

## **Symposium: Digital Archives: Navigating the Legal Shoals**

### **Email as Archives: You Have to Have it Before You Worry About It**

Francis X. Blouin, Jr.\*

I am going to talk about the Digital Dark Age—and I love that concept—for reasons I'll say in a minute. And I also want to thank the Kernochan Center and the Rockefeller Archives Center for the invitation to be here today. For reasons that will be apparent in my talk, I've been chomping at the bit to talk to lawyers about some of the problems we're having in our current work. It's been interesting to reflect on the sessions earlier in the focus on intellectual property and copyright. I think that sort of pertains to stuff we already have, and my focus is looking to the future, anticipating collections that we're going to get. In this context, I think it's important to distinguish between archives and special collections, because the context of what I'm talking about is our integral work with the archives of our own institution. As Bill was saying at the beginning, institutional archives have a sort of organic relationship to their records, and we're entering into that kind of organic relationship. It's a different kind of relationship that we have with our own materials versus what we might have with a historical collection that we go out and get and bring in. So we're finding that it's really more—in dealing with this—a cultural problem than a technical problem, for reasons that I'll mention. So I'm calling this "Email as Archives," but I think the subtitle is: "You Have To Have It Before You Worry About It."

So, just by way of background, the Bentley Library at the University of Michigan is both a state-based special collection, which I am not going to talk about, and we are the university archives, which I am going to talk about. And we are independent of the university library; we stand alone and report directly to the provost, and that has interesting implications. Our clients, if we want to look at it that way—I thought I'd use that term here; I don't use it very often—our clients are the deans and executive officers, the principal administrative officers who are

---

\* Francis X. Blouin, Jr. is a professor in the School of Information and a professor in the Department of History at the University of Michigan. He is a fellow of the Society of American Archivists. Since 1981, Professor Blouin has served as the director of the University of Michigan's Bentley Historical Library, which houses the archives of the university and the principle historical collections of the State of Michigan. His most recent work, *VATICAN ARCHIVES: AN INVENTORY AND GUIDE TO THE HISTORICAL DOCUMENTS OF THE HOLY SEE* (1998), presents the first-ever conceptual overview of papal archival records. Professor Blouin has also lectured widely on archival issues throughout the United States, Europe and Asia.

generating records of importance for the archives, and with whom we have been working in building the archives over a period of seventy-five years.

So what I am going to do today—I don't have any legal advice; I am here to get legal advice—but I am going to tell you what I told our clients as to why we need to worry about archives in this Digital Dark Age. I remind them, to begin with, that the university archives is three things: it's a source of information, and it's the evidence of the reputation of the university, that after all of us are long gone and we can't toot our own horns, it's really the archives that holds the evidence of what we've accomplished and what we've achieved. And then thirdly, and it's kind of connected to that, is that we do have to think about how we maintain institutional memory in a complex bureaucratic institution. So, because I don't have a lot of time, I'll just say that we have a lot of information in the archives, going back to 1817. We have a lot of stuff that people come and use our archives for, in terms of the role that the university has played in the intellectual life of the nation and the world, and we gather this information from a variety of offices and places around the campus. As anyone who works in the archives of a major university will say, you put this collection together from a multitude of sources.

But now we have a problem in this Digital Dark Age. A lot of our records are born digital, and increasingly we're finding that the kinds of records we used to get in paper form are being generated in a variety of digital formats and within digital programs and communicated back and forth through email, among other systems. The University of Michigan, I am told, now in all its combined email systems generates 12,000,000 email messages a day, and we're not interested in having all of them. So, at the same time, the archive of the future is embedded in these emails and the attachments that they hold. What's been interesting to me as I've looked over the past thirty years is that we no longer at the university have the infrastructure of file clerks and assistants to maintain the archive in the semiactive phase of the records cycle. Those people have just gradually been whittled away, with budget cutting and so forth. So, what we find is that every person has become sort of his or her own records manager, and that's a challenge. And increasingly now, archives are of course born digital and are stored digital; they're not even printed out. That phase ended maybe about five years ago.

So, the records production and evaluation and preservation in the digital environment of what we consider our essential records for the university archives for institutional memory are now being kept extremely haphazardly. That's why I love the phrase "Digital Dark Age," because I did a considerable amount of work at the Vatican at one time, and the records-keeping practices of our university these days in the digital environment are very similar to the record-keeping of the Vatican in the ninth century, so literally a "dark age" analysis. There are just too many incentives in the system right now, either to destroy too much or save too much. Which is to say we're saving too much of the wrong stuff, and we're destroying too much of the good stuff. So, the availability of cheap storage of course encourages saving, but there are a lot of issues that are encouraging destruction. So, we have seen the need for an institutional—a virtual archive.

So, I asked my clients, you know, "How many of you have called on a

predecessor to find an email of consequence?" All the hands go up. Say some problem has arisen during the current dean's term, and they can't find anything about it. So, they call the predecessor who says, "Yeah, I'll look at my laptop; maybe I've got it here somewhere," and they kind of piece together the record. Whereas you know in the paper environment—not to say that the paper environment was ever perfect—but in the paper environment, there would be a kind of central files in departments that might have accomplished retention in a more systematic way. Then the second thing I like to ask is, "How many of you really value the ability to tell the story of achievements and innovations of your school and college?" *Then*, all the hands go up. So, we have a conflicting situation.

We were grateful about six months ago to receive a grant from the Andrew W. Mellon Foundation to try to get something going that will preserve email on the campus of enduring value. The idea is to have a system—a technology—that will separate the ephemeral from the valuable. And this is, in our view, the first step in trying to create a digital archive. So, what we have developed—and I'm not going to go into the technology of it because this is about the law and not about technology—a one-click mechanism that will enable people to separate email of ephemeral value from that of, what we call, enduring value. The mechanism allows full review and management by the person. It doesn't get immediately sucked into the archives. People can deal with it, and it also is connected to and will help us integrate our archival interest into larger content management systems, such as SharePoint and so forth. So, we are currently rolling this out.

But what I want to talk about really is the environment in which we are rolling this out and what my discussions with my clients often reveal. First of all, what will be the result of this? We tell them that we want to create a system that is manageable over time but subject to the review of the particular client. So, it's not supposed to be and will not be intrusive.

We also want to make sure that it fits into the work pattern. Every school and college is different. Some have a multitude of deans and associate deans and so forth, and others are smaller and, you know, have tighter or smaller central offices. So, it has to be able to be customized and we're working with that. But the big question that always comes up for our clients is, "How do they know what to save?" I mean, "What does the archive want?" Their concern is to work with us individually to identify a genre of records. So, what we are doing in the archives is looking at the history of our relationship with particular schools, colleges and offices, identifying the kinds of material that have flown into the archives in the past in the "pre-Dark Ages" and informing them with a kind of a list or a template of the kinds of material that we would look for in the future. And that's fine. We have these conversations and they go well.

But we have also the problem of points of resistance within the university environment. The most notable was a brochure that was widely circulated at the university and gave me heartburn three years ago that simply said to everybody in the university, "Once you've read it, shred it." So, the provost of the university and the head of the library and the vice-president of communications who sits on our executive committee and I pulled out this brochure and I said, "I don't think this is

sound archival policy. If it is, we are not going to have an archive.” So, that actually helped spur the effort in which we are currently engaged.

So, there is a problem in the inclination to “shred and be safe,” rather than to “save and be sorry.” This is, of course, increased by certain legal frameworks that the university has to deal with. So, there needs to be guidelines about what should be saved and what should not be saved.

But the question still is, “What does the archive want?” And we kind of lay it all out. But then the reaction is “How do we retain it when the *lawyers* are telling us—and that’s why I’m here today and really glad to be here—the *lawyers* are telling us that it should be destroyed?” So how do we deal with this? How are we going to address the problems that face us in this Digital Dark Age when it is so easy to press the delete button and simply destroy it and just remove the problem? It’s an issue that I think we definitely have to address. And then the problem is determining through a kind of rational sense of analysis what to save and what to destroy. We don’t have, in the digital environment—at least on our campus I don’t know about anybody else’s—we don’t have on our campus a really developed sense of the long-term value of information. We save it because it is cheap to save it; we destroy it because we are afraid to keep it. And that’s not an information policy in my view.

We have developed some categories of information that are not a problem in any legal sense and that do inform the future of what’s happened in the past. So, we are identifying things that point to the value of sound (i.e. in the historian’s sense balanced with the legal sense of the term) retention policies. I’ll pull out a great proposal that we have from the 1920’s on the development of the anthropology department or the original curriculum plan of 1817. And I say, “Look at how important these things are to our history.” And they look at it and say, “Oh, we can save that.” Then I say, “Save it then and let’s not get rid of it in some sort of mass cleaning of a hard drive.” So, there are categories of information that are not a problem. But what we need to do is have better guidance on how to define those categories within the legal environment in which we are working.

There are several implications of all of this from the point of view of the archives. One is—and as Bill was talking about at the beginning—archives used to be considered an organic process where records flowed through administrative processes. They were created, read, evaluated and ultimately filed, and then at some time in the future were transferred to the archives. And the idea was that if the records were organic to the administrative process they were really, in a sense, true to it. They were an authentic record of the entire proceeding. There are a lot of problems with that, but that was the theory.

Well, now we’re moving into a situation, into a technology, where people are actually choosing what to put in and what not to put in. And even if we do get up to the point where people are actually maintaining a digital archive, the question will be, “What kind of an archive is that?” How authentic is it to the total administrative process and to what extent is it really kind of a manufactured history? This is not really a legal question. But we may be designing the archive by creating more of an historical consciousness in the explicit guidelines for

retention of records for historical purposes.

However, we still have this legal environment problem. The legal environment for records creation and retention needs to be redefined with digital processes in mind. We have new types of documentation, new processes of authenticating administrative processes and we have to figure out how we are going to maintain the record of those processes while respecting all these various conflicting laws of, privacy, right to know, etc. I call this the legal value of institutional memory. We have to determine what the legal value of institutional memory is over the long term. And that value has to be balanced between the very highly and very precisely defined value of records destruction over the short term as implemented through legal imperatives for institutional records management. So I'm here—happy to be here—to plead to all the lawyers in the room who came here out of an interest in this issue, to help us define the legal value of records preservation of institutional memory over the long term.

The critical point now, too, is that selection has to be done at the point of creation, more often than not, by the person generating the document. That is something entirely new, where we put the responsibility right at the point when the document is being created. And that's a culture change. And trying to guide that process, and trying to give people a certain comfort level in doing that is an administrative concern, but it is also a legal concern, because those decisions often have the specter, if not the strong arm of the law, behind them. And the difficulty is that if we don't have some system behind the process, we will have problems. We are finding now that people are simply accumulating messages; they come to the end of their term, they say, "I've got all this email. I've got 100,000 messages, 250,000 messages, and I don't have time to go through it and I can't give it all to you, so I'm just going to get rid of it." And I think that's not sound information policy. So we need clear legal guidelines that respect institutional memory as a context for this selection process.

So, in sum, I'm here really to say that the problem of this Digital Dark Age is to establish a context for preserving institutional memory, and the sum total of preserving institutional memory of all the institutions represented of all the archives is essentially preserving the national memory. So, this is a huge issue in how we're going to understand the twenty-first century. I'm happy to be here in a law school at last to plead for action on behalf of digital archives, the digital archives of the future. I think it's important for us as we think about the law—and I certainly agree that these copyright issues are extremely important—but copyright is going to be a moot point if we don't have the records at all. We need to develop a sound infrastructure for the retention of records in the digital environment. And this will mean balancing the right to privacy and the right to know; the right to restrict versus freedom of information; the necessity of records policies for retention and disposition versus the freedom to be relieved from records policies on retention and disposition. And all this has to be resolved before the concerns of copyright enter the story. Thank you.