



Robert Carrigan,
Chairman of the Chairmen's Committee

Ron Milton,
Vice-Chairman of the Chairmen's
Committee

Dan Morrow,
Chief Historian

IBIBLIO

2005 COMPUTERWORLD HONORS CASE STUDY

EDUCATION & ACADEMIA

ORIGINALLY A PLACE FOR THE SHARING AND SUPPORT ALL KINDS OF FREE SOFTWARE, AN INFORMATION COMMONS AND CONTRIBUTOR-RUN DIGITAL LIBRARY ALSO HOSTS OVER 1500 NON-SOFTWARE RELATED SHARING PROJECTS ON ALMOST EVERY CONCEIVABLE SUBJECT, HANDLING OVER 12 MILLION REQUESTS/DAY. [20055412]

SUMMARY

ibiblio, which began as sunsite.unc.edu in 1992, is an information commons and contributor-run digital library. Originally a place for the sharing and support all kinds of free software, ibiblio also hosts over 1500 non-software related sharing projects on almost every conceivable subject, handling over 12 million requests/day.

APPLICATION

Since 1992, ibiblio.org has assumed a leading role in the information commons, information access, digital library, and digital archives fields. As the Internet's contributor-run digital library, ibiblio.org helps organizations, groups, and individuals reach others with common interests anywhere in the world in new and creative manners. Located at the University of UNC Chapel Hill, ibiblio maintains its original philosophy of supporting and expanding information sharing communities and techniques while maintaining one of the nation's oldest and largest digital libraries on the Internet today.

One of the first World Wide Web sites in the world, as sunsite.unc.edu, ibiblio built on UNC's experience in research in the area of searching and retrieval, especially with the Wide Area Information Servers (WAIS). SunSITE began with publicly accessible Internet bulletin boards and established a model for the international SunSITE program which grew to over 40 sites world-wide by 1998.

Paul Jones, ibiblio's director, wrote the following goals for the SunSITE program in the early 90's:

- ? Provide easy, global access to free software and tools
- ? Act as a repository for freely shared information
- ? Promote development and research of new Internet tools
- ? Launch hot new Internet-based applications

More recently, Jones described his vision in an article in The Communications of the ACM entitled "Open(source)ing the Doors for Contributor-run Digital Libraries" (May 2001). He called for applying open source software development principles, including strong contributor involvement, to be brought to use for digital libraries and archives.

Since that beginning, ibiblio has focused on support and learning from our contributors — in short, letting the inmates run the asylum. Keeping an open conversation among the contributors and the ibiblio team leads to discovery of opportunities and of new approaches. As a result, Buddhist monks learn from Brazilian software developers, music tapers learn from GPS specialists, filmmakers learn from ibiblio team, and everyone learns from each other while providing free information to the world.

Starting with only 1.3 gigabytes of information in 1992 and growing to over 12 terabytes in January of 2005, ibiblio.org houses a diverse and creative collection of material. Groups that call ibiblio home include Project Gutenberg (public domain texts), the Linux Documentation Project (Multi-lingual resources for software support and implementation), etree.org (legally sharable music), the Online Burma Library (free information about life in Myanmar), Science for Monks (Tibetan Monks learn science), Documenting the American South (first person narratives of the American South), NASA Educational Video, Folkstreams (video documentaries of American folks practices), as well as Bible translation projects, citizens' groups local to UNC-Chapel Hill and individual folk musicians like Roger McGuinn. In addition to over 1500 collections of sharable information, ibiblio distributes over a terabyte of open-source and free software.

Ibiblio is not limited to WWW or FTP. It pioneered Internet radio (since 1994) and now streams 6 not-for-profit radio stations in various formats including experimental streams on Internet 2. For example, ibiblio helps WCPE, a North Carolina 24x7 classical music radio station, broadcast its content to Internet listeners in Ogg, MP3, and Quicktime formats.

The diverse set of contributors is not the only ones to benefit from ibiblio's information commons. As of January 2005, over 12.5 million information requests are handled by the ibiblio servers every day. Teachers and students use ibiblio collections for historical and geographic research; music teachers use the collections for their students; genealogists use databases to find their relatives; translators meet and work together; adventurers record their exploits; poets have access to Nobel Prize winning poets reading their work; software developers and administrators have access to the world's best software; even regional dart leagues organize their efforts on ibiblio.

During the course of ibiblio's history, ibiblio has received support from or worked closely with Sun Microsystems, Cisco, IBM, Red Hat Software, Mandrakesoft, Sourceforge, NASA, the Institute for Museum and Library Services, Creative Commons, the International Union for Pure and Applied Chemistry, and Network Appliance. ibiblio also collaborates strongly with the Internet Archive, archive.org. The most significant support to date came in the form of a gift from the Center for the Public Domain. Of course, ibiblio's strongest partners are its contributors, who vary from highly organized groups of scholars to highly-dedicated individuals. Throughout its existence, ibiblio has received consistent and strong support from three units of the University of North Carolina: the School of Journalism and Mass Communication, the School of Information and Library Science, and the Office of the Vice Chancellor for Information Technology.

BENEFITS

When ibiblio began as SunSITE in 1992, it focused on support and growth for the open source software community. It employed emerging technologies such as gopher, WAIS, and the newly created World Wide Web. ibiblio's vision, however, was not limited to free software or to the software community. It always included a variety of information resources being made available in new and creative ways. A primary goal was to improve and grow the nascent information commons that had already been cultivated on the pre-Web Internet. The scope of this goal was geographically wide, serving locally the University of North Carolina, regionally the American South, and globally Europe and Asia. Within a few months of launching, ibiblio (as SunSITE) had contributors from and delivered information to every continent.

An early contributor, Henriette Kress, AHG of Helsinki, Finland writes:

In the early 1990s, the Internet was for academics, and private persons simply could not have a website — unless they were affiliated with an academic institution. The first such institution to open up to the great unwashed masses was ibiblio (then sunsite.unc.edu).

What would have been impossible without unc.edu was suddenly easily done. One of the largest herbal information sites on the web, Henriette's herbal homepage, would not exist without ibiblio. It started small, like ibiblio. It's grown enormously, like ibiblio. And it's still free, still available at the original address — like ibiblio."

Ibiblio makes resources available to individuals, groups and entrepreneurs.

Another early contributor, Nicholas Pioch of France writes, "I've just found out that Microsoft Encarta Deluxe 2001 (the copy I just happened to find out and install) has direct links ('Web Links') from each artist's article to the WebMuseum (on metalab.unc.edu at the time) and that's actually the only weblink provided in that 2001 edition."

Ibiblio acts as a readily accessible archive of Internet information sharing resources.

Joe McMahon writes: "ibiblio stepped in when the archives of EMUSIC-L were about to be flushed, early in its lifetime. It still hosts them, even though the list is now long gone, providing a valuable historical reference and a repository of information that is still valuable to many."

Ibiblio connects collaborators and communities world-wide, helping people at many levels of skills and interests.

Jeffrey Elkner, a Computer Science Instructor at Yorktown High School and creator of the Open Book Project writes, "ibiblio gives me the opportunity to be part of an on-line community of learners which spans the world, and to share efforts to develop free, high quality learning materials with this community."

Ibiblio leads in introducing new technologies and new uses of technologies.

In 1994, ibiblio pioneered Internet radio by putting online a digitized simulcast of WXYC, the University of North Carolina's student radio station. This constituted the first 24x7 Internet radio broadcast in the world. WCPE, a classical music radio station was once limited to a small regional listenership. Ibiblio began offering Internet streams for WCPE in 2001 and now rebroadcasts its content in three formats. Today, WCPE receives most of its support from Internet listeners who live all over the world. In November 2004, WXYC and WCPE became the first radio stations to be simulcast on Internet2. ibiblio currently supports four other not-for-profit stations and has plans to add two more during 2005.

The model for information sharing and for an information commons was replicated worldwide under the SunSITE program in more than 40 sites. ibiblio exists as an information commons for all without national boundaries. Nancy Kranich wrote in "The Information Commons: a Public Policy Report":

ibiblio is a heavily used conservatory of freely available information in the fields of music, literature, art, history, science, politics, and cultural studies. A collaboration between the Center for the Public Domain and the University of North Carolina-Chapel Hill, ibiblio encourages users not only to view and utilize the collection and free software, but to critique and expand it, and to create and manage a new collection in their own area of interest.

The information commons concept may be not be obvious to grasp. In fact, many argue that it cannot be done or will not work and if it were to work at all it would be no more than an oddity. Ibiblio has shown over a dozen years of sustainability and innovation that the information commons has value, is workable, is extensive, and can be very inclusive.

There is still important work to be done in understanding cooperative work, using emerging technologies, maintaining interoperability of collections and of software. The areas of digital preservation and archiving are just beginning to receive academic research attention. Ibiblio's extensive experience and large collections will be essential to those new understandings.

ibiblio is not merely a research project, however. It is a working and heavily-used public information resource with great benefit to all.

IMPORTANCE

Ibiblio has been about implementing and extending information technology breakthroughs and best practices from its conception. It began using open source technologies wherever possible. The ibiblio servers are entirely Linux (or GNU/Linux), mainly RedHat's distribution. Application software includes

- ? Apache
- ? Squid
- ? ProFTPD
- ? MySQL
- ? PostgreSQL
- ? Mailman
- ? Icecast
- ? Tomcat
- ? PHP
- ? Perl
- ? OpenSSL
- ? mod_SSL
- ? LVS (Linux Virtual Server)
- ? keepalived
- ? Nagios
- ? Cricket

ibiblio supports many open source development projects, as well as the Linux Documentation Project. Recently, it has launched two projects of its own, which contribute software to the open source community: Osprey and Lyceum.

Osprey is a peer-to-peer enabled content distribution system. It combines a flexible metadata management system with the BitTorrent peer-to-peer protocol. A Web database application provides searching and browsing of collection objects, and the peer-to-peer component lowers bandwidth costs by employing

distributed downloading. The Permaseed application supplies reliable, persistent peer-to-peer access to files in a digital repository. Osprey is designed for digital repositories that house large volume content, such as multimedia files or research data sets. <http://osprey.ibiblio.org>

Lyceum is a multi-function application for weblog research and deployment. Developed using the extensible tools, PHP and MySQL, Lyceum is a free, open-source software package that aims to provide simple and automatic blogosphere generation with the ability to observe the activities and links of interconnected nodes (or blogs). <http://lyceum.ibiblio.org>

New technologies factor heavily into the services ibiblio provides to the public. Likewise, ibiblio contributors are also enabled to create or recombine new technologies in their projects. For example, Project Gutenberg built a distributed proofreading system to facilitate its digital text generation. The online music community, etree.org, uses new compression formats such as Shorten and FLAC for near-lossless digital audio.

To aid in interoperability, ibiblio favors the use of open standards. Its projects and services conform to metadata standards for information exchange including Open Archives Initiative, Dublin Core, and various appropriate XML standards. A recent addition to ibiblio.org is a group of XML RSS feeds collected from several ibiblio contributor sites.

Less technically complex applications of information technologies have a broad impact as well. Many of ibiblio's contributors simply provide material that cannot be obtained from another source. With hosting and assistance from ibiblio, groups of Tibetans in exile in Dharmasala have created over a dozen Web sites. Topics of these sites include science, politics, religion, philosophy, and education. One site, kunphen.org, focuses on helping Tibetans with substance dependence and HIV/AIDS.

ibiblio's focus has been not just on information technologies, but also on allowing its contributors to explore. Groups are trusted and encouraged to experiment. As a consequence, the use of information technology has arisen in unusual places, sometimes leading to broad social change. A box of tools and a few spirited examples placed in front of creative people yield the best unanticipated results. Such is the history of ibiblio.

ORIGINALITY

Ibiblio began uniquely as one of the first websites in the world and arguably the first in North America outside the high energy physics community. Ibiblio has long been unique in its longevity, size, growth, and even in its eclecticism. As ibiblio has evolved over time, its attention to emerging technologies and its relationships with developers and with creative contributors has led it to be among the first, if not the first, in many areas. As previously mentioned, it pioneered simulcast radio on the Internet. In celebration of the 10th year of Internet radio, it was the first to produce a live radio broadcast stream for IPv6 on Internet 2.

One distinguishing feature of ibiblio is its continued focus, obsession really, with freely available information. Since 1992, ibiblio has remained consistent across the ebbs and flows of fashion. Over time, other groups, especially the SunSITE groups, have adopted many of the ibiblio practices.

Ibiblio has constantly reached out to innovators. It has collaborated not only with technology innovators such as Linux developers but also with policy innovators such as the Creative Commons and the Center for the Public Domain, with content process innovators such as Project Gutenberg and Documenting the American South, with digital format innovators such as etree.org and folkstreams.net, and with content innovators such as the Afro-Louisiana Slave Database.

ibiblio continues to be unique in its collections and in its processes. It attempts to integrate best practices and new research findings for digital archives into the management of a very large collection of collections. Ibiblio will likely continue to be a trendsetter for digital collections and efforts in the preservation of our international information commons.

SUCCESS

Ibiblio's successes began early on by exceeding the expectations of everyone involved. Fully functional and accessible world-wide since 1992, ibiblio serves over 13 million information requests per day (as of January 2005).

The timeline below shows how ibiblio has evolved over the past dozen years. Selected collections are noted each year to give a sense of the variety among the 1500 collections (other than software) and of the variety of publics served.

1992 - sunsite begun with a goal 10,000 downloads a month; we exceeded that in 2 days before we were officially announced.
Linux Software Archive begun (still ongoing). WebMuseum joined sunsite (still ongoing).
Tim Berners-Lee drops by and we set up a WWW server, one of the first in the US and likely the first outside high energy physics.

1993 - Presidential Archives work with Clinton White House using WAIS.
Linux Documentation Project begun (still on going).
Dr. Fun, first daily cartoon on the Web, begun (still ongoing). CEARC (Cisco Educational Archive) started to provide information about networking in education.

1994 - WXYC first live 7/24 radio simulcast on the Net (still ongoing).
Virtual Shtetl, the first known Yiddish site on the Web (still ongoing).
Internet Poetry Archive begun with Nobel Prize winner Czeslaw Milosz's contribution in English and in Polish, in audio and in text (still ongoing).

1995 - Henriette Kress's herbal page is one of the oldest herb sites on the web (still ongoing).
Invited presentation to President and Mrs. Jimmy Carter at The Carter Presidential Library on digital Presidential Library futures.

1996 - "Documenting the American South" begun with first person narratives by enslaved people in the American South (still ongoing).

1997 - Roger McGuinn's FolkDen joins sunsite and changes to streaming formats for his audio collection (begun in 1995 and still ongoing).
Tibetan Centre for Human Rights and Democracy comes to ibiblio. First of dozens of Tibetan-owned sites on ibiblio (still ongoing).

1998 - Nobel prizewinner Aung San Suu Kyi's videotaped address on the occasion of the 50th Anniversary of the UN Declaration of Human Rights (smuggled out from her house in Rangoon where she'd been detained for a decade) made available on sunsite.
Sunsite becomes MetaLab. 50 million files served in the month of October 1998.

1999 - Open Book Project begun (still ongoing).
Presidency Research Group and Pew Charitable Trusts establish web resources to assist the 2001 Presidential transition. (still active whitehousetransition.org) Served two presidential transitions and nine emerging democracies.

2000 - MetaLab becomes ibiblio with large gift from Center for the Public Domain.
Afro-Louisiana Slave Database by Gwendolyn Menlo Hall made available (still ongoing).

2001 - Interactive Native American Archaeology Project, Excavating Occaneechi Town, made available online for the first time (still ongoing).
The "Internet Studio" Academic Program initiative is directed to create strategic relationships with Latin American Schools of Architecture, institutions and corporations involved in the urban development of the Latin American and Caribbean region by working in collaborative semester-studio projects via the Internet. More than 450 architectural students from Miami, Argentina, Chile, Ecuador and Venezuela collaborated in semester-long design studios via the Internet and videoconference technology during Fall, 2001. (still ongoing).

2002 - VolunteerTibet officially launched its service on December 10th, 2002, commemorating the date His Holiness the Dalai Lama received the Nobel Peace Prize (still ongoing).
Mental Health Association in Durham County (NC) founded and joins ibiblio (still ongoing).
A Living Stage: Reflections from South Africa Music and storytellers from the 2002 National Arts Festival in Grahamstown, South Africa.
www+ftp traffic now averages 3 million files/day.
Data flow averages 150 Mbs.

2003 - Khawa Karpo-Tibet Culture Centre (KKTCC), Dharamsala joins ibiblio (still ongoing).
GrokLaw: Digging for truth: the law and the facts with regard to Free and Open Source Software joins ibiblio (still ongoing).
librarian.net: Jessamyn West's library themed weblog of interest to rarified librarians and other information junkies joins ibiblio (still ongoing).

2004 - www+ftp traffic now averages over 10 million files/day. Data flow averages 250 Mbps.
NASA Educational Video added (still ongoing).
WCPE and WXYC become the first simulcast radio stations on Internet2 usingIPv6.
Institute for Museum and Library Services awards ibiblio and Folkstreams.net grant for digitization of documentaries on American folk cultures.

2005 - www+ftp traffic now averages over 13 million files/day.
Data flow averages 350 - 400 Mbps.

Ibiblio has been successful in building a brand and in building relationships that help collections and contributors find new publics.

Contributor Tim Poe describes his experience:

“Open Studios is a small 501(c)(3) organization located in Reno, Nevada, USA. Our interest in creating a web site stemmed from the idea that there was a worldwide audience out there, and to reach them, to communicate with them, to share information about our work with them, we needed a web presence. Without a budget to work with, hosting for our web site was a considerable cost consideration. Ibiblio.org, not only represents a huge contribution to our organization by providing free hosting, but also represents an enormous jump in a marketing sense. On our web site, hosted by Ibiblio.org, we have experienced a rapid entry (and to the top of) on search engines for the information on our site. That would not happen, if we did not have the association with Ibiblio.org. As a result of the work Ibiblio.org does on behalf of the sites they host, Open Studios has achieved far greater recognition around the world, than we ever expected before we started with them.”

A note from Pamela Jones of the legal discussion site, GrokLaw, tells a more specific story:

"I want to tell you how much you have meant to Groklaw. Without ibiblio, Groklaw could never have become what it has. We are all volunteers, with no budget at all, except our own lean pockets. ibiblio made it possible for Groklaw to reach a wider and much larger audience. We have won a number of journalism awards, and I absolutely know that I owe it in part to ibiblio's hosting us. Thank you, ibiblio. If you only could read my email, with daily thank you notes from readers all over the world, you'd realize how many people enjoy Groklaw and appreciate that it exists."

And a note from an individual user of Roger McGuinn's Folk Den tells the story of how ibiblio collections reach beyond computer networks to a more intimate level:

"Thanks for setting up the Folkden site. I've been downloading the songs a few at a time (modem connection) for the past week. I plan to use them with some of my private guitar students. These are the tunes and strums I grew up with. What a perfect way to pass the tradition along to the next generation."

Ibiblio is not only a collection of collections; it is also a site for research and for software development and design. Papers by ibiblio staff have been presented at academic conferences including the Joint Conference on Digital Libraries and have been published in The Communications of the ACM (one of which was translated into Japanese), Journal of Information Science and Technology, Library Trends and others.

Two projects, Osprey and Lyceum, are described in the Importance of Information Technology section. Both are due to be in beta release during 2005. Both represent a future direction for ibiblio to expand the research and development mission of the project.

Ibiblio will continue to lead in collections and will seek to lead in digital archives development in many media especially in digital video and audio. All along the way, ibiblio will continue to promote and expand our Information Commons and sharing concepts and practices.

DIFFICULTY

Personnel resources

ibiblio has only two full-time employees. All other employees are part-time student workers, whose number has never exceeded 8 at any one time. Project management, including that of research projects and of development, is done by one professional faculty member who also carries a half teaching.

A major challenge of a large student staff is a constantly changing workforce. Ibiblio turns over between one third and one half of its student staff every year. After a dozen years of such turn over, this turnover is not

only expected, but prepared for — as best as such change can be managed.

Infrastructure resources

ibiblio is not entirely dependent on corporate donations for hardware and bandwidth resources. However, donations and partnerships have made it possible for ibiblio to meet the growing needs of its contributors and to advance into new technology areas. ibiblio's bandwidth usage matches or exceeds that of the rest of the UNC-CH campus. While it can be provisioned with the necessary bandwidth, some of that usage has had to be funded by contributions.

Policy and copyright

ibiblio has encountered very few true copyright violations from our contributors. These have been easy to deal with and in the three specific cases were solved amicably and quickly. However, dealing with false claims of copyright has proved quite time consuming. The most notable false claim centered on the use of the Dewey Decimal System.

Ibiblio has encountered several claims of inappropriate materials. Few have been what one might expect. Chinese objected to Tibetan sites. Socialists objected to each others' sites. Scientists objected to other scientists' work. In one case, a man in Arizona objected to the scholarly collection of slave narratives, Documenting the American South. His objection was that he could find no former slaves that celebrated their time in slavery!

There is some evidence that ibiblio.org is blocked by some sites in China. But there is also evidence that ibiblio is not blocked at other sites there.

Growth

Since sunsite.unc.edu was launched, before it was even announced in 1992, use of ibiblio has been greater than expected. Bandwidth, processing cycles, storage, and even physical space in a data center have always been under pressure. While one loves the success and popularity, managing ibiblio is always a challenge.

Sustainability

Putting digital materials online is a small problem compared to preserving, cataloging, creating basic metadata, and migrating materials as necessary. Ibiblio has one of the longest records of successful digital access on the Internet today. However, only a basic understanding of how to sustain digital materials exists. Much research is still needed in this area, which guides many of ibiblio's projects.