of Social Publishing," *InfoAboutInfo Briefing*, Outsell, Vol. 7, February 6, 2004, p. 2.
19. "Trend Alert: Power to the People," p. 16.
20. Jeffrey Pomerantz and Frederic Stutzman, "Lyceum: A Blogosphere for Library Reference," n.d., (posted on ResourceShelf, February 29, 2004), <http://www.ils.unc.edu/~jpom/conf/Lyceum.pdf> (viewed July 18, 2004).
21. Anton Zuiker, "Blogs—A Short History," [blog] *Bloggling 101*, February 2004, <http://www.unc.edu/~zuiker/bloggling101/> (viewed July 18, 2004); Tim Berners-Lee's page is here: <http://www.w3.org/History/19921103-hypertext/hypertext/WWW/News/9201.html> (viewed July 18, 2004).
22. Clay Shirky, "Weblogs and the Mass Amateurization of Publishing" [blog], October 3, 2002, http://www.shirky.com/writings/weblogs_publishing.html (viewed July 18, 2004).
23. "Trend Alert: The Future Of Libraries," *InfoAboutInfo Briefing*, Outsell, Vol. 7, January 9, 2004, p.2.
24. "U.S. Book Production Soars to 175,000 New Titles in 2003; Trade Up, University Presses Down," Press release, Bowker, May 27, 2004, http://www.bowker.com/press/2004_0527_bowker.htm (viewed July 19, 2004).
25. As reported by Hillel Italie. "New Study Shows Big Drop in Book Sales," Associated Press, May 12, 2004, http://www.boston.com/business/articles/2004/05/12/new_study_shows_big_drop_in_book_sales?mode=PF (viewed August 18, 2004).
26. Susan Thea Posnock, "Winning in a Losing Battle," *Folio: First Day*, April 1, 2004, http://foliomag.com/circulation/marketing_winning_losing_battle/ (subscription required).
27. "Booksellers, Libraries See Growth in eBooks," InternetWeek.com, July 19, 2004, <http://www.internetweek.com/story/showArticle.jhtml?articleID=23901985> (viewed August 19, 2004).
28. Jim McClellan, "Under the Cover Story," *The Guardian*, [Guardian Unlimited Online] Thursday, January 29, 2004, <http://www.guardian.co.uk/online/story/0,3605,1133013,00.html> (viewed July 19, 2004).
29. *Library Futures Quarterly*, 3:4, Winter 2003, p. 2.
30. Peter McCracken, "A Comparison of Print and Electronic Journal Holdings in Academic and Public Libraries," *Libri* (53), 2003, p. 239.
31. Jennifer Saba, "Study: Teens Have Little 'Attachment' to Newspapers," *Editor & Publisher*, http://www.editorandpublisher.com/eandp/search/article_display.jsp?vnu_content_id=1000490563 (viewed July 19, 2004).
32. Christopher Null, "How Netflix is Fixing Hollywood," *Business 2.0*, July 1, 2003, pp. 41–43, <http://www.business2.com/b2/web/articles/0,17863,515760,00.html> (viewed July 19, 2004).
33. Josh Bernoff, Forrester Research, "From Discs to Downloads," as reported by Information Technology Association of America press release, September 2003, <http://www.itaa.org/isecc/pubs/e20039-03.pdf> (viewed July 19, 2004).
34. "CD singles to go 'in three years,'" BBC News, World Edition, January 15, 2004, <http://news.bbc.co.uk/2/hi/entertainment/3396887.stm> (viewed July 19, 2004).
35. Martha Kyrillidou and Mark Young, ARL Statistics: 2001-2002: *Research Library Trends*, <http://www.arl.org/stats/arlstat/02pub/intro02.html> (viewed July 19, 2004).

18 2004 Information Format Trends: Content, Not Containers

36. Primary Research Group, *The Survey of Academic Libraries*, 2004, Press release, PRWeb, March 2004, <http://www.prweb.com/releases/2004/3/prweb112699.htm> (viewed July 19, 2004).
37. Peter Monaghan, "U. of Idaho Press to Shut down, and U. of Georgia Press to See State Support Halved," *Chronicle of Higher Education*, March 12, 2004, n.p.
38. Association of American University Presses, "Universities and Their Presses in Hard Economic Times," Press Release, April 2004, <http://aaupnet.org/aboutup/hardtimes.html>.
39. Cathy N. Davidson, "Understanding the Economic Burden of Scholarly Publishing," *Chronicle of Higher Education*, October 3, 2003, n.p.
40. "U.S. University Press Book Production" [chart], statistics compiled by Andrew Grabois, R.R. Bowker, Inc., May 27, 2004, <http://www.bookwire.com/bookwire/> (viewed August 18, 2004).
41. Carol Tenopir, "Online Scholarly Journals: How Many?" *Library Journal*, Feb. 1, 2004, <http://www.libraryjournal.com/article/CA374956> (viewed July 19, 2004).
42. *Knowledge Lost in Information: Report of the NSF Workshop On Research Directions for Digital Libraries*, June 15–17, 2003, Chatham, MA, National Science Foundation, 2004, p. iv, <http://www.sis.pitt.edu/~dlwkshop/report.pdf> (viewed July 19, 2004).
43. Personal e-mail message from Kat Hagedorn, OAlster, University of Michigan, February 5, 2004.
44. Scholarly Communications Group, "Scholarly Communication at Risk" [Web site], 2003, <http://openaccess.jhu.edu/threats.cfm> (viewed July 19, 2004).
45. ECAR [EDUCAUSE Center for Applied Research], "Evolving Campus Support Models for E-learning Courses," ECAR Respondent Summary, *EDUCAUSE*, March 2003, http://www.educause.edu/ir/library/pdf/ecar_so/ers/ERS0303/ekfo303.pdf (viewed July 19, 2004).
46. "Corinthian Colleges Sets Third Quarter Fiscal Quarter Records for Revenues, Operating and Net Income," April 28, 2004, Press release, Corinthian Colleges, http://www.irconnect.com/coco/pages/news_releases.html?d=56612 (viewed July 19, 2004).
47. Information provided by OCLC Office of Research, April 2004 (e-mail from Thom Hickey, April 1, 2004).
48. Scott Carlson, "Cornell Tries New Publishing Model: Scholarship on Demand," *Chronicle of Higher Education*, March 5, 2004.
49. Brock Read, "Video Service Gives It that New College Try," Press release, Cflix, January 27, 2004, <http://www.cdigix.com/website/news/news0127004.asp> (viewed August 18, 2004).
50. "Trend Alert: The Future of Libraries," *InfoAboutInfo Briefing*, Vol. 7, January 9, 2004, p. 11.
51. "Content Software Technology: Microsoft Office System 2003," *InfoAboutInfo Briefing*, Vol. 7, January 23, 2004.
52. Wendy Pradt Lougee, "Scholarly Communication & Libraries Unbound: The Opportunity of the Commons," Presented at the Workshop on Scholarly Communication as a Commons, Workshop in Political Theory and Policy Analysis, Indiana University, Bloomington, IN, March 31–April 2, 2004, p. 1, <http://dlc.dlib.indiana.edu/archive/00001250/> (viewed August 18, 2004).
53. Mani Shabrang, Dow Chemical Business Intelligence Center as quoted in Drew Robb, "Text Mining Tools Take On Unstructured Data," *Computerworld*, June 21, 2004, n.p., <http://www.computerworld.com/databasetopics/businessintelligence/story/0,10801,93968,00.html> (viewed July 18, 2004).

All OCLC product and service names are trademarks or service marks of OCLC Online Computer Library Center, Inc. In addition, the WorldCat symbol and OCLC symbol are service marks of OCLC. Third-party product and service names are trademarks or service marks of their respective companies. OCLC grants permission to photocopy this publication as needed.

For updates and more information about
The 2004 Five-Year Information Format Trends Report
please visit:
www.oclc.org/info/2004trends



OCLC Online Computer Library Center, Inc.
6565 Frantz Road
Dublin, Ohio 43017-3395
1-800-848-5878 +1-614-764-6000
Fax: +1-614-764-6096
www.oclc.org

ISO 9001 Certified