P R O C E E D I N G S

PRESIDENT POWERS: We had the idea of bringing figures in research libraries and digital scholarship to our campus to talk about these issues, and we succeeded. And we succeeded because all of you are here today. We have a truly, tremendous group.

I know you're very busy in your areas -- for most of you or many of you -- on your campuses. I very much appreciate the time that you've taken to be here.

Jim, I want to give a special welcome to you as the, I think, the only president emeritus here. That has a nice ring to it. Are you enjoying yourself?

Again, thank you all for being here. I hope this conference is a tremendous benefit to you. We're being a little bit selfish about it. We're facing these issues on our campus. And we thought, what better way to get advice than bring leaders from all around the country on these issues.

And I want to say on behalf of myself and our campus, we're very grateful for the help that you're going to give us over this period of time, pointing us in the right direction. And we have a challenge.

The challenge we all have is how we can best leverage our investments, the investments we make in our libraries and digital research, and to leverage those
investments to advance research and teaching and learning.

We simply can't recruit faculty in this day and age or develop academic programs, and we could not have done it on our campus -- and I know that's true on your campuses -- without tremendous resources in our libraries and collections. We're very proud here at UT. I think we've done that.

At is turns out since 1963 -- this according to the Association of Research Libraries -- on our campus we've spent over $650 million since 1963. And that doesn't include the bricks and the mortar -- $650 million. That's a tremendous investment.

And I know on your campuses you've had similar investments. We need to think carefully how to leverage those investments -- and those investments as we go into the future -- to make sure that we're using them in the way that develops our faculty, develops our students, develops our academic programs and our research.

We're very proud on our campus -- as I know you are on your campuses -- of what we've done. Our collections rank in the top dozen in America in terms of volume count. They have tremendous richness and uniqueness and diversity.

And we do find that these collections help us attract the very best faculty and in many cases the very
best students. Let me just give two examples we're proud of on our campus.

David Oshinsky is a faculty member in our history department. He won the 2006 Pulitzer Prize in history. J.M. Coetzee, an alumnus and a graduate of UT, was the winner of the 2003 Nobel Prize in literature. And they are examples of people who have been able to use our collections, as a student in one case, as a faculty member in another case, to advance that work and use these great resources.

Coetzee wrote an essay entitled "How I learned about America and Africa in Texas." And he wrote about the time he had on this campus in the 1980s. And he said this, "I'd had the run of a great library, where I stumbled on books whose existence I might not otherwise never had guessed."

Coetzee credits the UT library as a tremendous resource of insight into history and culture and languages of southwest Africa. He used our libraries. And as I said, we're all proud of those libraries and those collections.

And I hope while you're here on our campus you'll have a chance to visit and see some of the richness of our collections. I do hope you will be able to explore them. We've had very acquisitive directors and great
leadership in our libraries.

Let me just give some examples of things we've been able to acquire in the last few years alone. At the Ransom Center we've acquired the Watergate papers of Bob Woodward and Carl Bernstein, as well as the archives of the film legend Bob De Niro, the renowned portrait photographer Arnold Newman and the American icon, and ironically iconoclast, Norman Mailer.

The Center for American History just received 17 tractor trailer trucks full of Exxon/Mobile historical archives. And that goes back to the Standard Oil Trust in the 1860's.

We've actually got the documents, the letters that form the Standard Oil Trust, what a great bit of reading -- and if you get a chance to see that -- of what they had in mind and the thinking that went into forming the Standard Oil Trust in that very early formative period of our business and corporate life.

The Nettie Lee Benson Latin American Collection is one of the leading collections and a collection of record for an entire continent and indeed beyond an entire continent.

And of course no cold war history or history of the formative years of the civil rights movement would be complete without consulting the richness of the Lyndon
Baines Johnson Presidential Library here on our campus. I hope many of you, who have not had a chance to see that, will have a chance to see that while you're here.

So we know what libraries and collections can do. And there are similar stories on your campuses. For those of you who come from campuses you know what great collections and great libraries can do for an institution of higher learning.

But we have challenges. What does the future hold for great universities and great research universities who face the challenges of advancing these libraries into the 21st century?

Will we be able to respond to the increasing and steadily rising cost of books and journals and databases and the demands that our faculty and our students put upon us?

Our scholarly and commercial publications here at UT alone were increased by a cost of almost $1 million this year as we renew them for the calendar year. Digital scholarship brings additional challenges and opportunities.

How will our universities respond to make sure that we're getting information and not just data? How will we organize this information? How will we preserve it and carry digital data across the millennia and make
sure that people in future generations can access it and use it?

And all of our great university research libraries face these issues together. So I hope this symposium will help us better understand these issues and understand the needs for scholarship in a digital age and the paths that we should take.

And I hope you will emphasize that in your discussions, the paths that we should all take on our campuses as we move into the years ahead. So I look forward to seeing how these discussions come out.

And again I want to thank you for being here, being on our campus and especially giving us wisdom and insight as to -- at least on our campus -- how we should be confronting these issues as important and challenging issues as we go ahead.

And I hope you'll take back to your campuses or your firms or your organizations some wisdom as well. So I do welcome you, and I hope you enjoy your time here.

It's now my great pleasure to introduce a dear friend of mine and a -- for at least me -- 29-year colleague at the Law School.

Roy, you've been there a great deal longer -- that's Roy Mersky, who has been a professor at the Law School, a chair holder at our Law School, and Director of
Legal Research and our librarian for many more than 29 years.

And I do hope you'll get over and see the law library. I was the Dean of the Law School for five and a half years. I was on the faculty for 29 years. And it is one of the truly premier law libraries in the country, a wonderful rare books collection.

And like so many of our libraries and institutions it owes it to great leadership over a long period of time. And Roy, you've given us that leadership. Thank you for what you've done, and welcome to the conference.

Thank you very much.

ROY MERSKY: I'd just like to say something. When we were organizing this conference I was told that I was going to introduce the president. So I spent a week preparing an introduction for Bill. And it's somewhat ironic, he introduced me instead, and I'm going to introduce another president.

I'll send you the introduction I wrote though. I'm pleased to introduce to you Dr. James Johnson Duderstadt, president emeritus and university professor of science and engineering at the University of Michigan. His auspicious academic career started on a note of uncertainty.
During the late 1950s in the small town of Carrolltown, Missouri, James Duderstadt played football on his high school's team. One day a telegram arrived for him, the first one he had ever received. It was from the football coach at Yale, with an offer of admission.

He was unfamiliar with the school. He thought it was in England, but he accepted. As a freshman at Yale he considered football to be of paramount importance, which led to poor grades. Resolving to improve academically the following year, he chose to major in electrical engineering and then regularly skipped football practice to study.

That's kind of a no-no here at Texas.

Eventually he quit the team, much to the dismay of his coaches, teammates and Yale football supporters. He went on to win the Chester Harding Plimpton prize given to Yale's top engineering senior and graduated summa cum laude.

From Yale he also received a master's in Engineering Science. His PhD is from the California Institute of Technology. As an Atomic Energy Commission post-doctoral fellow at Cal Tech, he developed nuclear-powered rocket engines for a manned mission to Mars.

He then joined the University of Michigan's Department of Nuclear Engineering. He served as dean of
the College of Engineering, as provost, vice president and vice president for academic affairs.

Interviewed in 1988 as the final candidate for the University of Michigan's presidency, he spoke in an open meeting about the country's need for a new model for research universities and stressed that Michigan should take the initiative to develop that model.

By a vote of six to zero he was elected president of the University of Michigan. He was instrumental in establishing Michigan's Millennium Project, a laboratory where new paradigms of learning institutions can be designed, constructed and studied.

Its mission is to provide an environment in which creative students and faculty can join with colleagues far beyond the campus to develop and test new paradigms for the academic community.

An analogue to corporate research and development laboratories, it's an incubation center, where new ideas concerning the fundamental missions of the university teaching research service can be developed -- not simply a think tank, but a do tank, where ideas lead to actual creation of working models or prototypes to explore the possible future of the university.

In 2004 the research center was named after Dr. Duderstadt and his wife Anne. Dr. Duderstadt's teachings
and research interests spanned a wide range of subjects in science, mathematics and engineering, including work in areas such as nuclear fusion reactors, thermonuclear fusion, high-powered lasers, computer simulation, science policy, higher education and information technology.

He has written numerous works including "A University for the 21st Century," in which he said, "Not only will social and technical change be a challenge to the American university, it will be the watchword for the years ahead.

"And with change will come unprecedented opportunities for those universities with the vision, the wisdom and the courage to lead in the 21st century." It's my esteemed honor to introduce Dr. James Johnson Duderstadt.

His vision inspires us to consider not whether a higher education will be transformed, but rather how those changes will take place and who will lead us in effecting them. Dr. Duderstadt.

JIM DUDERSTADT: Thank you very much for those kind remarks. I have to use Google myself, because my memory has gone blank on many of those issues. It's great to be back in Austin.

Unfortunately -- perhaps fortunately -- I was unable to accept the invitation from one of your former
regents, Charles Miller, to attend the Saturday night football game. But quite frankly I see enough of Ohio State from year to year anyway that I don't need to see them again.

For the past five years I've had the great privilege of serving on the selection committee for the Harrington Fellowship Program at the University of Texas. Beyond the extraordinary quality of the scholars that you've been able to recruit to Texas, either as visitors or more permanently for your faculty, what really impressed me in that effort was how every one of those applications pointed to the great quality and drawing power of your research libraries at the University of Texas.

I think it provides a very strong evidence of just how important these research libraries are for universities and for attracting the kind of outstanding scholars and teachers that we all aspire to. I faced a bit of a quandary about how to approach this particular set of discussions.

Since stepping down as president I no longer can master the alphabet soup of the library business: OCLC, ARL, RLG and so forth and so on. I read books; sometimes I write them. But I'm a little bit off track. But then it suddenly dawned on me that for the last ten
years I've worked and many ways lived in a library of the future.

During the last couple of years of my presidency we built this great big 250,000-square-foot facility that we intentionally tried to design as a library of the future. I got some very talented people, stayed out of their way and let them pursue their vision.

It does have books in it -- about a million books. But they're in the sub-basements and high-density shelving, and I doubt if they're used very much. It has about 1,000 workstations, a sound stage, recording studios, design studios, a very large pipe and Internet2.

And it's so popular to students that we have to operate it seven by 24, 365 days a year. It's the only building on our campus, other than our university hospitals, that's open all the time. It houses the servers for the J-STOR Project and another project that I'll tell you a little bit about, Project Sakai.

It also has a Starbucks -- actually not a Starbucks; we have our own coffee shop that I think is better. But that's very important in libraries of the future. So in a sense I've experienced firsthand over the last decade being stashed away, out of sight, out of mind by my university, what these are like.

But rather than present homespun stories of
what it's like to live in one of these places, I thought I'd move through a sequence of different perspectives that draw on much more recent activities, first at the level of the National Academies.

As some of you know, for the last five years I've chaired a series of studies by the National Academies, trying to understand the impact of information technology on the future of the research university.

In fact we had one of our sessions here a couple of years ago in this very building, in which The University of Texas hosted university presidents and executive officers from a number of other institutions for one of these discussions.

Next I'd like to move to the federal government, to the National Science Foundation, where the buzzword these days is cyberinfrastructure. I currently chair the Advisory Committee on Cyberinfrastructure for the National Science Foundation.

And it also provides input to the President's Council of Advisors on Science and Technology. So that will have something to say about this. Higher education - I know you're all waiting with bated breath to find out what the Spellings Commission report says, as it's released formally next week.

Although if you go to the website for Inside
Higher Education you can find it, because the report is considered a public document. As a commissioner on this commission charged by one of Texas's own, Margaret Spellings, Secretary of Education, I'd like to point out some of the implications it may have for research libraries.

I can't help but move to Michigan for just a moment. I'll say a few words about Google, but I'm more of a user than involved in the intricacies.

But then I'm going to link it to some other things: the Sakai Project, the OpenCourseWare initiative at MIT, the open learning initiative and a series of other activities that may be building the scaffolding for what Chuck Vest at MIT calls the metauniversity, that may be more effective in responding to the changing educational needs of this country and indeed the world.

The key themes. The first one is obvious. The technology which is driving much of this change is indeed a disruptive technology for the university. And in many ways the library's become the poster child for these changes.

Second, cyberinfrastructure -- that strange word that means hardware, software, people, organizations, policies that roll all the technology together -- is actually beginning to build learning environments and
scholarly environments that are not only ubiquitous and interactive, they are functionally complete. That carries a certain implication.

The open paradigm, open source, open content, open learning, open institutions, and of course open libraries -- another one of those themes. Universal access to knowledge and learning, and then finally whence and whether the university and the library itself.

The National Academy studies. Over the last several years -- actually throughout the late 1990's and the early part of the new century, there was a growing concern about what this fairly rapid evolution of information and communication technologies was doing to different parts of the intellectual infrastructure of the United States.

The National Academies had a very clear interest in that. And of particular interest was its impact on research universities.

There was a sense that, although the technology would present many challenges and opportunities, many of the significant issues were neither well understood, or, in many cases even recognized by leaders of our universities or those constituencies who support and benefit from them.

So they decided to launch a project in two
phases to explore this in somewhat more detail.

The first phase -- the acronym ITFRU, Information Technology and the Future of the Research University -- really involved a group of people that were people immersed in the technology, the chief technology officers of companies like IBM, Xerox and so forth, as well as university leaders and people in public policy.

The first phase was to identify what might happen over the next decade or two and then move from that to looking at implications. That first phase surprised us a bit, because looking out as far as we could see into the future, we saw no signs that the exponential increase of power in this technology -- the so-called Moore's Law -- would slow down.

In fact all the signs seemed to suggest it was accelerating. The term was riding the exponential -- that is, that part of the technology -- whether it's processing power, bandwidth, storage, whatever -- which was increasing the most rapid was changing the nature of technology and how it affected society.

In thinking far ahead we saw the evolution from giga to tera, and now, as Texas is well aware with its recent success, the petascale initiative of ten to the 15th operations per second, or storage or whatever. And in fact people are now beginning to talk about exaflop
machines, ten to the 18th. That'd be about a hundred times faster than the brain is capable of processing.

Of more interest perhaps to the research library community is data storage. Today's current generation of desktop machines -- the machine on my desk has a terabyte worth of memory, and that's expanding rapidly.

One of the members of our group, Stu Feldman, who was one of the senior technical officers of IBM at the time, got the presidents' attention at an AAU meeting by asking them how they would approach the construction of new libraries if he could offer them a device the size of a football that would contain the Library of Congress. And he guaranteed he would be able to do that within a decade. Most university presidents understand how big a football is, you understand. They might not understand the iPod, but the football they understand. More to the point, as I'm sure you're aware, the entire holdings of all academic research libraries is about one to two petabytes.

And that's the iPod of about a decade from now. And if you really want to go to the extreme, Hal Varian, who many of you know at Berkeley, estimates the total number of words spoken in the history of humanity at five exabytes.
That's an extrapolation of memory perhaps two
to three decades into the future. How do we ship this
about? Well, a couple of weeks ago Internet2 announced it
was combining with the Department of Energy's Science Net
to build terabit-per-second networks. Larry Faulkner, as
you may know, is the chair of the governing board for
Internet2 and has played a very significant role in that.

But bringing it back down with WiMAX, we’ll get
70 megabits per second over five- to ten-mile distances.
So it really will be ubiquitous.

So, first finding: the technology is moving as
fast as it has over the last several decades, and may in
fact accelerate. Second finding: the impact of IT on the
university is likely to be profound, rapid and disruptive.

That's a phrase that comes from again one of
Texas's former faculty members and leaders, Marianne Fox.
It will change teaching, research service organizations,
academic structure, faculty culture, financing management
and so forth.

Now, we all know that if change is gradual,
institutions can adapt. But this disruptive technology
term really goes back to Clayton Christianson's book, The
Innovator's Dilemma, in which it suggests that in rapidly
changing technologies, they may look at first inadequate
to displace existing technology, but they explosively
later displace traditional applications, because they enable new ways of doing things.

Third finding indicated that, although information technology would present many complex challenges and opportunities to universities, creating considerable uncertainty -- procrastination and inaction, which of course are decisions in and of themselves, are perhaps the most dangerous course of all to take.

So that moved us into the second part of this, which was to create a forum involving: the leadership of research universities across the country through AAU, the presidents, its companion organization for provosts, and then a series of workshops, which we call executive leadership core workshops, where we would bring together four to five institutions that would commit to a day-and-a-half -- president, provost, head of the libraries, CIO, deans, chief financial officer -- kind of sit around the table, and in a structured dialogue with the blinds pulled down, compared their very candid assessment and their strategies.

We hosted a number of these around the country and involved about 25 institutions in this. And as I said, one of these was at The University of Texas. One of the things that came out of these time after time after time was the degree at which people were looking at the
research library -- as I said early, the poster child -- as the harbinger of the kind of changes that would occur.

And to give you a sense of this, let me relate to you a discussion that we had, not in Austin, but rather at Chapel Hill, and you can guess the institutions. The question that was asked was, how many of the presidents around the table were currently planning to build more libraries.

Of course all the hands went up. Well, what are you going to put in them? Are you going to put books in them? Well, no, we're going to put books off campus in high-density storage, retrievable facilities. Well, what are you going to put in your libraries?

And after a little bit of discussion it turned out the only common denominator was a Starbucks. All of the libraries would have a coffee shop. And that's because the libraries of the future that they were building were seen as community centers, where students and faculty came to study and to learn, but where books were largely absent.

In a sense the library was becoming a people place that provided the tools, the services, the expertise to support learning and scholarship, but along with that, an environment for social interaction. So it raised the obvious question of what is the university library in the
digital age.

Is it built around stacks or around Starbucks?

Is it a repository of knowledge? Or is it a student union for learning? And in fact we concluded that maybe we weren't talking about libraries at all. Maybe it was the kind of physical spaces that universities would require for learning communities.

In fact one of our participants suggested that, if today the common element of every library is a Starbucks, perhaps with massive digitization and distribution of library holdings, soon every Starbucks will have a library.

That is, indeed you can take your coffee cup and either your iPod or your WiMAX connectivity and access the knowledge of the world through the internet. But beyond that it was also clear that the library itself has become the most important observation post for understanding how students really learn these days, and how faculty utilized knowledge.

In fact if the core competency of the university is the capacity to build collaborative spaces, then the changing nature of the library itself may be the paradigm for the changing nature of the university.

Second subject: National Science Foundation.

My next-door neighbor and the only other
faculty member that has an office in this big library of the future is Dan Atkins, who many of you know was the founder of our School of Information.

But beyond that he chaired a very important blue-ribbon study several years ago for the National Science Foundation on the nature of information technology, which utilized this term, cyberinfrastructure.

The concern was that we were approaching an inflection point in which the potential of this technology was going to transform how scholarship and learning occurred, particularly in the sciences.

Well, the conclusion of the report that they came out with I think is quite important. And let me quote, "A new age has dawned in scientific and engineering research, pushed by continuing progress in computing, information and communication technology and pulled by the expanding complexity, scope and scale of today's challenges.

"The capacity of this technology has crossed thresholds that now make possible a comprehensive cyberinfrastructure on which to build new types of scientific knowledge environments and pursue research and learning in new ways and with an increased efficacy.

"The emerging vision is to use cyberinfrastructure to build more ubiquitous,
comprehensive digital environments that become interactive and functionally complete for research communities in terms of people, data, information, tools and instruments."

Stress the terms again: comprehensive, functionally complete, ubiquitous. Now, they pointed out that as with most of technology evolutions the real impact occurs not through the technology change, but through human behavior and organizational change.

And therefore they stressed that the National Science Foundation efforts in the future should look at all of these components as they try to build the appropriate environments for scholarship and for learning. Well, of course no good deed goes unpunished.

And in return for leading this very successful panel, Dan Atkins this last spring was named the first director of the Office of Cyberinfrastructure at the National Science Foundation, where he, among other things, leads efforts such as the petascale initiative.

I in turn -- in part because of Dan and in part because of another Michigander, Arden Bement, who's the director of the Foundation -- was asked to chair the advisory committee for this. Now, we've only had one meeting.

But that meeting, I think, suggests that this
will have a profound implication for libraries, because one of the most difficult aspects of this is the explosion in the amount of knowledge being generated by scientific activities, by scholarship on campuses and so forth.

And the belief is increasingly that creating and managing the new types of archival and support material will really be the responsibility of institutions themselves. In a sense we should not expect national repositories for this kind of petabyte-scale information; but rather universities, either on their own or through consortia, will have to develop these. And these will be a component of what you might call the library of the future.

Third topic: the National Commission on the Future of Higher Education in America. That was launched by Secretary of Education Margaret Spellings almost exactly a year ago. Although the final report will not be presented to her until, I think, it's next week, since the report itself is now in the public domain, let me make a couple of comments about it, and then highlight a couple of things.

It's kind of a good news-bad news story. The good news I suppose is best reflected by that pithy little paragraph in the Economist a year ago. "There's no shortage of things to marvel at in America's higher
education system, from its robustness in the face of external shocks to its overall excellence. However, what particularly stands out is the system's flexibility and its sheer diversity. It is all too easy to mock American academia, but it's easy to lose sight of the real story, that America has the best system of higher education in the world."

The bad news is this isn't good enough -- not nearly good enough. "Despite these achievements" -- here and now I'm quoting from the report -- "the Commission believes U.S. higher education needs to improve in dramatic ways.

"Our year-long examination of the challenges facing higher education in America has brought us to the uneasy conclusion that the sector's past attainments have led our nation to unwarranted complacency about its future.

"What we've learned over the last year makes it clear that American higher education has become what in the business world would be called a mature industry, increasingly risk adverse, at times self-satisfied and unduly expensive, and has yet to successfully confront the impact of globalization, rapidly evolving technologies, an increasingly diverse and aging population, and an evolving marketplace, characterized by new needs and new
paradigms."

Now, there's a lot of fine detail with a lot of devils throughout the report, because in part it was written by consultants that had very little to do with the commissioners. But the commissioners themselves did, essentially through revolution, take control of the process and converged unanimously on six major recommendations.

I'll summarize them very briefly, and then I'll connect them back to the issue here. The six are: to remove barriers to access and success; every student in the nation should have the opportunity to pursue post-secondary education; to restructure financial aid -- the entire student financial aid system, federal, state and institution, is simply not doing the job, because it's channeling dollars to those students who don't need help at the expense of those who do.

Transparency, accountability and public purpose -- and you should thank Regent Miller for this -- to urge the creation of a robust culture of accountability and transparency throughout higher education.

Investing in innovation both on the part of institutions and on the part of the government. Lifelong learning, recommending the development of a national strategy for lifelong learning.
And finally, responding to the needs of a global knowledge economy -- and there we put in a recommendation which kind of gets behind the existing National Academy efforts -- it goes under the acronym of Rising Above the Gathering Storm or the American competitiveness initiative.

The key themes here for the purpose of this discussion, I believe, are the following. That every student in the nation should have the opportunity to pursue post-secondary education, number one.

And number two, every American should have the opportunity for whatever post-secondary education they need, aspire to, merit in an increasingly competitive, global knowledge-driven society, throughout their lives.

Put another way, I think one of the themes in this is to take the final step in America's sequence of steps over the last two centuries to expand educational opportunity -- from the Land Grant Act to universal secondary education to the GI Bill, to the Equal Opportunity Acts -- and provide all Americans with universal access to lifelong learning as essentially a civil right appropriate for a knowledge society.

A very bold vision, a very challenging goal. How do you achieve it? Well, it's not in that report. So let me go on and provide some other --
Now back to Michigan. Several things going on at Michigan and at other institutions around the country, including The University of Texas, I think, may point the way to how this particular challenge might be faced.

One of them clearly is the Google print library project. I'll just briefly summarize it, because many of you know and have actually read the Michigan contracts in much more detail that I have. But five institutions in December 2004 announced an agreement with Google to digitize a substantial part of their book collections -- Michigan, Stanford, Harvard, Oxford and the New York Public Library.

Last month the G5 became G6 with the addition of the University of California. At Michigan we committed to digitizing our entire book collection -- about 7.8 million volumes. Now, the fact that Larry Page was an undergraduate engineering student at Michigan during the 1980's I assure you had everything to do with this project.

He had been bouncing this off of us for a couple of years. We thought first, that's a crazy idea. Then we said, we'll, I don't know. We were involved in the J-STOR Project sponsored by the Mellon Foundation. So that gave us a good head start.

The original thought was that it would take a
decade. That estimate is now down to about six years. And Google is moving up the learning curve very, very fast. So it may happen much sooner. Digitized materials from the collection came online in June.

And in fact through its own copy, Michigan receives as part of the deal, called MBooks, it's now beginning to actually download full-text copies of books that are in the public domain, no longer under copyright.

I went through this -- I might add -- because I had a 30-year-old textbook that still is under copyright, that the publisher wanted me to do an update on. Since it had been put together before the days of word processors I didn't know what to do about it.

So I said to the librarians, Could you go ahead and hurry up and digitize that part. And sure enough, there it is scanned by Google. Anyhow our provost, Paul Courant, at the time, who was very instrumental along with our associate library director John Price Wilkins, put it this way.

"Our purpose is to extend the realm of ideas in the broad service of society." Part of this was the concern that most of our vast collection, as in the vast collection of your research libraries, are orphan works -- works that can no longer be easily identified in terms of copyright ownership.
In our case on the order of about 80 to 85 percent. And therefore in the kind of world in which people -- students and scholars -- approach searching for information digitally -- Google, Wikipedia, whatever students use these days -- they were likely to kind of pass out of sight.

So part of it was simply to provide access to these materials. But I will tell you that our provost believes that part of the reason as well was to Michigan to take head-on the grand challenge to digital access that is copyright.

Now, that is nothing new for my institution. I ended up as a named defendant in the case that made it before the Supreme Court in 2003 over affirmative action. University of Texas with the Hopwood case also faced some of that challenge.

I can look at other examples: NSF Net in the 1980s which led in part to the internet; Sakai, which may also end up in court -- as we'll see in a moment -- challenging the intellectual property claims of monopolies such as Blackboard.

So we're used to going to court about things and getting sued. So maybe we will on this one as well. Fortunately I'm not at the helm on something like this. Courant believed very strongly in pushing this. As he put
it, the world of Google is really emblematic of something very important.

And that's the world of ubiquitous, indexed digit content. Google's arrangements with Michigan and other libraries essentially are going to teach us how we have to live in that world, that will encompass both material that started as print, as well as material that was or will be born digitally.

It's quite a challenge obviously. I'll come back and make a couple of comments about it in just a moment.

The Sakai Project, again yet another one of these open sources projects -- in this case a consortia of universities and corporations: IBM, Apple, Cisco, Unisys -- that is building the open source to support learning and scholarship, with some efforts to move that up to the enterprise level -- more broader activities involving universities -- and perhaps down to the desktop level for scholars.

That one actually is right above me in the building I live in. And they in fact right now are in the process of developing yet another tool which will support yet one-third of these open initiatives, which is the OpenCourseWare initiative, initially promoted through MIT.

As some of you may be aware, several years ago
when asked, the MIT faculty decided that their commercial or proprietary rights over the digital assets behind their instruction was more valuable in the public domain than it was to them personally.

And therefore with the help of the Mellon and Hewlett Foundations they have put roughly 2,000 of their courses, the background material, into the public domain. As Chuck Vest notes this is not a distance-learning method effort, but rather a web-base publishing venture.

But it's one that's now being used by millions of students around the world. And increasingly it's being a model that's followed by a number of institutions in the United States and abroad.

Coming back again to Sakai, that's the reason why -- since most universities won't have the $20 million that you get from the Hewlett Foundation and the Mellon Foundation -- we have to provide cost effective ways that they can capture those materials and make them available to the world.

All of this together with other open initiates -- the Open Learning Initiative at Carnegie-Mellon, the Open Knowledge Initiative which was a precursor to this -- even long-standing efforts such as the open university philosophy of the British Open University, which removes the traditional barriers and
constraints on admission and enrollment -- all of it is coming together to create what Chuck Vest likes to call the metauniversity.

Open source, open content, open learning and other open technologies raise the possibility of developing a scaffolding on which one would build truly global universities.

After all when you look at the need for higher education at the global level, the huge scale of it, the great diversity of cultural, political, economic context, the financial issues and so forth, suggest that the old paradigms of campus-based learning just are not going to be sufficient.

In fact the metauniversity -- again as Vest puts it -- a transcendent, accessible, empowering, dynamic, communal-constructed framework of open materials and platforms on which much of the higher education worldwide can be constructed or enhanced -- is in fact taking advantage of the great strength of cyberinfrastructure to distribute knowledge and learning opportunities and to do it to the world.

But -- which takes me to the next part -- to achieve this vision it's going to have to overcome not only some formidable obstacles in terms of existing mind-set, but also some very powerful monopolies. Let me give
you a case in point.

Two weeks ago -- actually a couple of weeks after the Spellings Commission had approved the end of a very difficult process to construct a final report, which we felt we could sign off on -- and 18 of the 19 commissioners signed off on it with the understanding that nothing will be changed -- late one Friday night we got a message from one of the commissioners that apparently read the report once again in great detail, or perhaps had someone from her organization read it once again in great detail, and found that she could no longer accept some of the language in the report.

The language was this: "The Commission encourages the creation of incentives to promote the development of open source and open content projects at universities and colleges across the United States, enabling the open sharing of educational materials from a variety of institutions, disciplines and educational perspectives.

"Such a portal could stimulate innovation and serve as a leading resource for teaching and learning. New initiatives such as OpenCourseWare, the Open Learning Initiative, the Sakai Project and the Google book project -- uh oh, trouble there -- hold out the potential for providing universal access to both general knowledge
and to higher education."

Well, we did have significant representation from industry on this among the commissioners. This was a company in Seattle that has a little bit of difficulty with a couple of sensitive words in this: open source and the name of another company.

And so after a wild weekend of email negotiation we were finally able to get the Commission to agree on compromise language. And again I quote -- and this is what's in the report -- "The Commission encourages the creation of incentives to promote the development of information technology-based, collaborative tools and capabilities at universities and colleges across the United States, enabling access, interaction and sharing of educational materials from a variety of institutions, disciplines and educational perspectives.

"Both commercial development and new collaborative paradigms, such as open source, open content and open learning, will be important in building the next generation of learning environments for the knowledge economy."

So we kind of take out the names of specific programs. But open, open, open is in there because we feel that may be the quickest route to the future of higher education, at least on the global level. That was
Well, this takes us into almost the metaphysics of the issue of ownership. Here I would remind you of that great quote from John Perry Barlow in Wired magazine many years ago.

"The enigma is this. If our intellectual property can be infinitely reproduced and instantaneously distributed all over the planet without cost, without our knowledge and without its even leaving our possession, how can we protect it?

"How are we going to get paid for the work we do with our minds? And if we can't get paid, what will assure the continued creation of distribution of such work?" Or even better Thomas Jefferson. "Ideas should freely spread from one to another over the globe for the moral and mutual instruction of man and improvement of his conditions. This seems to have been particularly and benevolently designed by nature, when she made them like fire, expansible over all space without lessening their density at any point, and like air in which we breathe, move and have our physical being incapable of confinement or exclusive appropriation."

It seems like a clash between the irresistible force of the open knowledge movement against the immovable object of intellectual property ownership. But another
quote from Kevin Kelly, also from Wired, who writes very provocatively this time in spring's New York Times.

"The courts may haggle forever as this complex issue works its way through the top. In the end it won't matter. Technology will resolve the discontinuity first. The Chinese scanning factories, which operate under their own looser intellectual property assumptions, will keep churning out digital books.

"As scanning technology becomes faster, better and cheaper fans may do what they did to music and simply digitize their own libraries. But it will adapt eventually. The reign of the copy is no match for the bias of technology. All new works will be born digital. And they'll flow into the universal libraries you might add more works to the long story." And of course Kelly's concept here is the age-old quest for universal access to all human learning.

The Library of Alexandria, which for a very brief time he points out, had between 30 to 70 percent of all books in existence, disappeared, obviously. But perhaps with massive digitization we're beginning to actually approach that possibility once again.

WorldCat suggests there are 32 million books in the world. And the analysis of the G5 suggested that the unique volumes that could potentially be digitized from
those libraries amounted to about one-third of the world collection. And to that of course is now added the University of California. As Kelly said, if we can provide all of the works of humankind to all of the people of the world, it will be an achievement remembered for all time, like putting a man on the moon.

Well, what is the future of the library? Maybe the future of the library is not simply that of the university but of civilization itself. The key here is not providing open access to the digital records of all of the works of humankind. But rather it's the linking together of a substantial part of the world's population with limitless access to knowledge and learning opportunities that's enabled by this rapidly evolving cyberinfrastructure, again increasing 100 to 1,000-fold every decade, decade after decade after decade.

I think that if you go to the cinema you'll see any number of science fiction movies that point out the horrors of artificial intelligence, the possibility that even the Terminator will become governor of the State of California.

But in reality what we're talking about here is not artificial intelligence, but a new form of collective intelligence, in which billions of people are interacting in a very direct and robust way with access to the
knowledge of humankind, unconstrained by today's monopolies on knowledge or learning opportunities.

And maybe this is the most exciting vision for the future of the library and perhaps for the university itself, because no longer constrained by space, time or monopoly, but rather unleashed by cyberinfrastructure, perhaps this is the new kind of a global learning community that will evolve over the course of the next several decades.

Well, enough said. I can make some further comments about the future of the university. But I think it's probably better to open it up for discussion, because I probably provoked a number of you with these comments. So let me stop right now and answer questions if I may.

VOICE: [inaudible] role of the library --

JIM DUDERSTADT: I think what you'll see is that clearly the responsibility to collect unique materials will define and characterize certain research libraries. And that gives them their uniqueness. And as I mentioned at the outset, that's what attracts very unusual scholars to campuses.

On the other hand, the broader knowledge will be a collective knowledge and increasingly will not be viewed at a particular place. But rather will be in the ether maintained by a collection of institutions. How
those collections come together -- obviously it'll involve the commercial sector to some degree, whether Google is the glue that pulls it together, whether it's through federal activities.

To build the petabyte per week kinds of repositories that the large hadron colliders are going to require from Geneva as it comes online is something that the federal government is going to have to figure out how to do in order to support high energy physics.

So it may come in part from the federal government, and it may come from institutions themselves. So I think the thing is to just aggregate those very unique materials -- our papyrus collection at Michigan for example or the kinds of things you have -- from that roughly one-third and growing fraction of the world's collected wisdom in books, and then adding to that other kinds of multimedia materials that I think will have to be a shared resource.

And institutions in the way that they cooperate today and make major investments will have to continue to do that in the future. And that requires some kind of collective agreement.

VOICE: [inaudible] language -- describing.

JIM DUDERSTADT: Well, good. You mean we did something right. We're probably going to take as much
heat for the Spellings report as you folks will take for losing to Ohio State on Saturday night.

VOICE: [inaudible] is there anything you could tell us?

JIM DUDERSTADT: There are 19 commissioners on the Spellings Commission. And five of them were kind of out of higher ed: David Ward, Chuck Vest from MIT, Bob Zemsky from Penn, myself and Charlene Nunley from a large community college in Maryland.

So we were kind of the higher ed mafia that were horrified at the release of the first draft of the report that was done entirely by consultants, did our best to keep it from ever seeing the light of day -- which of course it did -- and then went through a series of efforts to patch it up.

We had enormous help from people like Sara Martinez Tucker, who will actually have to implement much of this as the new undersecretary-designate of education and many others. But in part it was to recognize that it was a very, very diverse commission.

They saw things in many different ways. Now, what we didn't want to happen is for the report to be so hostile -- which of course the first draft was -- that those of us from higher education would have to repudiate it and either write a minority report or simply refuse to
sign it.

David Ward was with us as we moved -- I mean there was a tremendous amount of writing and rewriting and negotiation over every bit and piece of it. And there's still a lot in it we don't like. But nevertheless we felt we could support the recommendations, as did David Ward, president of the American Council of Education.

In the end however, some of the language in it that we tried to get changed, we really couldn't get changed, because of the time scale. I think if we'd had a little bit more time, we might have gotten it changed. David was sufficiently uncomfortable, as were the organizations that essentially he reports to through ACE, that he decided it was best not to sign.

In retrospect I tend to agree with that as being a wise decision, because it sends out the signal that while the 100,000-foot part of the report is something we can support, there are a lot of details there that we do not support.

And therefore it should not be portrayed as something that has unanimous consent. And that allows us then to challenge some of those parts. In all candor it's not a very well written report. Usually things written by committee are not.

The summary of the report was not written by
the Commission. But the findings and the recommendations were written by the commissioners themselves. They took control of that process.

As to what will happen, again since the federal government really provides only about 20-25 percent of the support of higher education in this country, their leverage is pretty limited, and particularly with a Congress that has to approve anything.

So I'm not sure that you're going to see them having great influence over it. It's our hope that the principal response to it will come from higher education itself, because the two themes in it that drove everything were first the themes of social justice, the fact that there's a significant fraction of the American population -- particularly those in the lower two economic quartiles -- that are simply not being served by the way higher education is currently structured and is approached in this country.

And second, that more broadly the challenges of the global knowledge economy -- Friedman's flat earth -- are very compelling and will require a very significant change to meet the changing needs of the nation. And that was the second major theme that drove this.

So I'm hoping that if we can keep it at the 100,000-foot level through higher education itself, that
we can get some significant things accomplished.

    Just nuking the federal financial aid system
would be positive -- right -- if we could restructure
everything in terms of Pell Grants and so forth and get
the dollars to people with need, rather than subsidizing
the wealthy, which the commercial lending -- I'll be
provocative -- which it does right now.

    But nevertheless we'll see more as people
become more aware of it. There will be rollout over the
next two or three months, with the Secretary carrying much
of the load. And we'll see what the Secretary's decided
to push during the last two years of the Bush
administration.

    And those of us on the Commission are already
being lined up to appear before a number of groups. If
you hear from 19 different commissioners, you will hear 19
different perspectives of what really is said and what's
missing.

    VOICE: [inaudible] career -- individual
faculty? Where the rubber meets the road, so to speak.

    JIM DUDERSTADT: There are several aspects in
it. Let me say the Spellings Commission, and then I'll go
back to the library of the future. The Spellings
Commission -- what probably got the faculty members'
attention most earliest was the suggestion that maybe this
would be a No Child Left Behind -- kind of a combination
of a Nation at Risk and No Child Left Behind, brought to
you by a prominent Texan who led those in the State of
Texas.

And in fact at one point those consultants even
floated a particular testing device, which is this
collegiate learning assessment tool. That was strongly
opposed by most of the commissioners.

But what remains in it is a challenge to
faculty, that is their obligation to clearly define in a
public way the objectives of their educational program,
and then to devise a way to provide evidence, publicly
once again, of how they are doing and meeting those
objectives, but leaving that up to the faculty to
determine.

Now, that's not really that unique. There are
a couple of the regional accreditation groups that already
have language like that. They put it in two or three
years ago. And they are holding faculty's feet to the
fire of those institutions undergoing accreditation review
right now.

I led one of those for UC Santa Cruz. And I
must say the faculty response to that was very, very
encouraging, very sophisticated and very important I think
for the future of that institution. So in any event I
think to the degree that the faculty's challenged, this is their responsibility. It's not something for the government to tell them what to do. But it's something they have to do. That could be constructive.

Back to the library of the future and so forth, you know one of the strange things is that we look at this as a revolution in higher education. But most of the university hasn't changed very much over the last -- I mean the academic programs are the same way as they've always been. We finance it the same way. We organize our degree programs. But what has changed is the day-to-day work of the faculty.

I mean, how many times do faculty visit the library any more? The first thing they do is they go to a search engine -- sometimes they go to Wikipedia for that matter -- to get some guidance and move beyond that. They no longer talk to the person across the hall.

They're in almost instantaneous interaction with people all around the world. So the life of the faculty member is changing enormously. And from our experiences in visiting a number of campuses, the same is true with the students.

This is the plug-and-play generation. They are always on, always connected, forming their own learning groups, much more active in defining their learning
environments and much more reluctant to accept the kind of passive learning that all too often the lecturer and our curriculum seeks on their part.

So things are changing on the part of faculty and students very, very dramatically right now and very fast. It's only a question of time before the institutions themselves change. Let me kind of conclude by going back to again one of these experiences.

We had two round-table sessions, two days each, of the AAU presidents and the AAU provosts. Actually to get the presidents in the right mood, we had Lou Gerstner, former CEO of IBM, to come in and talk a little bit about how IBM was a company that was just about to fall apart when he came in as CEO, in large part because although they developed the technology, they didn't take it seriously themselves.

And he suggested that maybe universities were doing the same. Well, the best way to have a meeting of presidents is to let them do the talking. Then they'll listen to themselves. They don't respond very well if you talk to them.

But I think there was some constructive activities that came out of that. But the interesting ones were the provosts. We were actually able to get onto the table in the provost discussions -- we had about 50
provosts there -- the ultimate question of whether the research university would even exist -- at least in terms that we could understand it -- a generation from now. Presidents would never touch that. But provosts were at least willing to touch it. The comment that came up at that point, actually from Susan Lohmann who's a professor at UCLA, that if you look at that period from the Civil War to a generation afterwards in this country, higher education changed in almost every way that it could, from the kind of British boarding school models of the colonial colleges to the great land grant universities, faculty empowerment with academic freedom, from institutions of several hundred to several thousand students, the research mission, the service mission and so forth.

Everything that could change changed. It all changed in a generation. Most of the provosts believe that higher education is entering a similar period of change, in which the university as we understand it -- and that's kind of the research university -- will be so totally different a generation from now, that we probably won't even be able to recognize it or describe it in the terms we use today.

I thought that was very provocative. Maybe that's because provosts see institutions in a much more
hands-on way and kind of sense some of these changes that are occurring at the grassroots level. At the front of that change will be the research library.

It's not only going to be the poster child of the thing we look at. But it could well be the organization that is driving much of the change. What an exciting time to be in this business. Right? Thank you.

[TWENTY MINUTE BREAK]

DON CARLETON: It's my privilege and delight to introduce our next speaker, Clifford Lynch, who is going to give us his thoughts on the impact of digital scholarship on research libraries and archives.

It's hard to imagine any attempt to conduct a serious discussion about the research library or archive in the 21st century without having Clifford Lynch play a leading role in that discussion.

Since 1997 Dr. Lynch has been the director of the Coalition for Networked Information or CNI as many of you know, which is jointly sponsored by the Association of Research Libraries and EDUCAUSE. CNI includes approximately 200 member organizations concerned with the use of information technology and networked information to enhance scholarship and intellectual productivity.

Prior to joining CNI Clifford Lynch worked for 18 years in the office of the president of the University
of California, the last ten as director of library
automation. Dr. Lynch continues to serve as an adjunct
professor at UC Berkeley's School of Information
Management and Systems.

Clifford Lynch is a past president of the
American Society for Information Science and a fellow of
the American Association for the Advancement of Science
and the National Information Standards Organization. He
currently serves on the National Digital Preservation
Strategy Advisory Board of the Library of Congress.

And he was a member of the National Research
Council group that published The Digital Dilemma:
Intellectual Property in the Information Infrastructure
and Broadband: Bringing Home the Bits. He now serves on
the National Research Council's Committee on Digital
Archiving.

Clifford Lynch is much published in a variety
of media. In fact it would take me the rest of the
afternoon to list Dr. Lynch's publications, not to mention
his lectures, some of which are on podcasts. Happily for
us however, we do not have to hear Dr. Lynch on our iPods,
because we have him here in the flesh.

Please join me in welcoming Clifford Lynch.

CLIFFORD LYNCH: Podcasts. I never thought
about actually the need to enumerate those -- a new worry.
Something we can all worry about now is bibliographies of podcasts. It's a pleasure to be here.

It's always a pleasure to follow Jim, who I think has done an incredible job of laying out this enormous, complicated landscape that we're operating within. And I think that backdrop is going to be very helpful to my comments.

I want to talk a bit this afternoon about the futures of the research library. And I use that word, futures, plural, very deliberately.

I think that we may see a world ten, 20 years from now where there is considerably greater diversity among our research libraries, as a result of deliberate policy choices that the leadership of those research libraries and the universities in which they're embedded make in the next, let's say, about five years.

That's my suspicion. And I don't want to necessarily argue that one specific future is always the right future, and everybody not choosing that is making a dreadful mistake, but rather to illuminate some of those choices, and leave it to all of us to talk about the pros and cons of those choices.

Now, before going any further, I want to first constrain my comments a little bit. I'm going to talk about the future of research libraries and specifically
research libraries that are embedded within research
universities. There are, to be sure, research libraries
that are not embedded within research universities. And
we indeed have at least one of those represented here.
Those institutions I think face some somewhat different
challenges, although certainly some common challenges and
opportunities as well. But I think that there are some
very specific questions there, which I'm not going to try
and cover.

Secondly, I'm going to deal with research
libraries and not academic libraries broadly. I think
that the gap between most research libraries and most
academic libraries or most libraries in higher education
institutions -- and let us just remember there are a whole
lot more higher education institutions in this country
than there are research universities. Those academic
libraries are going to change in dramatic and sometimes, I
think, in almost unrecognizable ways rather quickly. When
you start thinking about what a library embedded in, say,
a community college going to look like in another five or
ten years.

I think that the answers there get very, very
different than the kinds of things that we think about
with the future of research libraries. And I'm again
going to leave those off the table.
I do want to sort of note the question of the future and the purpose of the research university itself, which was well framed in the earlier remarks, and to say that what happens there is obviously going to be a factor here.

But it's one that again I want to leave largely out of scope in my remarks. Will there be as many research universities in the next generation as there are today? I'm not sure.

It does seem very clear that, when we look at the sort of much broader system of higher education and how that connects to the system of K through 12 education in this country, that there are some things there that are pretty acutely stressed and that -- at least from my rather ignorant view -- there's liable to be some changes here over the next decade.

But changes or not, I think that we will continue to see a core of institutions that are recognizable as research universities, and within them we will continue to see great research libraries.

My last environmental caveat -- and this is one that again I think was touched on, but I just want to underscore -- there some questions that are starting to be framed both in the United States and elsewhere, particularly to my knowledge in parts of Europe, about
what really is the purpose of the research university.

Clearly there's consensus that one of the purposes is to create knowledge. Another is to pass that knowledge on through teaching. There's I think a more debatable issue about how much its mission encompasses the dissemination of knowledge, particularly the broad dissemination of knowledge to the public.

How much, for example, one can think of things that have historically been publishing, largely outside the purview of the university, as moving into core missions of the university. And I do hear that debate starting to surface in various areas.

The resolution of that debate -- and again I would expect that it will not be resolved uniformly, but rather the different research universities will make different choices -- is I think going to be a factor in shaping the future of research libraries at those universities.

And I'll have a little more to say about that shortly. So let me move on, and let me connect this question of the future of the research library with questions of what I believe the program calls digital scholarship or some such.

The fundamental point is that changes in scholarly practice writ large -- and that includes changes
not just in the practices of doing research, but indeed of
teaching and learning as well in higher education -- are
going to shape the future of the research library.

Talking about “how will the research library
evolve” as if it has a sort of an intrinsic right to exist
and an ongoing destiny, I think is really misleading.
That destiny is, I think, going to be absolutely shaped by
these environmental changes.

Paul Courant, who is I guess starting to emerge
as a presence at this meeting, even though he's not here,
has made the statement a number of times that there's no
scholarship without scholarly communication. And I think
that that's an important observation.

It's clear that the changes in scholarly
practice are profoundly intertwined with changes in
scholarly communication. And there again this brings in
the question of roles of libraries.

I think that one absolutely critical point to
underscore is that the gap between scholarly communication
and -- if you'll permit me the term -- traditional
scholarly publishing is growing ever wider.

With the broad landscape of scholarly
communication dwarfing the much more narrow confines of
the province of traditional scholarly publishing, that gap
is going to become of profound and growing interest I
I also want to underscore the theme of convergence here. I think when we speak of the future of the research library in the setting we're talking about, we're equally focused, and it's equally important to think about the future of the entire portfolio of cultural memory organizations within our institutions.

That means not just the libraries, not just the archives, but also the museums, and beyond the museums some of the museum-like collections that are embedded in departments. It also means trying to sort out and understand the roles of some things that have emerged in the last half century, that have certain kinds of cultural memory behaviors, but don't fit comfortably in here.

A leading example that I can think of is public broadcasting, which often is sitting on enormous, often almost completely unexplored archives of key scholarly source material. Those need to go into the question as well.

So when we think about what the future of the research library is in our institutions, I think it's critical that we think about that in the broader context of the institutional strategy for all of its content management organizations, all of its organizations for dealing with the management of scholarly evidence and of
I think looking at individual bits of that in isolation is increasingly painful, particularly as it becomes a little bit more difficult every year to confidently say what ultimately distinguishes one of these institutions from another.

Okay. With all that setting and groundwork let's talk just a little bit about what's going on with scholarly practice. We heard a discussion already about cyberinfrastructure or e-Science as they prefer to call it in most of the rest of the world perhaps a little bit more clearly.

I don't really think I have to trot out all of the poster children of e-Science for you, all of the new particle accelerators and space telescopes and other national-scale or international-scale facilities.

I think that it is worth reminding you though that in so many fields of science, this has permeated not just the, you know, the sort of hero marquee projects, but in fact the work of the vast majority of scientists working on the scale of, you know, individuals labs, modest-sized research groups.

If you look at the current scholarly practices in molecular biology, in chemistry, in engineering, in astronomy, you'll find that more and more and more there
is an adoption, an infusion of information technology into the day-to-day working practices of those scholars.

And beyond that there is an increasing move to build up community databases of information out of those tools. Think of things like the various molecular biology databases. Think of things like the national virtual observatories that are currently being federated in astronomy.

Yes, you can still find, you know, the scholar in the sciences whose involvement with information technology is primarily searching literature databases and cranking up the word processor and trading electronic mail and using the web browser to look at things.

But I think that that cadre in most of the sciences is diminishing. Furthermore I think it's important to underscore that, at least my sense is, broadly across the sciences the move to e-Science has been accepted on a disciplinary basis.

It is not controversial to be using these tools. It is expected that you'll use these tools when appropriate, and that these are good tools for doing good science and advancing scholarship.

If we look though across the rest of our research universities, beyond the sciences and engineering, I believe that you will also find a rapidly
growing use and reliance upon the tools of information
technology, of cyberinfrastructure, of massive data
resources to advance work in the humanities and the
qualitative, rather than quantitative, social sciences as
well.

We heard earlier about the Atkins report, which
at least in the United States is sort of the touchstone
for articulating that vision of how information
technology, advanced networks, high-performance computing,
shared observational and experimental apparatus and data
management to accompany all this is transfiguring
scientific research.

What wasn't mentioned earlier -- and I would
also direct to your attention -- is a study that was done
under the auspices of the American Council of Learned
Societies with funding from the Andrew Mellon Foundation
and which was chaired by John Unsworth, and which has put
its final draft report now up on the net.

This Commission -- and I should, in the
interest of disclosure, mention that I did serve as an
advisor to it -- took a pretty broad look at what's going
on in the humanities and the social sciences in terms of
information technology.

And I think that the picture that the
Commission came away with, is that these kinds of
scholarly practices are starting to make a major
difference there as well.

    They are not as uniformly accepted. They are
not as uniformly distributed. But indeed if you want to
identify marquee projects -- not necessarily marquee
projects in the sense of having a marquee price tag that
would impress the National Science Foundation, but in
terms of the kind of scholarly creativity that they
demonstrate -- you can find good examples of this,
startling, impressive examples of this in the humanities.

    And indeed we have several people here who have
been pioneers in developing exactly those kinds of
examples, modeling cities, buildings, societies, analyzing
cultural objects using information technology tools.
There's a tremendous amount that's going on there.

    So I think that one of the many messages of
that report is that we should not think of these changes
of scholarly practice as limited to the sciences and
engineering. Indeed they are happening in medical area.
They are happening in humanities. They are happening in
the social sciences. You can even find fascinating bits
of them starting to turn up in other places, the
performing arts for example, where they are starting to
have some impact. I would not overlook also the fact
that, even in the humanities, even in the social sciences,
there's new floods of data turning up.

This is not just about dealing with old material in new ways. It's not about, we can do a spectrographic analysis of the Mona Lisa and now understand how that painting happened in ways that we couldn't before, or try and digitally restore it to the way it probably looked when it was painted.

Just think about things like this. If you wanted to do a biography of Alexander Hamilton's years in office, this is a pretty manageable thing. It's not insane to think of a scholar actually over the course of a career familiarizing him or herself with most, if not all, of the source documents, assuming they can get into the right archives and see them. The scale is a human scale.

When you start thinking about someone who wants to write a history of Bob Rubin's role in the Clinton administration -- to take kind of a parallel thing -- the volume of evidence is unthinkable.

The notion that some human being can actually look at all of this becomes more and more tenuous. We start thinking about having to have automated tools in order to simply deal with the flood of evidence. This by the way I'll just note is not limited to people trying to understand history in academic settings.

I would invite you to have a look at the way
that text-mining technologies and database technologies and information-retrieval technologies have found an enormously profitable niche in things like large-scale inter-corporate litigation.

When something like Enron decides to have a meltdown, and you go in and take possession of all of the computer files of an organization that size, this is a very formidable problem to sift through.

And we're starting to see technologies show up there, which I think are going to show up increasingly in many of the social sciences, as the flood of evidence -- those 17 tractor trailers full of stuff -- start showing up in digital form.

Okay. So those are a few of the things that are starting to change the way scholarship is practiced. I'll just mention two other things that are also having an effect in here. It's not just about what scholars do. It's also related to how they interact with their own scholarly record.

We're starting to see labs, research groups, even individuals wanting to amass significant personal literature collections or to do so at a lab level, to link that up to public databases, to private database, to do computation on it. All of a sudden the literature is permeating
into the laboratory. It's permeating into all of the places where scholars work in ways that it couldn't before, and really becoming much more of a sort of an interactive loop, I think.

It's no longer sort of, read the literature, go research, write, go back to step one. The whole matter has become much more iterative. And there's some complicated things that are showing up as a consequence of this.

Here's one good example. I think just about everybody here knows all about the shift that has taken place in libraries from purchase to site license, particularly for the journal literature, over the past two decades.

Now, what this typically does is basically sets up arrangements where the library licenses and pays for some community to have access. And that community is typically defined by language that says something like, the faculty, students and staff of The University of Texas and -- parenthetically since it's a public institution -- other people who are incidentally physically present on their premises.

You can probably negotiate something like that in. Well, okay, that's great. There's been a technical challenge which we are mostly finished fighting our way
through to make sure that members of the licensed
institution have access to this content, wherever they may
be, whether they're physically on campus, whether they're
at home, whether they're out at a research facility,
whether on sabbatical at some other university.

Well, now we have the emergence of what's being
called in some circles virtual organizations as part of
the whole move to e-Science, cyberinfrastructure,
e-Research.

Virtual organizations are basically nothing
more at their crudest level than a group of scholars that
come together to work on a specific problem, to interact
with a specific set of data and instrumentation, that may
have been drawn from an array of different institutions --
sort of standing institutions -- scattered all over the
globe.

So you may have a team of particle physicists
or astronomers or whatever forming a virtual organization.
Yet their access to the literature is tied up with their
source affiliations. So all of a sudden you have this
conflict between building collaboration environments for
these folks and being able to uniformly import the
literature you want as part of that collaboration
environment.

These are some of the kinds of complexities
that are starting to show up -- these needs to compute,
these needs to support virtual organizations -- as we look
at our ongoing relationship between the scholar and the
literature of his or her field.

We're also seeing the production of scholarly
works that go far, far beyond traditional articles,
traditional printed monographs -- that represent
databases, that represent simulations, that represent
complex mixtures of software and data and analysis.

Right now I think that we're in an environment
in most disciplines where people do both. Very few
scholars have completely forsaken the sort of traditional
article or traditional monograph, depending on their
discipline, for these new kinds of vehicles.

But these new kinds of vehicles are becoming
increasingly important. Even in areas where things like
the journal article are still absolutely essential stock
in trade for scholarly communication, those articles are
being supplemented and interlinked in many, many fields by
complex data sets or databases, often managed on a
community basis.

Often that data is augmented by complex or
specialized software that's necessary in order to
understand, manipulate and work with the data. So even in
fields where that traditional journal article remains
paramount, there is an ever-growing amount of supplementary material that's equally important.

And there are, by the way, fascinating things going on these areas. I don't have time in this conversation to do more than just sort of sample a couple of trends. But I think it is worth it at least sampling into a couple of trends.

There's always been a notion in science in particular about reproducibility. The idea is if you write up the results of an experiment, someone else who is sufficiently well motivated and has the appropriate laboratory equipment, should be able to reproduce it.

And in fact there's been a large gap between theory and practices in many cases, because the article summarizing the results often is very light in research methods and things like that. There's a lot of interest in making reproducibility easier.

And part of the notion there is, can you come up with more raw data that can be reanalyzed as part of a form of reproducibility. So you make the raw data available under sort of norms of scholarship in a field, so that someone else can work with it.

Can you even somehow structure not just the data, but things like the experimental setup in some representation? There's very interesting research work
going on in this. There are people who are trying to
codify theorem proving and things like that, so that you
can basically import an article and rerun the analysis,
rerun the steps of the proof.

I think rerunning the analysis is probably more
plausible in terms of having real impact in the near term
than rerunning the proof. But there clearly is this kind
of thinking about reproducibility. In the humanities, of
course, things get even more interesting, because in the
humanities it's less about reproducibility than citing
sources.

But you could cite some really obscure sources
in the humanities, sources that would require trips across
the ocean, special pleading with curators, to get into see
things. Maybe you could get it; maybe you couldn't. Now,
it starts to become possible to make the evidence -- or at
least digital versions of the evidence -- an integral part
of your piece.

We may increasingly see norms for archival-
based research calling for digitizing the relevant pieces
of the archive if they're not digitized and connecting
them up to the article in some ways.

Finally in this area -- just to mention one
more trend that I think is of critical importance -- we
are seeing in various disciplines a sort of a revisitation
of the norms about data sharing, the expectations about when you share data and why and who can take credit for it and under what circumstances.

And obviously these are very different conversations with very different outcomes from discipline to discipline. But I think that by and large -- and institutional review boards notwithstanding in certain areas -- the tendency is towards greater sharing.

I'd also point out that the funding agencies that underwrite a good deal of the research that goes on -- and I don't just mean government agencies here, but also private agencies, private foundations -- are starting to take a much stronger interest in the data component of research that they fund, and how that data is going to be shared.

I've been intrigued, for example, to see that certain private foundations that fund work in the management and cure of specific diseases are now requiring researchers that they fund to share data with other researchers in advance of publication.

They're basically saying that our primary goal is to accelerate the progress of research in this area, not to accelerate the prestige of the researchers that we're funding. And I think that those kinds of pressures are going to continue to put focus on these questions of
data sharing.

So data, I think, is becoming an increasingly critical part of the scholarly record. Now, how does this connect up to libraries? Well, let's just point at a couple of specific developments. First and most obviously the question of data, which was mentioned earlier -- who's going to take responsibility for it?

Where does that fit institutionally, given that data is of growing importance as evidence, as documentation and as direct scholarly communication? Who's going to manage it? I think that more and more it's being suggested that libraries are going to carry at least a very significant part of this burden.

And it's worth understanding at least a couple of sub-pieces of this. Jim talked about some of the cyberinfrastructure vision of data management. And I'd urge you to read the relevant documents from the National Science Foundation and elsewhere.

But they're being very cagey about how much funding they're going to put into this and about how much investment we're going to see. Really this issue is not just for the NSF. It's for all of the federal funding agencies.

How much investment we're going to see in national-scale data management strategies. We have some
tradition of this in place in areas around the life sciences already, because of the ongoing investment of the National Institutes of Health, in the National Center for Biotechnology Information, within the National Library of Medicine.

That served as sort of a de facto national focal point for the management of a good deal of genomic and related information. We have some similar kinds of things in planetary data from NASA. NSF in fact -- although it doesn't always like to admit it -- does have a considerable history of funding a certain collection of community information resources in the fields.

It funds things like the protein databank. But the need is enormously larger than what's in place. And an open question right now is, how much of that is going to be done on a national disciplinary basis. How much of it is going to be done on an institutional basis?

How much of it is going to be done through various kinds of intermediate or alternative approaches, for example consortia of universities, things perhaps modeled after something like ICPSR in the social sciences. We have a few examples of that.

It's clear there are some beneficial economies of scale in terms of the ability to afford to construct specialized access and management tools in individual
disciplines. For those who would like the government to undertake a lot more of this on a central basis, it's interesting to notice that we are seeing models of that elsewhere.

The UK has done that for a long time through its data centers, which it set up in the mid-'90s in various disciplinary areas -- things like the arts and humanities data center, the environmental sciences data center.

I just recently saw some announcements of an EU-wide attempt to build up a network of repositories to deal with some of this. I'll just make a couple of final points about data and about the role of data here and the role of libraries in dealing with this.

First one -- and this is an observation which I think I also owe to Paul Courant. If you look across our research universities, across the range of disciplines represented there, it's important to remember that a very significant number of faculty really don't get grants.

Or at least they don't get big grants. They might get grants that buy out their teaching for a semester or two to work on a book or to go abroad to lecture. But they're not big sort of systems grants to sequence a species genome or something like that, into which you can build data curation and people, and out of
which you can take overhead to move to the library.

A lot of that research, as far as I can tell, falls to the university, because it's the university's faculty, and the university underwrites it, and they are really the only agency there to do it. It's also worth being aware, in the broader university context, that I think funding agencies are becoming increasingly conscious of the value that research data represents.

They are starting to worry about -- as good funders, as responsible funders, as responsible allocators of public money or foundation money -- how do we ensure that this data gets appropriately preserved and shared. Both in the short-term, how do we ensure that it's being backed up and secured well, so that if there is a natural disaster we don't lose 15 years of scientific work. And in the longer term, how do we assure it gets archived.

You're starting to see funding agencies require data management and dissemination and preservation plans as part of the grants. And I believe you will increasingly see this considered as a factor in deciding who does and who doesn't get grants in an increasingly competitive environment.

And the story that a lot of these researchers are going to tell, with the encouragement of their department chairs, their deans, the vice president of
research. They say, Well, if we have to make a case for
this, we'll say the university and specifically the
library will do this. They'll take care of this on an
ongoing basis when we're done. We'll just deposit it
there.

So there's liable to be not just a need, but
perhaps also a growing mandate to deal with this. Okay.
A couple of other comments in this area before I try and
bring the threads together.

I think there's a real question of credibility
for libraries, particularly when we think about scientific
data. If you look at the ACLS report that I discussed
earlier, the notion the libraries represent the
laboratories of humanities, that they are an essential and
central part of cyberinfrastructure for a future vision of
the humanities, is widely accepted.

Great respect for libraries within the
humanities. When you look in the sciences things become
much more ambivalent. Certainly libraries still continue
to expend massive segments of their budget licensing
access to scientific journals and making them available.

Although I think that despite the best efforts
of libraries, working scientists in our universities are
still remarkably uninformed about just how big that
investment is. Graduate students are even more clueless.
They find them through Google. They think they're all free. Life is wonderful.

Beyond this sort of writing checks for the scientific literature, many of our research libraries have not had very deep interaction with a lot of their scientists over the last decades.

There is a real potential challenge when we start thinking about an enormous role in data management and data curation and data stewardship in the scientific areas in establishing the kind of relationships and the kind of credibility that will be necessary to bring that off.

And I think it's important really to sort of address that very clearly up front, because I think it ties into some of the questions I raised at the very beginning of my talk about the futures of libraries.

I think one future that some libraries may elect to take is to very much back away from the demands of particularly scientific data, and to build largely upon the alliance with the humanities in dealing with data kinds of things, to really continue to restrict their engagement with the sciences in licensing literature.

I think that we will see alternative methods for supporting scientists who are at institutions, who are in that position. I think that's a very strategic
question with a very high price tag though that libraries need to consider.

It's not just data. Another piece of the relationship between libraries and scholarship deals with dissemination, particularly of these things that don't fit into the traditional paradigms of monographs, of journal articles.

I think that -- and I'm sure there are people here who will disagree with me -- university presses particularly have not by and large been aggressive in moving into the digital area.

They have, largely when they have translated things so they're really using the digital environment for distribution rather than getting into the support of authoring of things that are intrinsically digital -- very much by the way, the same pathway that journal publishers followed when they moved to the digital world.

There are a lot of reasons for that. Some of the reasons have to do with a lack of availability of venture capital of various kinds, either grants or funding. Some of it has to do with other priorities. But leave that as it may.

What we've seen as a partial response to that is many libraries now getting involved often in deep ways with individual departments in things that look a lot like
reinventing publishing in the digital world. That's going to be a very crucial role, it looks to me, for libraries. And it's a role that also implies preservation in ways that are very different than the kind of preservation that we had in the print world. In the print world publishers were not responsible for preservation. Publishers disseminated.

Libraries acquired and preserved. In this new digital world you may have a phenomenon that's very much you as a research library collaborating in the construction and dissemination of this thing. You bear a moral responsibility for either preserving it or deciding that it's not worth preserving at some point -- very, very different situation -- with much more pressure perhaps on individual libraries making choices and the entire scholarly community living by the results of those choices.

So let me try and tie this together now into a few looks at the roles of libraries and a couple of other things that are liable to change. The first thing that I think needs to be said is, the massive digitization programs that are under way, among the Google Six notwithstanding, change in research libraries is going to be somewhat incremental.

They are stewards of enormous physical
collections. And those enormous physical collections aren't going to go away. Some of them may move into low-access storage. Some of them may be merged in various ways with those of other institutions to reduce the amount of redundancy in the physical holdings.

Massive digitization will open up access to things in ways that were just unimaginable only a few years ago. But change in here is going to be somewhat gradual. It's just the nature of the beast. We're not going to get rid of a lot of the old really instantaneously.

Indeed perhaps the biggest challenge here is that for every voice advocating and demanding investment in something new, there's still one advocating and demanding continued investment in something that you're already doing.

There are very few things I think today that research libraries do that don't have a constituency within their institutions. So there's a really large reallocation challenge here. It's not a question of, we just stop A and start B.

Let's talk just a tiny bit about the sort of four functions that are typically attributed to a research library: collecting, organizing, providing access and archiving.
Clearly in the organization area things are changing. Things are changing in two different ways that push us in opposite directions. At the same time we are digitizing and we are using in scholarship a tremendous amount more unique material. Images for example have become much more of a first-class object in a much broader array of teaching and scholarship over the last, let's say, 20 years than they had in the past.

Most of this is relatively unique material right now in the sense that we haven't yet built up sort of canons and replicated it in various ways. Although I think you can argue that systems like ARTstor are making the first tiny steps towards that.

These are very hard to deal with because we really don't know how to teach machines to describe and organize images very well yet. So we face very challenging organizational problems in this area.

When we look at collecting it's clear that the sort of traditional published materials -- while they will continue to be important -- are likely to be a smaller part of the collecting universe. Let me just say a quick word or two about those traditional published materials.

You can divide those at least in a sort of a vague way -- and just bear with me; I'm going to be vague here, but I think the point is real -- into materials
where the primary audience is the scholarly community, and materials where the primary audience is not the scholarly community, but they are important to the scholarly community, as evidence that can drive scholarly inquiry and analysis.

Now, the materials inside the scholarly community, we are working out pretty effectively acquisitions and archiving and things like that. Yes. There's still more work to do. But we can point to a lot of progress in the area over the past decade -- progress both with commercial and non-commercial players in that field.

One of the reasons we can do that is because there's a common set of values and interests around that scholarly material. And that includes the values around access, archiving, preservation, maintenance of integrity.

When you look at that vast assortment of material which is not targeted for higher ed, but which higher ed is interested in as evidence -- popular film, television, radio, newspapers, popular books and magazines, all of the tracings of popular culture and discourse -- we have an enormous problem.

The commitment to ensure archiving scholarly access and use of that is not there in our society.

Indeed one of the great challenges that the research
library community faces, I believe, is making the case to the society that we've got to clear a path to be able to do that.

And that's implicit in these discussions around orphan works. It's implicit in many other policy discussions. But that is a central policy problem. If you want to see just how bad that can get, I invite you to have a talk with your film and media studies folks.

They're probably one of the most severely challenged academic fields right now because of the difficulties they face in getting access to the source material they need and being able to weave it into their work.

Beyond the published literature, clearly we are going to see libraries taking more and more stewardship responsibility for all of these other parts of scholarly communication and evidence that I've been describing -- the data sets, the new genre works.

These will have very different financial properties. They will have the property, for example, that in many cases they are very inexpensive to acquire, unlike published materials. But they are very expensive to maintain, to provide continuing access to, to preserve.

The whole economics of how you allocate resource between acquisition and preservation or ongoing
access is liable to shift in this world. And I think we can already see evidence of it. I cannot resist a very quick anecdote here.

In some discussions over the past few years I have had conversations where I've talked with people about: Well, what happens to the next generation of special collections; that today, you know, you can often find a donor that will help you acquire, for example, a set of personal papers that come onto the market; and that usually the primary issue there is getting the money to acquire it. Sometimes there's a secondary issue about finding some money to organize it or to stabilize some of it. But the big problem is usually the acquisition.

Tomorrow you may be able to get many people to say, Sure, take my disks. The problem will be how to handle it going forward. And what you may wind up doing is seeking donors who will essentially endow a collection, that somebody is being kind enough to give you. But you need an endowment to keep it alive.

I actually saw a company -- and I have to thank Neal Beagrie from the UK for the pointer to this last week called Arkhold. They are advertising a service to the consumer now, where you can purchase an endowed website.

So you can pass it on to your children, you know, those precious memories and things like that. And
they've worked out the interest rate and the cash flow and things like that. You'll be interested to know that they're figuring at a 4 percent return the price tag on an eternal website is $2,500 a gigabyte. Fascinating.

So this kind of issue is starting to move into the consumer consciousness as well. But I think that there's a powerful message there. Okay. So those are a couple of the points about how these functions may rebalance.

I want to say one word, too, about providing access and preservation. And that's that I think that there's a growing imperative -- and perhaps this is a particularly appropriate thing to say on September 11 -- to reconsider a bit what we mean by good stewardship.

When we think in our museums, our archives, to some extent our libraries, of unique works and unique collections, I'm not sure that good stewardship is just about better theft protection and environmental control anymore.

I think we need to be honest about how far we have come with digital capture technologies, digital imagining of two and three-dimensional objects. We're at the point now where we can produce, if we want to, good enough replicas for most scholarly purposes.

I would not say all. I'm not sure we'll ever
be able to say all, at least not until we get to the real
Star Trek kind of technology, but a very, very large
percentage. Now, this is starting to emerge, for example,
in the museum community as a serious question.

If you are a museum holding a collection of
public domain works, what really constitutes good
stewardship here? I think you can make a compelling
argument that one answer to that is you make high quality,
digital replicas of this material available very broadly.

That is one way to ensure, not just access, but
survival of that material in the face of all kinds of
inevitable catastrophes. I think that this is also
starting to come into the mix.

I was very struck -- when I was at the
University of Michigan early this year they held a
symposium about the implications of mass digitization.
Mary Sue Coleman got up, and she gave a talk which in part
I believe reprised some of the remarks she made to the
American Association of Publishers, where she pointed out
that one of the results of this digitizing program was
that for the first time they had a real disaster recovery
and business continuity alternative starting to emerge for
their library collections.

I think we need to recognize that these kinds
of issues are on the table. So those are a few thoughts
about how the mix of these long-standing functions, which I don't expect to change -- organizing, collecting, providing access and preserving -- are likely to be rebalanced and reinterpreted.

I think it's important not to lose sight of course of the points that Jim made about library as place, about library as a place that is sort of a social interaction crossroads across the many disciplines represented at our campus.

These elements are coming on as well. But I've chosen in my remarks here to focus more on the collection side. I would just underscore that the question really in my mind shouldn't be just, What's the future of the research library in the 21st century university.

It should be, In the 21st century research university what are our collective strategies for organizing, collecting, providing access and preserving the evidence, the output, the communication of scholarship.

That's really, I believe, the way to frame the question. I think as soon as we do, it points us at the need for a coherent, systematic, institutional strategy that encompasses not just our libraries, but all of the cultural and scholarly memory activities that are based on our campuses. Thanks.
And I would be pleased of some comments.

VOICE: [inaudible]

CLIFFORD LYNCH: Yes. Let me repeat that, because I see a few people trying to hear. Perhaps that's something we should keep in mind as we have our discussions going forward as well, is talk loud. The issue is -- I talked a lot about things on an institutional level.

What about libraries collectively? How do they come into play here? It seems to me that we are moving into a world where libraries are going to need to act more collectively. Many people over the years have commented on this sort of strange mixture of cooperation and competition that characterizes relationships between our major universities.

They, to a very great extent, share common values and objectives. But they also in some cases compete vigorously against each other for talent and for prestige in various areas. I think that libraries, you know, have always picked up on some of that as well.

They collaborate. But they also are protective of their resources and of their relative positions. The quality of the libraries reflects on the quality of the university as a whole in an important way. It seems to me that in this digital world what we're going to find
inevitably is that many kinds of collections are going to 
become more uniformly accessible.

The notion that you cannot attract someone 
because of the depths of your special collections may get 
shakier, because you've got those special collections 
digitized, and in fact they're being used by scholars 
around the world.

I think that that's going to be a tough thing 
to get over. But we're just going to have to move past 
it. I think that really the requirements of scholarship 
demand it in the same way that I think that some of the 
discussion earlier is pointing to rethinking a number of 
things about higher education and how we do it.

It's clear to me that for many kinds of data we 
will be in much better shape if we work together than if 
we work individually. It's just too hard to work at this 
individually. It's too expensive. We haven't got the 

scale.

It seems to me that we end up in a world where 
inevitably we're going to have to learn to trust each 
other as institutions in more critical ways than we do 
today perhaps, and in less geographical ways.

I mean right now you see a history of 
cooperation that's often rooted in geography -- regional 
storage centers, reciprocal borrowing privileges, things
like that -- that mostly make sense on a geographic basis.

We're going to, I suspect, see alliances show up in some cases on strength of academic programs in specific areas, rather than geographic proximity.

You know a few schools with premier interests and capabilities and discipline acts coming together to deal with some of the data management issues around it.

So it that's one thing you'll see. The other thing -- and I think you can already point to the work that ARL and other library organizations, but especially ARL have done in speaking out on some of the public policy issues that are crucial to libraries.

I think that collectively research libraries are going to some extent live or die by their ability to make the case on a public policy basis around some of these intellectual property issues. They've got to continue to work collectively through organizations like ARL to do that.

VOICE: [inaudible]

CLIFFORD LYNCH: That's a very complicated question that probably to some extent goes way beyond my competence to answer very well. Maybe I'll just say one or two things that come to mind as little bits of this.

It seems that in the sciences at least we are certainly seeing lots of work going on to operate on a
coordinated, international basis, both in the developed world and, where relevant, into the developing world. You certainly see that in many of the data management strategies in environmental studies, meteorology, earth sciences.

They're federating the national virtual observatories into an international system. Certainly the interchange agreements in Genomex have been in place for a long time. So within the sciences there's always been some attempt -- not always but certainly in the last 40 years there's been some attempt to think systematically.

It tends, I think, to be more of a federation kind of a thing than a truly multinational that people contribute to. There probably needs to be more of that kind of conversation in the humanities than I'm aware is taking place, although I may have simply exposed my massive ignorance in that comment.

But I'm not aware of the same kinds of organizations as bring together scientific data management groups across most of the humanities. It would seem that there would be interesting prospects for libraries in other countries, both developing and developed, to contribute in very different ways to area studies and language studies and things like that.

If you've got, you know, the national library
of Sweden collecting everything in Swedish that was ever done in a very systematic way, and they can figure out how to make it available, it's not clear how much of that you really need to replicate, if they have got expertise to bring to bear on that that we may never have here in the States.

At the same time there's a lot of tricky issues about not wanting to leave the worldwide cultural records at the mercy of individual, political and social entities. I don't know how to square some of those. I think there are a lot of good questions to be asked there.

I do want to circle back and just make one comment on one of the sort of framing statements you made. I think that while collections may become somewhat more homogeneous, that the service portfolio and the expertise portfolio that is held by individual research libraries -- and here I am talking mostly in the States -- may become a more distinguishing feature.

It's not just the collections -- be they at your home institution or anywhere else among the collective research enterprise -- rather it's how capably you can help scholars to navigate and work with that entire constellation of resources.

VOICE: Could you talk a little bit about the issues relating to the collection and archiving of net
content as a [indiscernible], whether it be correspondence
or email, publications of net documents, selection of
websites. We all agree that these are pretty ephemeral
and pretty dynamic --

Could you talk about that in the context of the
role of research libraries?

CLIFFORD LYNCH: Yes. Happy to. The question
was about the archiving of web content and the role of
research libraries doing it. I think that first off this
is going to be an essential part of understanding many,
many aspects of our culture, society and activities going
forward.

And we absolutely have to collect and preserve
it -- just a no-brainer. I think that we want to
replicate this record fairly broadly, because it is
subject to lots of kinds of attack.

I think at the same time there is a good deal
of leverage in research libraries coming together to use a
relatively limited -- and I don't want to say, one -- but
a relatively limited set of collections tools and
capabilities, so that everybody's not out their trolling
the entire web every week, but rather we siphon it up.

I'm not a big fan of technical monocultures of
any kind. Maybe we have two or three distinct breeds of
siphoning tool. Then we do those passes and interchange
the results among our repositories and libraries. So I think that's really important.

Obviously, sort of adding on to the baseline of that, there are more detailed, more refined and more in-depth studies that you want to do to support specific scholarly disciplines or research groups. At one level I would say the really good news about dealing with the web and archiving the sort of surface web is that it's pretty cheap, and it's pretty automated.

If you look at the operating budget of the Internet Archive relative to the operating budget of the major research libraries of the United States, it's a pretty good trade, considering the significance of that content.

I think that we get into much more problematic areas, when we start talking about deep web material, which we really haven't dealt with very much -- databases and things. We really, I think, underestimate how much stuff is in there in the databases that we used to pick up in ephemera of various kinds that no longer is produced in ephemera.

So I think a focused look at that is necessary. I also think that dealing with things that are not exactly part of the public web, but that represent institutional records that are on a web inside an
institution or even personal records, is going to become critically important.

I mean, going back to the 17 trailers there, I'm aware of very few efforts to do that in current corporations where everything they do is increasingly represented in a digital way.

VOICE: I hope that libraries will get involved in -- catalog, and I hope we'll talk more about that tomorrow.

CLIFFORD LYNCH: We're just about out of time. I'll take maybe one more question or comment, if we have one. It looks like people are ready for a break. Thanks again.

DON CARLETON: Well, I'd like to thank James Duderstadt and Clifford Lynch, both our speakers today for fine presentations. Thank you.

(Whereupon the symposium was recessed, to reconvene on September 12, 2006.)