DOCUMENT RESUME

ED 362 358 RC 019 316

AUTHOR	Smith, Steve
TITLE	Project Jukebox: "We Are Digitizing Our Oral History Collection...and We're Including a Database."
PUB DATE	Oct 91
NOTE	10p.; In: The Chugach Conference: Finding Our Way in the Communication Age (3rd, Anchorage, AK, October 3-5, 1992); see RC 019 315.
PUB TYPE	Reports - Descriptive (141)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	*Access to Information; Alaska Natives; *Archives; *Audiotape Recordings; Computer Software; Databases; Higher Education; Information Retrieval; *Information Technology; Legal Responsibility; Libraries; *Library Collections; *Oral History; Preservation
IDENTIFIERS	Alaska; *Digitizing; University of Alaska Fairbanks

ABSTRACT

Project Jukebox is an enterprise of the Rasmuson Library at the University of Alaska. The project aims to digitize 6,000 hours of oral history. The special oral collection of the history of Alaska includes audiotapes of major and minor politicians in the state, average citizens, and Native leaders. The audiotapes created problems for guaranteeing preservation and ease of access to the information. Apple Computer provided the project's hardware and some software. One piece of software compresses the digitized sound to save storage space. The digitized audiotapes are stored on magneto-optical-erasable disk; they will eventually be available in CD-ROM format. Users of the program can search for information by such means as subject, keyword, or file number. On the screen, users find transcripts of the information with options to listen to the audio, view pictures, print a transcript, or view the release agreement. Converting information to new technologies entails problems that concern deciding what materials will be converted and who will have access to them. Questions by conference participants addressed the costs involved in the project, enhancing the recordings, and copyright issues. (KS)

Project Jukebox:

'We are digitizing our oral history collection . . . and we're including a database'

Steve Smith
is project manager and coordinates extended campus services at the University of Alaska Fairbanks Rasmuson Library. He has been responsible for design, installation and maintenance of network access to CD-ROM databases. He has an M.A. in communications from the University of Hawaii Manoa and is completing a Ph.D. in information science from Nova University.

I really have what I think is the best job in the world — I work at the Rasmuson Library at the University of Alaska Fairbanks — because my title is "project manager" and basically this means that I'm kind of the in-house consultant on technology. So every time new toys come in, I'm the guy that gets to play with them, and I'm involved with all the new projects there. We're making a real concerted effort at Rasmuson Library to open up access to things, and part of what I'm going to show you today, Project Jukebox, is one of the ways we're doing that.

You're really going to see two things that are happening simultaneously, and I think it's echoed in most of the things I've read; and one is we're moving into multi-media — there certainly is a big push on for that — and that means we're getting away from just the text which is kind of an archaic and foreign language. A lot of the programmers I work with, I say, "You guys are working on dead languages." They have to learn all those arcane commands, and all those little abbreviations that nobody else can understand. But we're moving into audio and video and the full spectrum of things. I started out in broadcasting; I got out of that because I thought it was too narrow a thing, and I got into computers and networks and I discover now that all of my broadcast and video and audio background is coming back into play again, because we're really getting into that. The second thing is, with this multi-media playing, we're networking and making access available to all these things.

We have a number of special collections at Rasmuson. One of the collections we have that's a special collection is an oral history collection — about 6,000 hours of tape right now, and it's growing as we continue to collect audiotapes. It also includes some videotapes — that's one of the specialized collections we have.

We have a problem with that, as we have a problem with all of our specialized collections. The problem is two-fold: one is preservation; the second is access to that material. A lot of the recordings we have are just on this type of tape, and any of you who have worked in this know that if these things just sit on the shelf, and you don't do anything to them, they're going to degrade. They're going to lose the quality from that magnetic recording. Most of the stuff we do now that's new is on cassette. We haven't quite made it to the digital-audiotape , but we're moving in that direction. But even those — if this stuff just sits on the shelf it's going to have problems. And with 6,000 hours, it becomes very labor intensive for us to go back and rotate the tapes and run them through a machine on any kind of a regular basis. Not to mention the storage that this starts to take up. And we face this problem of storage, of paper documents — we have an extensive photograph collection, several hundred thousand photographs in our archives — how do we store that stuff, how do we make access to it?

The access problem exists right now. As you walk into our archives, it looks very nice, very traditional, everyone

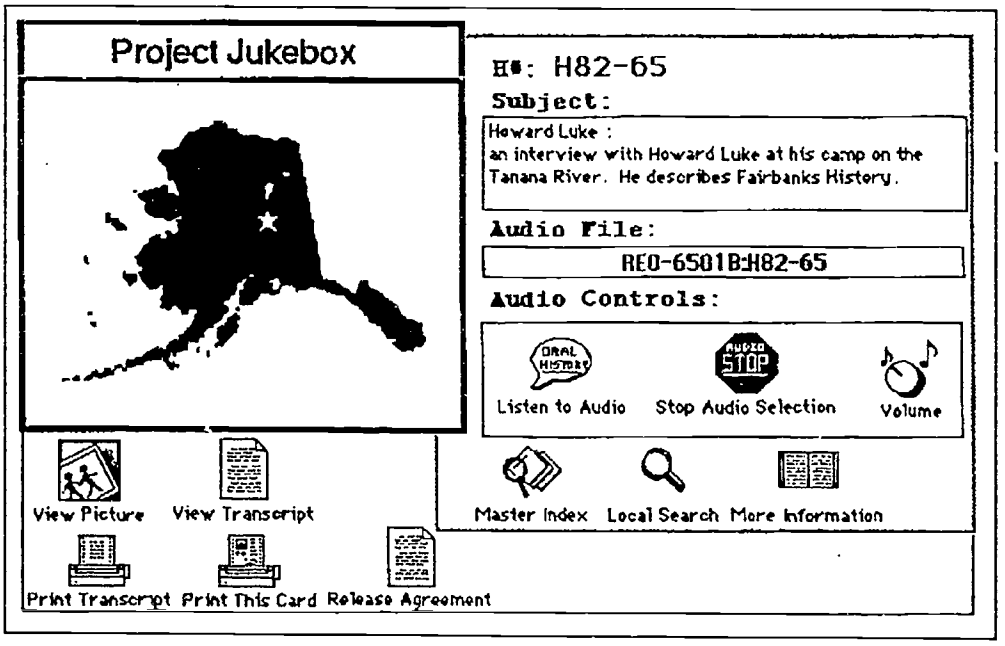
speaks in hushed tones and you sit down and you look through this index — we finally put the index on a computer — but it's hard to really understand what it's saying. And you search through and you might find a recording you want, and maybe if you're lucky a small abstract that tells what's on that recording. Our oral history collection covers every area of Alaska history, the past and the current. We have all the major politicians in the state, and the minor politicians in the state, and the average person in the street, in the cities and in the villages, the Native leaders. Any area you want to go into, we have coverage of that on oral tapes — some of them have been recorded by us; that's what we're doing now — but many of them are recordings that have been made elsewhere, that have been made on radio stations, that are coverage of something, some event — we collect those tapes and archive them. So there's a lot of interesting and important information there — cultural, political, economic, social, you name it.

But anyway, the person comes in, they may find the tape they want, they think, they find several tapes they want, and so they give a little slip of paper to a staff person, and that staff person goes back and finds the tape. And of course we can't let people listen to the original recording. Suppose they hit the wrong button on the tape recorder; they'd erase it. So we have to make a copy of that, and you can't make duplicate copies and just keep them on hand — there's 6,000 hours of recording. That would more than double our storage space, and just the time to do that is too expensive. So the staff member goes and finds the tape and goes and makes a copy, and if there's not too much else going on, you may get the copy the same day; you may not get it for several days. If that person were on vacation, or sick, or the right staff member's not there, you may have to wait longer.

You finally get the recording in your hands — it's a dubbed version — and those of you again that deal with the analog format know that every time you make a copy, you lose some quality. And some of these tapes are recorded on a little $20 cassette recorder, so by the time you make a copy of it, half the time you can barely hear what's going on there. You finally get the tape in your hand — if you're the researcher and trying to look up this stuff — and you go to play the tape, and you have to play through the whole tape to find this stuff, and it may not even be the recording you wanted. So then you go back and you repeat the whole process.

So we thought that — there are several things — one is it deters people from just getting access to this information. People will begin to not even use it. There's a lot of history of past and present here that's just being lost. Secondly, just from a preservation standpoint, how do we preserve this stuff? Well, we started talking about that — we had a real specific need — and were looking at ways to preserve that. It seemed to us that putting it into some kind of electronic, not analog but digital format, and making it accessible with a computer front-end — it starts to work.

Two years ago when I came to this conference, Bill Ahearn was here from IBM, and he was interested in the project, and we were carrying on discussions with IBM and with other folks, Apple Computer among them, and Apple Computer bit first and through Apple Computer and a program they have called "Apple Library of Tomorrow," they provided us with all of the hardware to set up this project.



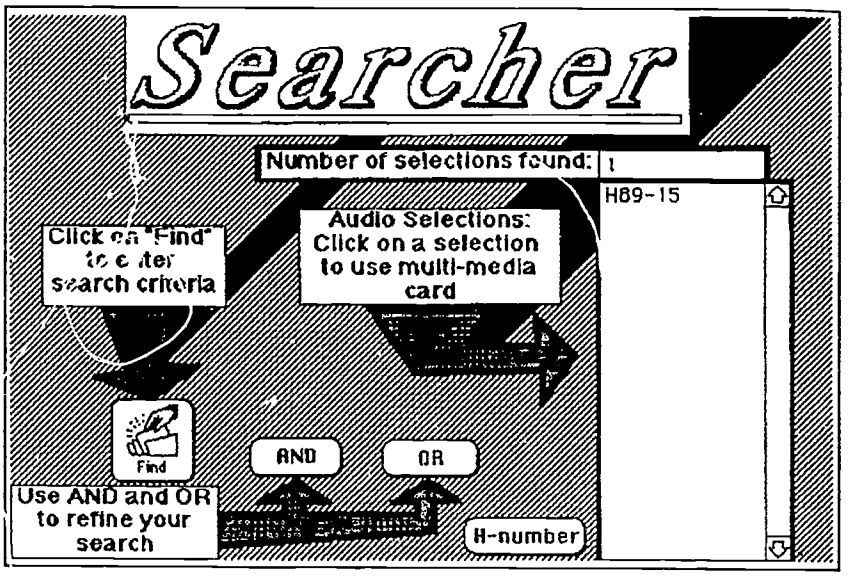
An introductory screen for an interview that has been formatted for the oral history CD-ROM collection.

Basically what we are doing is we are digitizing our oral history collection, and we are putting it into a digital format. Along with that, we're including a database so you can search through and find the tape you want and then play it back immediately.

Apple's interested in it, as IBM was, because it's a multi-media platform. And Apple really is giving us some proprietary software that deals with compression of the digitized sound, because digitizing anything like video-audio takes an enormous amount of storage space. So we're getting some software from them that'll probably be

on the marketplace in a year or two. And they really I think just saw us as a good practical application of this software they're developing. We digitize the audiotapes; right now we store them on what's called the magneto-optical-erasable disk — it's basically an optical disk you can record to, and you can erase it and just use it like any other storage media in that computer.

Our final goal will be to put this right now onto a CD-Rom, and then to put work stations around wherever people may be interested in accessing this information. We're using right now a Macintosh SE as our baseline work station. If you already had one of those you could just acquire the CD and a CD player, and plug it in. We're not married to the CD technology necessarily. Our feeling is once we digitize it, whatever the memory platform becomes in 5 or 10 or 20 years, or 50 years, once it's digital we can move to that platform. Once it's digital, we can make however many copies we want without any loss of quality in the original. Those of you who are interested, I won't go into the technical details now, but if you're interested, I can tell you some of technically what we're doing, and how we're sampling the audio and all of those kinds of things.



The Searcher screen allows the user to find records by specifying keywords.

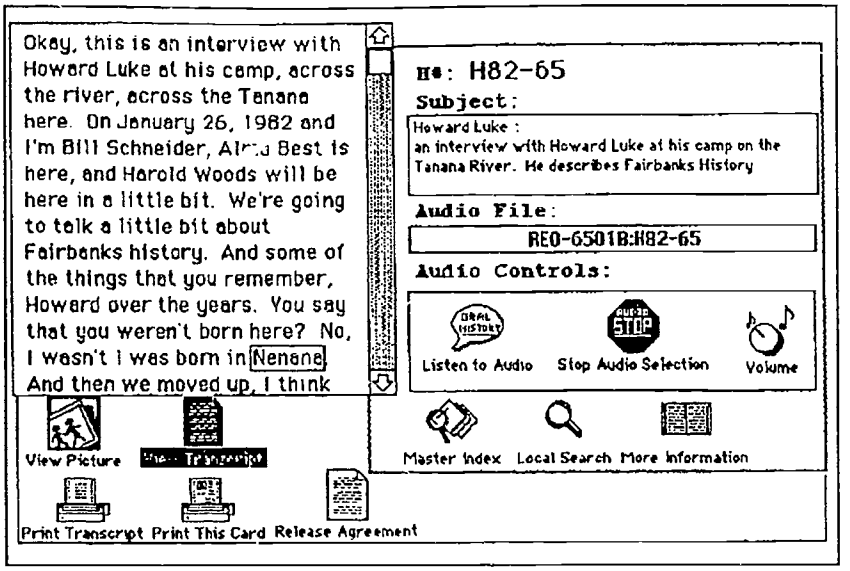
And that's really where we're spending a fair amount of work right now. What I'm going to show you today is a very early prototype of what we have and what we're working on. What we will then be able to do is the access will go directly to the individual who needs to access that information. They don't have to go through an intermediary; they can directly get to the tape and find what's on the tape and listen to it, and even pull a recording off. We can make specialized collections.

One of the things we're working on is a project with the National Park Service — they want to put these around in the parks, for instance in Alaska — that have information on the parks that will cover the range of things. If you're climbing a mountain, you can look at the specific mountain, and see if they've done it before, what routes they've taken. If you're just a casual tourist and want to know a little bit about where the name Denali comes from, you can look that up — you can hear an oral recording. We've also added to this a visual component, because we thought, well, while people are listening to this they want to look at something, they want to get something out of that, and we've added photographs where it's appropriate, when we have those available, of the individuals who are talking, or of the event that was recorded, so we really have a multi-media thing here.

Eventually we hope to move to a platform where we can include video in this as well. One of the reasons we wanted to digitize this and keep as high a quality as possible is that there's a very distinct difference between looking at a typed transcript of the oral recording, and hearing a recording in the environment it was made. So you can hear what was going on there. A lot of the recordings, for instance, were made with folks out in their cabins, in the villages, and in the background you can hear the other sounds of kids playing, with dishes clinking in the sink. That gives you a sense of place and a sense of history that you would lose otherwise.

Let me show you what we've got here. This is a very limited thing — I can't search through a lot of records on this because this is kind of our travelling road show — but I'll show you kind of what we have. This is all being done by basically Bill Schneider, who's director of our oral archives, and Dan Grahek, our programmer. And I kind of dip my hand in and kind of play around with the interface and other things from time to time, and then am the contact with Apple Computer. The project has gone in bits and pieces. Once I go through it, I'll tell you some of the implications that we're running into with funding something like this.

So here we have a searcher, and you say, well, you want to find some things and right now, it says "Nothing's found," up here because I haven't put anything in, but let's say I want to find something... Down here at the bottom it says "H number" — that's just our internal number — if you happen to have a list of those, you could search right through to the recording. But if you don't have that you come over here and you say, "Find," and it says, "What do you want to look for?" I happen to know that one of the three recordings I have in here today is from a guy named Howard Luke, so I'll just put in "Luke" and it says, "Do you want to look for the transcript?" What we're also doing with this is we take the recording and we put in a full typed transcript, so you can follow that along. Because the quality



A search for the word "Nenana" locates it in the Howard Luke interview transcript.

of many of the recordings is fairly low — with a transcript you can see word for word what the person was saying or talking about and, if I can find my pointer, I'll just say, "Look in the abstract," and it's a one-hit.

Now in the final version of this, of course, there'll be all sorts of numbers down here, and you pick the one you wanted. But we'll say we want to pick that one — Dean Gottehrer referred to the problems we have with copyright. One of the problems we have is we had specific agreements with the folks that we got these audio recordings from, and we have specific release agreements, and here it is, and we put it in there. And we've done everything we can to insure that the person will read that. And so there's the release agreement that's here, and of course a person can just bypass all of that, and just click on that "Read the release agreement" without reading it, but what can we do?

Okay, this now shows that this is the number it is. Here's a brief description: "An Interview with Howard Luke at His Camp on the Tanana River." And it describes Fairbanks history, that's some internal stuff that we're using. Over here it places approximately in the state where this recording is coming from. And what we hope to add later on here is a geographic locator, so if you say, "Gee, I'm interested in finding out things about Southeast," you could click on the Southeast and see what kinds of recordings are there, and then combine that with something else like "transportation in Southeast" for instance. So we have all of that down there.

Here's the release agreement again, which you can see — this is all in HyperCard by the way, so I just print this card — we will have a transcript which you can put up there, and there's the transcript of the recording. And over here it says, "Listen to the audio," and that's if we actually want to hear it, and we can adjust the volume. And we have photographs associated with it, and we can select the images. Here's one that's a map and it says, "Here's where we took it from," and we can go back and pick this other image and we click on that and it shows you Howard Luke himself — he's working on a boat. Now one of the interesting things — one of the problems we're wrestling with in terms of copyright — this is a very poor quality photograph: one of the reasons we're doing that is again we have a copyright problem with all of our photographs. We've specifically scanned in our photograph at what we call "reference-quality photograph"; later on, as we do this, we'll have a reference number down here so you can get to the actual photograph. Because one of the concerns we have is if we've scanned that in as a very high quality image someone could take this and hook it up to one of the printers. They could print out that photograph and use it in ways that we don't have a release that it can be used. It becomes a real problem. What we plan to do to archive this stuff is scan all this stuff in at a very high resolution, so at some point when copyright law changes and some of these issues become clarified, we can deal with that higher quality image. But for the work station right now we have a much lower quality image.

And that was Howard Luke. We could search locally in the transcript. I saw that "Nenana" is in there, so I could say "Nenana," and it comes right to "Nenana" and highlights it. One of the things we're getting from Apple Computer is controls — over here where it says, "Listen to the audio," we'll not only be able to listen to it but we'll be able to fast forward. So for instance, you could find this reference to "Nenana" and go right to that point in the tape. That's one of the things — we're working with what's called the "advanced technology group" at Apple Computer, and they're working on the software for us, which is a handy group of folks to have working on your software. So you can jump right to the place that you want to in the tape. For instance, if you had recordings of Jay Hammond, and you want to say, "What did Jay Hammond have to say about the Permanent Fund?" you could pull those pieces out. It makes all sorts of interesting things for people in radio working on documentaries to do research — to find things. The other way we do it, if you really want to be more traditional in your searching, you can go to what we call the "master index," and you can search here.

We borrowed almost everything we have here, and we've got money from different sources. By the way, one of the chunks of money we got to continue this project was from the Center for Information Technology to help us develop all of this, and we really appreciate that, and it's helped us to continue going with this.

You can search through this — this was developed by a guy named Mark Zimmerman, and he works for the CIA; he's their Russian computer expert, and he's the only one they have, so he tells me. But on the side he creates software like this, which is a searcher — so I can search for something. And it says. "What do you want to jump to?" And I happen to know we have a recording in here by Vuka Stepovich — those of you who are familiar with Alaskan history will recognize that name — and here it says, "Vuka Stepovich," and it jumps to the recordings, and there it highlights it, and luckily for me that happens to be the recording we have in there today, and I go over here and I highlight that. This is the way to search through things in a more literary traditional library format. And I highlight the number there, and I say, "jump to the audio card for me," and then that comes up again, and it's going to tell us to read that, and I've read that, and here's Vuka Stepovich, and again you have an audio recording which you can listen to — this one happened to be recorded in Fairbanks as well — and that will come up in a moment.

Highlighting one of the people provides the information this is Vuka R. Stepovich. This is a reference-quality photo.

While you're looking at that, we have again an image, now these are scrolled down — some things we have a great number of images on; some things we have only one image. Here's the interview with Vuka Stepovich. And again, you can go through and highlight and say, "Who was that?" There's Vuka right there, and there's Mrs. Dan Isabelle Egan — who's this down here? Margaret Egan. And you can go through and highlight all of the people... There's one here — Pat Egan. We don't know who that person was. Later when we know what this is we can highlight this and we can click on that and find out exactly where the recording was made, whose cabin this was. So we can jump around here and see all of those things. And if you wanted to, we could print the card. I mean, you could go and get the actual photograph. And you go back here to see the transcript, and you could follow through the transcript because you can't hear it very well. And I can just follow the bouncing finger, and you could follow the recording there.

I want to show you one other interface we have, and that's for the Parks. The National Park Service is very interested in this. We've done an interface for them that's a custom interface, and we're doing a series of recordings with them right now. So we've done this for the Yukon Charlie National Preserve. Right now you'll be able to look at places, subjects, and keyword search — if you're still into keyword searching — or we can look at people. Right now we just have the people interface, and there'd be a number of people. Right now we just have Don Chase, and we have two recordings by him, and I'll just click on this one, and up comes the recording on Don Chase. And there we can search through these various topics that Don Chase talks about. These will be set at information centers in the National Park Service parks. What we do here, instead of a full transcript, is just an outline of the points he talks about. We have a photograph of him; we don't have the audiotape in here for this. You can see here how we'll have the controls for fast forward and rewind, and you can jump right to where you want to go.

One of the problems we find we're facing with this, and it's a larger problem today within the library world, is that the National Park Service is paying for this. So we're paying a lot of attention to that right now because they're the ones paying us. We're doing new recordings for them. The recordings we're going to digitize and put onto them do not include our 6,000 hours of oral history tapes. We got another small chunk of money from the Fairbanks Native Association — we're going to concentrate on some tapes that they're interested in. But what happens is those 6,000 hours of recordings — Vuka Stepovich and Howard Luke —are getting left in the dust. Until we find the funds to pay for that, and we may not find those funds, people are going to come and they're going to use this, and they're going to look this up, but they're not going to be interested in finding Vuka Stepovich or Howard Luke, because that stuff's in an old format, and it's hard to use. And that becomes a real critical issue: Who's making the decisions as we move in to new formats and new and easier ways, in my opinion, to access information? Who makes the decisions on what things we put in that new format, and what things we leave behind?

For instance, one of the things I do is put CD-Rom databases onto local area networks so that people can access those. In fact, we now have dial-up access, so anywhere in the state, potentially, you could dial into this network I've put together in Fairbanks and get access to 12

different databases covering a broad range of topics with about 10 million citations in research. The problem is, we're finding that as students and faculty and researchers are coming in to use that, that is not comprehensive coverage, but it's easy to use, comparatively. It's certainly much easier to use than going through large books with small fine print. And we've actually had students come in to use something, and we've said, "This only has coverage on the CD-Rom for the last year of information." For instance, we have one called "statistical masterfile," but if you want to get coverage for the last 10 years, it's over here, and that may be more help for your research, and they say, "Oh, it's print. Forget it; I'm not even going to look at it." They're going to look just at the stuff that's there, that's in the nice pretty format on the computer. And how do we deal with that? Some of those things are going to drop away.

There's a series of research materials called "The Arctic Bibliography" — 15 volumes, print -- comprehensive social, economic, scientific research on the Arctic. They stopped printing it about 15 years ago — I'm not sure of the exact date — and they've gone to an online version. People use the online version. We've tried several times to get funding to digitize the print version and make that also electronically accessible. We've gone to people like the National Science Foundation and they say, "We're not interested in that; let's deal with this current stuff." Eventually that 15 volumes of information on the Arctic is going to get used less and less. And so the kinds of things we're doing here — I really believe in this kind of stuff — and we're moving ahead with this, and I really see that we're going to get to a point... we're going to have this available first on our computers.

We're probably six months away from having the National Parks Service stuff done; another year or so from actually having a product that's available in our library. But once we do that in Rasmuson, and then put it on the local area network that we've got, and then our next thing — we've already been experimenting with passing this through the university computer network so that you don't have to have a Macintosh computer, it could be an IBM or other flavor of computer — you can access this information on the network. The problem is, when the money comes in, are we going to have access to all of the tapes that are there or are we going to have a select few, and who's making those decisions on the select few? So from my point of view, it's not the technology — this is the fun part, and you play around with this, and talk about the user interface. And that's what we're doing now. We're moving ahead with that and that's going to happen one way or the other. But the real problem comes in with who's going to have access to it, what materials are we going to make accessible to people. That's a real quick overview of our Project Jukebox. I'll be happy to answer any questions anybody's got.

LARRY WIGET: What is the estimated cost for digitizing 6,000 hours worth of tape?

SMITH: For 6,000 hours worth of tape we're going to be probably somewhere around $150,000 to $250,000. A lot of this stuff is just manual things. We have to do transcripts for everything, the full transcript. It's one-to-one right now that we put the tapes on, but with the compression we can save a lot more of that. By the way, this digital storage of archival materials is still a controversy. The national archival organizations do not accept digital storage right now as a standard. We're going to present this at a conference in about two weeks — there's a national conference of oral historians and archivists, and it will probably create quite a controversy because this is not accepted as an appropriate means of preservation.

HOWARD WEAVER: How would those expenses be affected if you weren't retrofitting, essentially. That is if you were to start doing it now, would that reduce that appreciably?

SMITH: A lot of the costs would be reduced. I mean, right now, we've been working on this for over a year — and I only have received three recordings, and they're not even full recordings — because we've been doing a lot of development work and we continue to do that. So the costs will go down some. But it'd still be the basic costs of getting the transcript and paying the staff to sit there and put the recordings in, and quality control. But those should go down. One of the things we've been doing is developing here a kind of a cottage industry that I think somebody in the private sector should be interested in because the National Parks Service has given us two chunks of money to do things for them; the Fairbanks Native Association is interested in it. We've had some discussion with BP about this; I think there are a lot of organizations where you can do this kind of a thing. And what we're really moving to — we started out with specific problems with oral history tapes, audio recordings — but we're really migrating to a point where this is a multi-media product for preservation of records which would include photographic records, film, videotapes, print. It would include the whole spectrum of records.

LEN FRAZIER: Have you worked out a method for enhancing the recordings? Is it possible, for example, to remove background noises?

SMITH: Well, we can. Yes. We've been working, and the guys at Apple have given us some software, and another company... it's marvelous when you work with someone like Apple, because they talked to a smaller company in Silicon Valley, and this other company says, "Oh, they're working with Apple; here, have some software; here, have some support." We got a nice product from a place called

DigiDesign, and we can actually go in, and we can filter out certain frequencies, background noises at a certain frequency, and we can just filter that out, cut it all out, cut out the highs and the lows. The question it gets into though for us from an archival standpoint is do we want to change the original recording, do we want to change the way that that was done? If this was recorded in a noisy auditorium, do we want to filter out that fact or do we want to leave that in there? We can certainly do that. But what we're dealing with right now is we're going to make an archival copy of the tapes that's the highest quality we can go at, and we're going to store that probably not optically, but probably on one of the new digital tape backups.

Those of you who do computer stuff — 8mm tape and digital-audiotape , you can store gigabytes of information on that stuff. We'll probably do that and reduce our shelves and shelves of tapes down to one shelf with all of our 6,000 hours of recordings. And that will record everything — problems in the recording, all the noise and everything. We'll then use with the compression that we're using from Apple Computer — the stuff they're giving us is virtually lossless compression; it means you don't lose any of the quality. But we'll probably take that, probably cut out some of the other information for what we would call our reference-quality recording, change it to something that would be a high-fidelity recording, to something that might be low-quality voice-grade recording. So there we may deal with changing the original recording a little bit.

But again it becomes a problem. As you mentioned, there's the capability to fix photographs, and we have a lot of old archival photographs. I had a guy from Kodak come see me the other day who was showing me their new electronic imaging. He'd take an archival photograph, it had scratches through it, and was dated 1898. They digitized it and they cleaned that all up. It looks like it was a photograph that was taken yesterday except for the clothing that people were wearing. We may want to start doing that. What then do we believe is the original? We get into some real interesting ethical questions here with the preservation of materials.

MICHELLE CORDER: I'm curious about the genesis of this wonderful project. Apparently there wasn't actually a heavy user demand to use these fragile materials as much as it was an administrative decision to make an attempt to preserve for archival purposes your collection.

SMITH: We had, I don't know what the numbers are because I'm not directly involved with the oral history program day-to-day, Bill Schneider, though, the director of the program, felt there was a need — that people were trying to get into the tapes and it was cumbersome basically for us to find the recording somebody wanted and get it into their hands. And then the preservation is a real issue.

CORDER: So it was sort of the archival part of the university rather than the library part.

SMITH: It was the archival part. Will is an anthropologist by training, he works in the library, and he has a real keen interest in rural Alaska and Native Alaskans. He sensed from working with them that there's a whole culture here, an oral tradition, that we have preserved on a number of these tapes. He's gone out and sought out and collected tapes with elders in villages, and a lot of other people from the villages, and he has a vision that eventually there'll be a machine like this in the schools that the kids could go to in their village and find out what their grandfather or great-grandfather had to say — that this stuff will be there, and that the kid could sit down with this and open it up and find out the history of his or her village, or his or her region of the state. The whole constitutional convention is on audiotape. We put in a proposal last year — that didn't make it — to include some of that so you can sit down and hear the actual debates that went on at the constitutional convention that developed the constitution for our state, you can sit down and study that.

DEAN GOTTEHRER: I used to take students to the library every semester in my magazine-article writing class, and one of the places we would stop was the archives because it's a veritable goldmine of material for magazine free-lance writers who are looking to write stories about Alaska. And with stuff on tape you've got to go through the whole process that Steve was talking about. And if you're writing an article, you then have to do your own transcript to try to get that material out so you can use it in an article. It's far easier to use the photos although you have problems trying to locate what it is you're looking for and to have a copy made that you can use. I think that the demand was there. The problem was getting access to the materials, finding what you're looking for and being able to use it.

WEAVER: What will happen to the basic environment when you're done with it? Will it be proprietary? You guys will own it? Or Apple will own it?

SMITH: We'll own the product that we have. We plan to have one basic product. The first kind of thing I showed you is our oral history one, and we'll own that. We haven't really gotten to the point of discussing, say with Apple, what we do if we need to license that out, or if people just want to buy the engine that we have, and then use it. I think this would be a great one for somebody in the private sector to go to Science and Technology grant in the state, to take this on and develop that and then sell that to companies. Other places like the Park Service — they'll own the one that we do for them — we'll have all the program; they'll just have the end product. And we'll make a version where they could add their own tapes if they wanted to. But we're doing all the development work for them — so we're customizing it by client right now.

AUDIENCE MEMBER: What kind of legal or ethical hurdles do you have to get over archiving materials like that? Is it a legal problem or primarily an archiving one?

SMITH: It's both, but primarily a legal one. We're very careful with our archives department. It's called the Alaska Polar Regions. When we acquire collections from people, we'll get just a collection of documents or a collection of photographs or a collection of recordings spelling out — and basically what that release agreement you saw there was — what the rights of use for that is. And then we charge people to come in and use it. Some things you could use for commercial ventures. We'll have television producers come in and they want a bunch of our stills or some of our videotape or film to use in productions and some things aren't releasable for a commercial product; other things we can. There's a restaurant in Fairbanks called the Pump House but they went into our archives collection and made copies of the photographs, and they paid a fee to use the photographs. That's why we're really... Some people are very apprehensive, they're scared to death in fact of archives.

When you go to this kind of digital format, you can hook this up to a high-quality recording machine, you can hook this up to a high-quality printer, and if we have the high-quality images in there, and if we keep the recordings at a high quality, people can pull that out and use those things for whatever they want, and use them out of context. That's a real problem we have. We've made an intellectual decision at this point because we could make it so you could access the tapes anywhere you wanted. And you can do some of that now. We have a fast forward that I showed you that we'll be putting in there to look for a specific word. We made an intellectual commitment that you find the transcript for the tape, you see that, you see the abstract and then you access the whole tape. So that you get it at least in context. Because again, let me go back to the example of Jay Hammond, what he said about the Permanent Fund, you could pull that out of there, or we have recordings of Walter Hickel, you could pull that out of there and use it out of context very easily with this. That's a real danger.

JOHN McKAY: Do you see that as the role of the university? Sort of to monitor whether people might take university materials, reference materials, library materials and use them like that?

SMITH: No, we don't want to be a policeman for that. We do try to guard carefully the people who have given us materials, guard the rights that they have set up with us. Of course, once it's out of our hands, we really can't do that much. For instance, on every Xerox machine in the library we have a thing that says, "You are responsible to make sure you don't infringe on copyright law, Section X," but we really have no control if somebody wants to come in there and wholesale plagiarize by copying page after page out of a book. An instructor, say, making copies of that to pass out to a class is infringing copyright law. The only place where we really clamp down is — the way the copyright law reads now is that it says the person doing the duplication is going to be held liable — so for instance if we have an instructor bringing videotapes to us, and he says, "Here, copy these; make a hundred copies of these to go around," we say, "Where's your copyright release?" from the producer of this program. No copyright release, no copies.

AUDIENCE MEMBER: (Portions of question inaudible) There doesn't seem to be private enterprise doing embellishments of...

SMITH: Well, I think there's some private enterprises out there doing a lot of this, and there's a lot of efforts similar to this on the national level. Smithsonian's been working on this; Library of Congress has had several projects like this. There's a lot of activity, and there are private vendors out there that will do some of this for you. We haven't seen anything within the oral history specifically, the oral history group in the United States that's doing this. We have one of the largest oral history collections around at our university. There hasn't been that much. But I think there's a huge market — that's why Apple Computer gave us a bunch of equipment to do this because they want us to test out their software. There's a hot competition there.

HERBERT DORDICK: Are you thinking of going online? Would it be possible for me to access this from another city?

SMITH: That's our eventual goal, to go online with this, so that you could dial into this and you could listen to the recordings. Right now we have some severe limitations with the pipelines. In computer terms, if you go on at 1200 baud to listen to audio you're going to be sitting there all night long. But those things I really think are solvable. With compression, with the ways that we're moving into fiber for transmission — that's going to be solvable.

DORDICK: Then you've given up the copyright protection. The only means of copyright that will work is when you have some node which can control. Once you go online you don't have that.

SMITH: That's one of the reasons we went with the much lower resolution on the photographs — we call them "reference-quality" photographs. And if you want to get a high-quality reproduced photograph then you've got to get in touch with us, and we'll make that for you. I think it's the legal issues that are going to be the sticky questions in all of this. The technology's moving ahead. The networking stuff I'm doing with CD-Roms right now — I'm

dealing with the vendors on that, and some of them say, "No, you can't do it at all," some say, "Gee, we don't care," and others say, "There's a fee we charge you for that." It's all over the place; it's a very sticky wicket.

DORDICK: If I were really interested, I could take your photograph and enhance it after I've received it.

SMITH: It'd still be possible to enhance the thing and get around it. Yes.

AUDIENCE MEMBER: I guess I still don't understand. Watching an NFL football game, they tell you you can't record it, and if you record it, you're doing what they told you not to do in the first place.

DORDICK: But that's okay. The copyright act allows you to make one copy for your own use. It's if you make copiés for all of your friends and charge them for it.

AUDIENCE MEMBER: Like you said, stealing's stealing. Instead of reading that, you might both have them read it and bring it out in an audio — the agreement. That's part of the deal. I don't see the issue. Because we have that in books. What's to stop somebody from putting on a Xerox a photograph of a book — it's the same thing.

SMITH: There's no way to stop them. But again, if you make multiple copies... There have been universities, for instance, that have had major lawsuits against them when they make copies wholesale that are used in classes. Because then the publishers are not paid for students buying those books or buying the periodicals. It's the wholesale copying and distribution. And where we would have problems here is if somebody without paying for the rights took some of these audio recordings and used them in a commercial product. If someone made something comparable to the Civil War series on PBS about Alaska, and they took all those recordings and made the videotape. And they sold the videotape, and then we'd see the tape, and say, "Gee, we never got a call from these folks for permission to use all these oral history recordings." There's where the rub comes in.