

Interviewee: Flynn, Mark
Interviewer: Valerie Emhof
Date of interview: November 24, 2008
Category: Florida (archives digitization)
Status: Open
Tape location: Box #54 (CD)

Emhof: Mark, can you start by telling me about your position at the State Archives? What are your duties and, more specifically, what are your responsibilities in regards to all of the digital collection?

Flynn: Well, actually, I am a library program administrator for information technology. But really, what I was brought in here to do was to start the Florida Electronic Library program. Now, Florida Memory has actually been here much longer than Florida Electronic Library. And Florida Memory is a program that really focuses on the assets of the state archives. And it is really a program that has been run out of Gerard Clark, over-site management out of his area, headed up by Jody Norman. The work that we do with Florida Memory is more a support role, to assist them with basically putting together whatever technology they need to make their web service work well. So, we usually get involved when they need help with the back-end of some of their web services. For instance, putting together a database application that sits behind one of their website; or perhaps doing routine system administration and maintenance, looking after their servers. When they need space, helping to speck out space and help them plan for expansion. But it is really a support role. The other thing that we do, however, is that when we go out and promote our digital programs, whether it is Florida Electronic Library or whatever it is, we always try to promote Florida Memory as well. So we go out two or three times a year, to talk to the K-12 community, we go every year to the Fame Conference, which is for school librarians, and there is about three thousand people there, and we talk to them about the Florida Electronic Library as well as Florida Memory, and some of our other web services that are particularly suited to students' use. And then we go out every year to the Florida Education Technology Conference, which is a meeting of about twelve-thirteen thousand. And we do the same there; we promote the digital collections and try to drum up support for them. Pretty much, the day-to-day management of the Florida Memory is Jody Norman and her staff with over-sight from Gerard. But we try to step in whenever they need support, for making sure their web pages are available for the wide public, making sure they have the resources they need. Also, from time to time, with expansion of some of their web pages. We have web design staff that is there to help them too.

Emhof: What do you do with the Florida Electronic Library?

Flynn: Florida Electronic Library is a program that is also run with Federal dollars, just like the Florida Memory project. It is primarily for creating a digital library for Florida.

It is a 3.5 million dollar program, of which ninety percent of the funding goes to licensing software, or licensing data. So, through the Florida Electronic Library, we provide access to online magazines for the whole state of Florida, like Times, Newsweek. All total about eight thousand full-text magazines, journals. A lot of multi-media is contained in that too. There is also a lot of resources pertaining to other more specialized interests, like health. Readers' advisory services, like: 'I have read this book and I am looking for another book that is like it.' We also fund the "Ask A Librarian" service, which is Florida's virtual reference cooperative. That is a hundred libraries in the state that have come together and they cooperate with reference staff to make available live online help and reference assistance from ten in the morning until midnight. So, that gets paid out of my grant. The management for that is actually down in Tampa, the Tampa Bay Library Consortium. Then another thing that we do through the Florida Electronic Library is Florida on Florida. Florida Memory is just focused on digital archives; actually it is focused on archives, digitization and scanning of analog materials that have great research interests and making it available statewide. Florida on Florida is focused on all of the digital libraries in the whole state. So the purpose of it is to create a union catalogue of all the digital libraries in the State. One search and you could pull information from all of it. So Florida Memory could contribute records, the metadata from Florida Memory, to Florida on Florida. But we also collected from digital libraries' digitalization programs all around the state. So it is really nice because if you are looking for information on a key Florida historical topic, like the barge canal, you can find information related to that in repositories all over the state. And rather than having to drive from place to place to place, on search pulls everything to your desktop and you are able to view it. So, we are looking at more and more, as community colleges and universities digitize their collections, their archival collections, we do more to continue to build that resource as well.

Emhof: What is your background in the field? What did you do before coming here?

Flynn: Before coming here, I was head of special collections and archives in Georgia Mason University, and prior to that I was head of collection development in the archives of the Jesuits in the New Orleans province, working for many years there. And prior to that, I was a hand-press printer in Italy for a short period of time. That is in a nutshell my whole career up to this point. So basically, I went from Loyola University where I was there for a good long while, managing all collections, but I started the special collections in archives there, and the archives for the Jesuits. Their records cover the whole Southern United States, from Texas down to Florida up to Tennessee. And it goes all the way from the Colonial period up to the current time. And then, George Mason University, for the most part we were handling all of the records from the Federal Theater Project for the WPA, which during the last Great Depression, it was part of the work in progress of the administration to get people back to work, to get artists back to work. And so, they also had a lot of interesting collection archives about new towns in America, places like Reston, Virginia, some of the new towns in Maryland. There was a big program in the Nixon administration to encourage new town plan. So, a real different concept in urban

development in America. Although very common I think, more common than in Europe. So we had lots and lots of records related to urban planning, as well. And so I came here, originally got hired ... well, I did spend one year at FSU at the DIRAC Science library, working with the papers of Paul Dirac, the father of Quantum Mechanics. And then I came over here and took the job as head of technical services. And so automated the State library, brought up the first full integrated library system, and then I changed position and became a consultant, because here in Florida we have a ... [pause]

[Interruption by a co-worker with a question]

[Interview resumes]

Flynn: So, let's see. We went from managing technologies here to ... our division operates grants programs, we give out almost forty-five million dollars in grants every year. And a large part of that is the LSTA, the Library Services of Technologies Act. And a large part of that was providing federal grant funds for libraries to do automation projects. So, I worked with that program for about four years, being the chief technologist that monitored grants programs for technology. And then, at that point, I moved into this position to manage the Florida Electronic Library, which was a brand new plan we had just developed and needed to be implemented as well as IT operations for the State Library and the State Archives; for both.

Emhof: How does working for an Archives or Library institution change from working for any other company in your field?

Flynn: Well, it is hard to tell because I have not had a lot of experience other than working in libraries and records. I have worked in libraries and archives for thirty years. The only other job I ever had was at the post office, and working as a deck hand on a river, on the Mississippi. So, it was real different from that. [both laugh] But, you know, again ... everything we do is based upon planning. I think that is key to the success of whatever endeavor. We may not have the ... we are not as driven, obviously, by things like profit, yet we are driven by other measures which are just as important. Which is filling a need in society and being able to deliver on it. And, in one way, that has to be evaluated and measured. So ...

Emhof: What would be that need?

Flynn: In Florida, primarily, it is the need for access to web-based resources about government; information about government. And, you know, it is very important for the life of a democracy to have an informed voting population. But also, people use government services all the time. They are dependent upon government services, for a lot of things. Whether it is for their trash collection, marriage license, or aid dependent children, social services. There are a lot of things that governments do that really private industries are not interested in because there is no money in it for them. But serving up

government information, making information available to broad ray of the population is possible. Especially information about their government, so that they can participate in government. Things like the Florida Administrative Weekly, for instance. It is a good example. It is a web-based service that we started a couple years ago; it had been published for many many years, but had not really been as accessible as it is now through the web. Basically, you know, the legislature passes laws, dealing with every aspect of life here, in our community and in every community of Florida. Those laws are implemented by agencies through a process of rule making, where the law becomes interpreted. And it could deal with anything from the size of your ... if you own a water-front property and you want to build a wharf to hook your boat up to. You know, the size that you can build it. To the designs of street corners in your neighborhood. Every aspect of what you do is governed by rules through rule making and through the legislative process. And so, this website, the Florida Administrative Weekly, allows people to participate in that rule making process. So that when an agency decides to pass a rule on a subject, a notice is prepared and goes out, and through google, or through signing up for this process, you can find out information about rules as they pertain to your interest, whether it is fishing, hunting, or whatever, driving cars, the size of your lot, whatever. All of these rules that you may have an interest in, you can participate in their development. So, that is an example of where what we do has a direct impact on people in the state. And then, as people come to us, because they want to know the history of their community, and we are the ones who collect the records that provide access to that. Legislative intent is another example where people come to us and access the records to the legislature, to find out what was really meant when they passed the bill twenty-thirty years ago. Or as the legislature begins the process of looking at new legislation. People will interact with our records and with our web services to participate in government and the development of new legislation, or passage of legislation. So these are all things that have pretty broad impact. And, really, the private sector has no interest in it, and there is no money in it for them.

Emhof: Now, to come back to Florida Memory, can you tell me about the storage system that you have put up for them?

Flynn: Well, as you know, Florida Memory took on the Florida Folk Life collection, and that collection includes massive amount of audio visual material. Both photographs, and a wide array of different kinds of files types, obviously a lot of audio and sound recordings. We knew that space was always going to be something of great importance. So, there is quite an array of storage housed at the central computing facility for being able to serve out these different web services, these file types that are in such huge demands. And there are in huge demands. But, the trick is, again, every year as Florida Memory project submits a grant, where they detail what their annual plan of service is going to be, what new things they are going to take on, is to try to stay ahead of their needs, and to be able to plan for their need of additional storage space as much as possible. I think that before they got into audio visual, they were relying on single USB ... Well, first of all they were relying at one point on backing up their digital images

using platinum CD-ROMs and DVDs. They went from that to using USB hard drives. They would just fill them up and then kind of store them. But, you know, both of those have pretty high failure rate. Even if you buy a stack of CDS, DVDs, there is a failure rate. And if you store materials and you don't test it over time, or don't regenerate them, there is problem over time. With the USB Portable drives, if you don't keep them plugged in and working continuously, there is a tendency to fail. So we were looking for a different way where ... Primarily what we were looking for was space to allow Jamie just space to work; where he could park a digital file while he worked on it. So we came up with a fairly low-cost approach to providing that work space for him, through these little single tera-bites raid array that are sitting out here. And they were low-cost. We were actually looking for something that we could purchase and keep it down to under \$1500. We were able to provide the four tera-bites. And again, that is primarily works storage space for him. We are currently now working with the managers of the central computing facility to define a real expanded approach to data storage for the Florida Memory that we hope will last them for at least two years. I think ultimately, given their growth and their need for storage, we would love to provide them with something like half a pedabite and three-quarter of a pedabite, or more, to allow them storage for a good 4-5 year period of time. But I am not sure our budget will allow. But we are now searching what we can do; we are probably looking at an expenditure around \$35,000 to put in place in the CCF area a storage array that is really all about housing the production files that will be served out to the public. So we are not talking necessarily about housing digital files for long-term preservation, or preservation copies. But we are talking primarily about digital data that we want to serve out to the public. So what we have now, the four tera-bites that you have out here, is for Jamie to work with, that is his workspace. We are looking at a much larger ray where we can put all the Florida Memory stuff in one area and make it available to the public, and hopefully have plenty of room for growth, at least two years, maybe three. Then, there is another project that we are working on, called the ... I always forget, I always have to look it up [searching through his papers on the desk] ... It is the PeDALS Project, which stands for Persistent Digital Archives and Library System. And what this is about is to create an architecture for digital storage for long-term [emphasizes on that term] preservation. So you are really talking about dark storage. Not necessary ... there is aspect of this program that will allow public access, but primarily what you are talking about is the ability to collect digital material that has archival value. We are starting out by just focusing on born digital material, electronic records of government, and born digital state documents, and being able to store them in a LOCKSS system, L-O-C-K-S-S, which means Lots of Copies Keep Stuff Safe. And this would be long-term permanent storage of these digital assets. And it is designed to be redundant; we have four other state partners, so there are five states all total working to develop this. And the LOCKSS boxes will be replicated, throughout the five states. So we will have all of Florida's records here, as well as Arizona, Wisconsin, New York and South Carolina. And we will have our data replicated in the other four other states. So if there was any disaster, we could recover quickly because of this redundant ray housed in other places. Again, it is a different kind of storage, and we are just starting with one tera-bite to kind of test it, to see how it works. And then the decision has to be with the

scale. But right now, it is just a pilot-project to test if it is a viable architecture for long term storage of our digital assets for preservation purposes.

Emhof: So, you are starting with the born-digital and then you are hoping that, if it works, projects like Florida Memory can also be stored that way?

Flynn: Well, ultimately, I think we would want to use it for storage of all our digital assets.

Emhof: So you think the LOCKSS system is probably the safest for preservation storage?

Flynn: I am not saying that [laughs] ... and there are other digital storage schemes, architectures, out there. The Universities of Florida have one. But this is certainly the most accessible for the money. And also it just so happens that the Library of Congress has given us a grant to try this. So, there are dollars here. It is an innovative new approach. But I think part of this is to really test its viability for long-term storage. But, also, will it scale to a lot more states? The truth of the matter is, I keep hearing different things throughout the course of this project, and right now we have been working for part of the year just simply planning this project. We are moving into a deployment and implementation phase in the next couples of weeks, where we actually are going to get the equipment. We have already gotten the software, and we are going to set up ... it is going to be about eleven servers all total over in our central computing area.

Emhof: So, where is that facility that you are talking about located?

Flynn: It is in the Northwood mall.

Emhof: So, would it be at the same location as the RAIDS for the Florida Memory access?

Flynn: No, those would be only the preservation copies. Now, the system has the ability to interact with our public access system. So if we were to take in, say a new state document or publication from Department of Transportation on whatever topic, and born digital; it has never been printed. It would be possible for us to take this and put a copy into the LOCKSS box for long-term storage. And at the same time move a copy over to our state e-dot program, where then it would become accessible to the public.

Emhof: So, is everybody working for Florida Memory moving to the RAID storage system? Is everybody using the same system now?

Flynn: Well, let me put it like this ... [thinks] ... No. Jamie is using the raid to raid for his workspace. The new system for serving all the digital assets out for the production system will probably more than likely be a raid-based system. The LOCKSS system is

not raid, it is very different. But it achieves redundancy through a different process than what raid uses. People like raid because it is recoverable very quickly. If you lose one drive, you can hot-swap it for another and immediately it will recover. The LOCKSS system works the same way, except your storage, your other redundant storage, is scattered throughout the LOCKSS network. And I have heard different things ... some people say you really need five replicated LOCKSS boxes to assure long-term preservation. Then they say seven, then they say fifteen. I think we don't know. But I think the more states that you can move to this kind of sharing, probably the more permanent. You know, there really truly is nothing permanent in this life. We all know that. But you do hope that through careful planning and collaboration with others, you can achieve more permanence than not.

Emhof: So, do you sit down on a regular basis with the archivists to discuss their standards, their needs in digital storage, what the ideal storage system would be, and what you can do right now?

Flynn: Yes, we do have regular communication. What we do is that in the instance of a ... for instance, the RAID to RAID that is out here. In essence, my staff researched it and made a proposal to Jody and Jamie regarding what we suggest and the money that they have. And then, basically, we would not have purchased it without their buying. The same thing will happen for this new production storage array. CCF will bring to us a suggestion, a proposal, based on their review of what is available on the marketplace and what they understand our needs to be. We will look at it; my technical staff here will look at it; Jamie and Jody and others will be free to look at it, and we will make a decision whether we want what they recommended or some version of it. And then, every year, the Florida Memory Project puts together their grant, same as the Florida Electronic Library. And all of these are actual written annual plans for what we are going to do in the next year. And these are shared among everyone, and reviewed, and commented on. So, there is a formal process of review of all these programs. And then, there is also on-going communication where we try to plan and build support through the years. And then, as things come up that are needed, Jody or any member of the team ... they may choose to come down and visit with my staff, talk. They may choose to send an email. They may choose to have a meeting with us. But, we are pretty much at their disposal whenever they need to communicate, or ... If we find out that there are issues ; if we get a call from CCF that one of their storage array is down and needs to be replaced, or there is a power outage, or there is some other aspect of the physical facility impinging upon the delivery of their services, we try to coordinate with them, so they know when things will be fixed. Or if they need assistance finding a work-around until things can be up and running, But, we try to keep them informed of anything that deals with their data and their services.

Emhof: Now, my last question would be, regardless of the money available, what would be your ideal storage system for a project like Florida Memory, especially for preservation storage?

Flynn: Well, I think again. We really are at the cross of learning about it. This is all brand new. I challenge anyone in the country to say that they know any more or less than what we do. I think we are trying something that has great appeal because it is economical, and has the potential for providing the kind of redundancy that can guarantee long-term access. It also has other characteristics of being ... automatic updates. If you pull something out of storage that was on Adobe 6, it asks you: "do you want to leave it with Adobe 6," the copy that you just pulled out, "or do you want Adobe 8?" So that it will work with the latest thing that you have on your browser. There are other aspects of it that are really, very automated. The automated workflow, that they are looking at developing with this system, I think has great utility. I think in the future archivists dealing with digital resources are primarily going to be concerned about the rules for handling large bulk of electronic records, rather than looking at things individually. And to have them automated where processing is going to be very very important. So it is really more than just digital storage, it is also having the middle-ware, the curatorial rationale, behind it. So that there is a lot of pieces and parts that have to fit together. And I think that LOCKSS works very well with what we have designed in terms of a curatorial rationale for this project, and what we are hoping to achieve in automated archives processing. Another ... You know, RAID to RAID may work, but still it has some failure points, that cannot be ignored. The fact that you may have a storage array, but it is located in one place. And if you lose it, because of a flood or something, it is not anywhere else. So, you know, for the fail-over of a system, they suggest that you put it at least three hundred miles from where your production box is located. And one of the aspects for LOCKSS, this LOCKSS network architecture that we have developed, is that we will have that redundant fail-over ability. And it will be at least in four other locations, scattered around the United States, from East to West.

[End of Transcription]