

OSDN : [SourceForge](#) - [Shop ThinkGeek](#) - [freshmeat](#) - [Newsletters](#) - [TechJobs](#) - [Personals](#) - [Slashdot Broadband](#) **Search**



[pjones](#)

[Preferences](#)

[Subscribe](#)

[Journal](#)

[Logout](#)

Sections

[Main](#)

[Apache](#)

[Apple](#)

[2 more](#)

[Askslashdot](#)

[14 more](#)

[Books](#)

[BSD](#)

[1 more](#)

[Developers](#)

[2 more](#)

[Games](#)

[14 more](#)

[Interviews](#)

[Science](#)

[7 more](#)

[YRO](#)

[3 more](#)

Help

[FAQ](#)

[Bugs](#)

Stories

[Old Stories](#)

[Old Polls](#)

[Topics](#)

[Hall of Fame](#)

[Submit Story](#)

About

[Supporters](#)

[Code](#)

[Awards](#)

Services

[Broadband](#)

[Online Books](#)

[Personals](#)

[PriceGrabber](#)

[Product News](#)

[Tech Jobs](#)

Ibiblio Director Paul Jones Answers

Posted by [Roblimo](#) on Wednesday
August 07, @11:00AM

from the
using-more-bandwidth-every-year
dept.

Okay, here are answers from [Paul Jones](#), director of [ibiblio.org](#). [You asked](#), and he responded -- and not always as seriously as you'd expect from someone who can ask us to call him "Professor Jones" or "Doctor Jones." But he's really "Just Paul," he says, "even in class." We hope a whole lot of you have a chance to meet Paul in person one day, because he's not only a warm and friendly guy, but one who has done a whole lot of good for Linux -- and for the Internet in general.

Paul:

Let me start out with a little overview of [sunsite.unc.edu/metalab.unc.edu](#). Or better yet to point you to [our annotated timeline](#). Then say that [ibiblio.org](#) began and has continued to be a way for the [University of North Carolina \(the original and still the best\)](#) to explore information sharing in the context of our missions of education, research and outreach. You folks using and contributing are the outreach part. In particular, we ["acquire, discover, preserve, synthesize, and transmit knowledge"](#) with all of your help.

We are a joint project of the [School of Information and Library Science](#) (there we are involved in digital archives and digital libraries), [The School of Journalism and Mass Communication](#) (there we are involved in electronic publishing and multimedia sharing), and the [Vice Chancellor for Information Technology](#).

Except for one and occasionally two full time employees, [our entire staff](#) consists of students or in my case part time (as I have faculty responsibilities). So be nice to all of us, we're always learning. No matter what Robin said in the article introducing me, none of this would have happened without some very good people on staff and contributing content.

But that brings us to:

Question of Money by too bad

One of the things that people frequently ask about sites like [ibiblio.org](#) is "They are great. But how long will they be around?" Do you see this as a concern (esp. after the

Advertisement

pjones (10800)

[pjones](#)

[pjones@ibiblio.org](#)

(email not shown publicly)

<http://www.ibiblio.org/pjones>

Karma: Excellent

AOL IM: [smalljones](#) ([Add Send Message](#))

Related Links

- [our annotated timeline](#)
- [University of North Carolina: original and still the best](#)
- ["acquire, discover, preserve, synthesize, and transmit knowledge"](#)
- [School of Information and Library Science](#)
- [The School of Journalism and Mass Communication](#)
- [Vice Chancellor for Information Technology](#)
- [our entire staff](#)
- [sponsors large and small](#)
- [top secret international fund as well](#)
- [Internet2/Distributed Storage node](#)
- [Archive.org](#)
- [Gentoo](#)
- [WebMuseum \(originally We](#)
- [Microsoft Encarta Deluxe 20](#)
- [The Linux Documentation Project](#)
- [Documenting the American](#)
- [Hong Kong Picture Archive](#)
- [Henriette's Herbal Homepag](#)
- [Hyperwar](#)
- [Creative Commons](#)
- [Roger McGuinn](#)
- [Folk Den](#)
- [Dr Fun](#)
- [Plan 9](#)
- [Collection Policy](#)
- [Internet Anonymous FTP Archive \(IAFA\)](#)
- [Linux Software Map \(LSM\)](#)
- [Jonathan Magid](#)
- [contributor-maintained Col Index](#)

LWN announcement) and do you have any comments regarding this. Are there any good approaches you suggest (like augmenting free usership with voluntary subscriptions, etc) for such free sites in general?

Paul:

We have been very lucky, since our beginning, to have generous and understanding support from The University of North Carolina and from [sponsors large and small](#) including Sun, IBM, Red Hat, VA Linux^{h^h^h^h^h}Software, Mandrake, Cisco and others.

We also do get some research contracts and grants, but most importantly for us in the past two years has been a large gift from the founders of Red Hat and the Center for the Public Domain.

We have some [top secret international funding sources as well](#). At the moment, we actually have a small endowment that if spent wisely should last several years. It is my hope that we will never have to charge the patrons of our digital archives.

BUT this brings me to my favorite question, which only got a rating of 4:

Donations?

by Anonymous Coward

Where do I send the cheque?

Paul:

Send your or your organization's tax-deductible contributions to:

Ibiblio.org
Campus Box 3456
University of North Carolina
Chapel Hill, NC 27599-3456

Moving on to:

Typical Questions

by suwain_2

I've downloaded my share of things, and find that the 3 Mbps cap on my cable modem is almost always my bottleneck. So my question is fairly simple (albeit broad) -- can you describe your setup a bit, in terms of bandwidth (both what you have for an Internet connection, and how much traffic you actually use), servers, storage (I'd venture to guess it's to the tune of several terabytes?), etc.

Paul:

We're on UNC's network. Our connections to the commodity and Internet2 networks are served by UNC's OC-48 network connection. We maintain a constant throughput of network traffic outbound in the 160-180Mbps/sec range.

- [Open Archives Initiative \(OAI\)](#)
- [Digital Resource in the Humanities conference](#)
- [Problem of Access in Contributor-Run Digital Libraries](#)
- [we worked with several others](#)
- [Open Source Metadata Framework \(OMF\)](#)
- [Scrollkeeper project](#)
- [Howard Rheingold](#)
- [Mickey Mouse](#)
- [American Open Technology Consortium \(AOTC\)](#)
- [others on the boards](#)
- [Stationers Guild](#)
- [WXYC](#)
- [Ms. Frizzle says, "Take chances, get messy and EXPLORE! Wahoo!"](#)
- [WAIS](#)
- [That was more fun than Wack-a-mole](#)
- [slack](#)
- [Paul Jones](#)
- [ibiblio.org](#)
- [You asked](#)
- [More on The Internet](#)
- [Also by Roblino](#)

Interviews

Welcome to the interviews section - this is place to come to read the assorted conversations that Slashdot and the readers have had with various people involved in the Internet, computers, or anything of interest.

Our current main servers were donated by IBM and serve content from a central fileserver with 2TB of disk attached. In our racks, we have approximately 5TB of space (with system disks, Sourceforge and an [Internet2/Distributed Storage Initiative node](#)). We do some load balancing between streaming services, web services, and large downloads like distros. On a typical day, we move over 1.5 terabytes of data off our servers. (Thanks to Fred Stutzman for much of this info.)

Backups

by Chris Pimlott

What's your backup strategy? I imagine it's hard to deal with both so much data as well as being under constant bombardment from clients around the world. How often is data archived? Have you had any major data loss incidents and, if so, how well were you able to deal with them?

Paul:

Like everyone else we rely on [Archive.org](#), but seriously... (Fred answers this since he did the restore).

We run managed backups on UNC's enterprise storage facilities. We run them every night and have incremental backups for three months. UNC uses StorageTek machines and Tivoli Distributed Storage Manager for enterprise backups. We have had major data loss incidents, in which a raid card failed and lost the array's configuration. One of the disks in the array died simultaneously, we were unable to re-import the configuration to the new card, so we had to restore from backup, which took a number of days.

I, Paul, can only say that in the past things were much worse and we did have one famous meltdown in 1995 that was not pretty. Since then the UNC enterprise backup has been our friend - and for the most part disks and RAID arrays have been increasingly more reliable.

What's your biggest area?

by Otter

I know ibiblio (I still think of it as SunSite) as a) a repository of Unix software, especially useful for pre-Freshmeat apps and b) a mirror provider. "Free online publisher" wouldn't have made the list, but looking at your main page I see all sorts of things I didn't realize you hosted. Which ones get the most traffic?

Paul:

For sheer bytes, ISOs rule. But then it doesn't take too many downloads to get a lot of bytes for an ISO. Source-based distros like [Gentoo](#) have seen a lot of activity lately.

One of our most visited sites is also one of our oldest, Nicholas Pioch's [WebMuseum \(originally WebLouvre\)](#). An amusing reason may be that, as Nicolas 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."

Among other favorites are:

- [The Linux Documentation Project](#), which began on sunsite
- [Documenting the American South](#)
- [Hong Kong Picture Archive](#)
- [Henriette's Herbal Homepage](#)
- [Hyperwar](#) A hypertext history of the Second World War

What about content producers?

by Fluid Donkey

In general how supportive have you found the producers of such content to be of your services? Do many if any really believe that something like this will cause them to starve to death?

Paul:

First, they are all with us voluntarily and can leave any time, taking their stuff with them. That alone pretty much says that they believe in what we are helping them do.

I should say also that not all material is copyleft. But all of it is free to view, listen to and to reference. We are working with [Creative Commons](#), which we also host, to develop a small but viable set of licenses for folks including our contributors who want to share their work on various terms (attribution, home or personal use, educational use, etc).

One important contributor, [Roger McGuinn](#), has been making one folk song a month available for download since November 1995 on his [Folk Den](#). He also sells CDs and performs concerts. He seems to be doing pretty well. Many contributors are scholars or students who understand the importance of sharing information.

Dave Farley, who does the wonderful [Dr Fun](#), has a book contract with [Plan 9](#), and we're looking forward to seeing what we've seen in electrons in print.

Relative importance of different material?

by kafka93

What is the center's view on the publishing of material that might be considered "offensive" or "dangerous",

and does the center make subjective judgements upon the importance of one piece of intellectual property over another on the basis of 'artistic worth', 'decency', etc.? With only limited resources available to promote the archiving of data, is there the risk that important fringe documents may be left by the wayside, or ignored due to political/social concerns?

Paul:

Like non-digital archives and libraries, we have a [Collection Policy](#). You'll note that we do not explicitly ban materials for content nor do we plan to. We do not maintain materials that are illegal, slanderous, libelous, or otherwise prohibited by law. Ultimately the contributors are responsible for their content and we do not review the content once a project is taken on.

Most rejections of content come about because the content is too commercial, just personal, or relies on advertising.

***Metadata and easy searching
by Ryan Muldoon***

iBiblio stands out as an excellent repository for a wide range of culturally valuable resources. As it and other sites grow in size, the importance of good searching and indexing becomes extremely relevant. Have you given any thought to how you might want to cope with this? Specifically, are there any metadata schemata that you are considering using? I would love to see iBiblio be used more like a content feed to research/cross-referencing applications.

Paul:

Interesting that you asked about this as this is an area that we've been working in for the past couple of years. Actually we go way back to pre-Web metadata to the [Internet Anonymous FTP Archive \(IAFA\)](#) files which were the model for the [Linux Software Map \(LSM\)](#). Thanks to [Jonathan Magid](#) for this innovation and for suggesting that we host Linux in the very beginning.

When we designed our [contributor-maintained Collection Index](#), we designed it to create and display metadata that could be shared via the [Open Archives Initiative \(OAI\)](#). Please note that this metadata is at the collection level - not at the item level. Item level metadata is for future work. Also since you asked: Miles Efron and I will be presenting a paper at the [Digital Resource in the Humanities conference](#) in September on the [Problem of Access in Contributor-Run Digital Libraries](#). Serena Fenton is co-author to this paper.

On the Linux Documentation Project front, [we worked with several others](#) to create the [Open Source Metadata Framework \(OMF\)](#).

The OMF aims to collect data about Open Source documentation, or metadata, that will be used to describe

the documentation. The idea is that the OMF will act as a sophisticated card catalog type of system for the numerous Open Source documentation projects that exist. The OMF offers a number of advantages over standard card catalog type systems, however. Chief among these is the fact that the OMF has been designed from the ground up to be completely open, standards based, and sharable. We will accomplish this by using pre-defined standards (XML and the Dublin Core description for metadata) and allowing all metadata generated to be accessed by anyone that wants it. Because the metadata itself is to be stored in XML files, anyone should be able to use it.

OMF support is included in the [Scrollkeeper project](#). Note that none of these metadata designs are overly complex. That is by design. The idea is to keep the metadata simple enough to be understood by the creator of the digital item or collection that it describes. If I could make one strong point about metadata design it is that simplicity is the key - and the hardest thing to pull off.

***Trust metric and online publishing
by Creosote***

I heard you talk at the Southern Presses conference last year about the use of trust metrics (like Slashdot's karma and Advogato's peer certification) as a possible alternative to the "top-down" means of filtering that scholarly and commercial publishers use, namely formal peer review and mass marketing, respectively. Are you more or less optimistic about the long-term viability of this model then you were then? (Especially in light of the powerful efforts to keep control of the gates we're seeing these days from Hollywood, the recording industry, and their political allies...)

Paul:

Beginning here I am speaking personally and not on behalf of ibiblio.org or any of its sponsors or supporters including but not limited to the University of North Carolina.

The Blog is one example of creator-empowerment that has gotten more attention since that talk and I think there will be plenty more examples to come. I still believe that people in constant communications will result in "Smart Mobs" (thank you, [Howard Rheingold](#), for naming and noticing and writing on this). This is not just about music or movies or about one country or even one age group. While I don't think that we will completely replace our reliance, however reluctant, on [Mickey Mouse](#), I do think that we are entering a time in which there are new opportunities for us to share information and to work together. The slew of misguided efforts by media and information cartels, especially the RIAA, which demonize their customers and clients, will make things tough but they also are signs that the old solutions are not working well and that newer, and I hope more

inclusive and more open, solutions are on the horizon.

GeekPAC and "When Congress Attacks"

by lunenburg

I noticed that you are one of the founders of the American Open Technology Consortium and/or GeekPAC - the lobbying group that got a bit of fanfare a few months back when it was formed, but has been pretty quiet since then. With Congress launching seemingly daily attacks on our technological freedom in order to support the revenue models of a few huge businesses, the need for a voice in Washington is growing urgent. Is the AOTC/GeekPAC working to get our voices heard? Is there a need for an umbrella group to tie together various groups like GeekPAC, Public Knowledge, Digital Consumer, etc.?

Paul:

Yes, (again speaking only as Paul) I am an officer of the [American Open Technology Consortium \(AOTC\)](#). But for various complex reasons, I am not a member of GeekPAC. As you might have guessed, getting these projects going has been no simple matter. Jeff Gerhard has been doing a wonderful job of making sure the legal and procedural steps are properly taken. So far, what you are seeing is some very motivated but very busy people learning how to work together to get the projects off the ground. The good news is that folks like Jeff, Doc Searles and [others on the boards](#) are smart, dedicated and experienced people who can and will play well with others (including Public Knowledge and Digital Consumer and EFF). We hope to represent slightly different voices than those already represented. If you are reading this, you know who you are and we need your help.

About the umbrella group, I think that a summit conference (or at least a summit listserv) would make more sense. This kind of looser structure, often called an Action Committee or Organizing Committee, has been very successfully used by both ends of the political spectrum in the past half century.

Two words...

by Anonymous Coward

DRM? Palladium?

What's your take on these two technologies?

Are you afraid they'll ultimately destroy what you have been working for, for the past 10 years? If not, why?

Optional question: What about the copyright extension we have seen?

Another optional question: Linux... or BSD? =)

Paul:

Not Linux vs BSD, but Digital Rights Management and

Microsoft's Palladium. DMR is the general term for the groups of solutions to the need for creators to be compensated for their work while allowing their audience to easily access those works. Or at least that would be ideally what DRM should do.

When DRM goes wrong, it tramples on the rights of the citizens to have access to information that they have legally purchased, want to criticize, parody, legally reuse or share.

When DRM goes wrong, it creates barriers to innovation and creativity. It biases access and reproduction of information to only certain technologies.

When DRM goes wrong, it creates and perpetrates closed markets and monopolies.

When DRM goes wrong, everyone suffers. It takes us back to the [Stationers Guild](#), a response to the printing press. "The Stationers Guild obtained monopoly rights in the printing and probably distribution of all books, a monopoly codified by the Tudors in a licensing system aimed at censoring religious dissent" which lasted until the early 1700s.

When DRM goes wrong, it is called Palladium.

The good news is that Palladium is vaporware - so far.

What is your greatest success/failure?

by burghurburg

Simple enough question in two parts:

Looking back on 10 years of doing this, what would classify as your greatest success, and your greatest failure?

Paul:

The simplest question is the hardest, of course. Luckily, you've narrowed the success/failure question to deal only with sunsite/metalab/ibiblio and not the past 10 years of my life.

One mark of great success is that we are still here hosting some of the original collections of information to be shared on the Net including the first 7/24 radio simulcast on the net, [WXYC](#). We've been a part of many innovations and I, personally, have been able to work with some brilliant folks who often surprised themselves with what they had accomplished. We're also funded and we enjoy support from some wonderful and diverse faculties at UNC.

There is no question in my mind that the most significant decision that I made in those ten years was to listen to Jonathan Magid when he suggested that we become the US site for an operating system that didn't even work yet - Linux. If you are reading this far and

are happy, you owe Jonathan. If you are unhappy, blame me.

In research, there is no such thing as failure. As I was explaining to our Interim Vice Chancellor, we are supposed to make mistakes. As [Ms. Frizzle says, "Take chances, get messy and EXPLORE! Wahoo!"](#).

Still, I do wish that we had found a way to use [WAIS](#) or another distributed search engine in a way that is still useful. There still seems to me to be something unfinished in that area. Killing gopher. [That was more fun than Wack-a-mole.](#)

And one final answer:

Slack.

by dsb3

You host a slew of subgenius content, so it must be asked ... do you have slack?

Paul:

While I do not profess to completely comprehend [slack](#), I have been assured by members of the Church that I do have it.

This discussion has been archived. No new comments can be posted.

Ibiblio Director Paul Jones Answers | Preferences | Top | **87** comments | Search Discussion

Threshold: Save: ☐

The Fine Print: The following comments are owned by whoever posted them. We are not responsible for them in any way.

Content producers (Score:4, Interesting)

by [hetta \(414084\)](#) on Wednesday August 07, @02:21PM ([#4027113](#))
(<http://www.ibiblio.org/herbmed>)

In general how supportive have you found the producers of such content to be of your services? Do many if any really believe that something like this will cause them to starve to death?

I'm one of those content providers. *Checks self:* Nope, not starving. In fact, I **love** ibiblio:

They give me unlimited non-commercial space in ftp and html (and that really is unlimited. I have zipfiles of herbal forums online, from 1992 onwards... couldn't do that if I had to pay monthly fees for the space.)

Ibiblio is in all the search engines.

You can still get my main page with the [same URL](#) [unc.edu] as that used back in 1995 - how many sites can you say that about?

There's smaller perks, too, like a shell account, setting up mailing lists (no ads!), and such.

So here's a big **Thank You** to both ibiblio.org and unc!

Cheers
Hetta

Re:My question was modded to a 5, but wasn't asked (Score:4, Informative)
by [pjones \(10800\)](#) on Wednesday August 07, @02:02PM ([#4026923](#))
(<http://www.ibiblio.org/pjones/>)

Just not one that I got from /. but not out of line

10 years ago we set a goal of about 10,000 downloads a month with Sun. We beat that in 2 days.

By two years later we completely saturated a t-1.

Now we average between 150 and 200 Mbs all the time.

(I may have answered this part above).

Setting a price is more difficult. We pay students, but they also get trained etc and several of the students are paid by research grants and gifts. All most all of our hardware is donated -- so setting a cost on that is imprecise. Our space, our machine room (7/24 controlled environment, monitoring, backups, and the like), and much more is not priced but contributed by the Univesity). We do pay for our portion of the network use of the commercial internet but that is bought by a university consortium and not at a regular rate. so costing out the project is not an easy task. we also support many research projects that return moneys indirectly to the school etc.

But let's just say it's not cheap and we greatly appreciate the support we get from UNC and from places like the Center for the Public Domain, and companies like IBM, Sun, Cisco, Red Hat, mandrake etc.

But especially from smaller local companies like webslingerZ, islandsedge and others who sponsor students

Certified Black Helicopter Pilot *** Unwitting Dupe of One World Gov'ment

[[Parent](#)]

- [1 reply](#) beneath your current threshold.
- [31 replies](#) beneath your current threshold.

Computers are useless. They can only give you answers.
-- Pablo Picasso

All trademarks and copyrights on this page are owned by their respective owners. Comments are owned by the Poster.
The Rest © 1997-2004 [OSDN](#).

[[home](#) | [awards](#) | [contribute](#) | [story](#) | [older articles](#) | [OSDN](#) | [advertise](#) | [self serve ad system](#) | [about](#) | [terms of service](#) | [privacy](#) | [faq](#) | [rss](#)]