Missouri J-School and the 'backstory'

Victoria McCargar Consulting 5057 Gaynor Avenue Encino, CA 91436 **T** 818.317.9515 **F** 818.501.4990 mccargar@mac.com

Table of Contents

Introduction 1

1908-2008: The Missourian 'backstory' 1

'SWOT' analysis 3

The project 'backstory' 4

Two years later 5

The Missourian 5

Reynolds Journalism Institute 6

A hefty responsibility 6

Toward a news repository at MU 7

The infrastructure problem 7

Eight steps 8

Step 1: Designate a project owner 8

Step 2: Form a project team 8

Step 3 (concurrent): Inform yourselves 10

Projects 10

Conferences 11

Step 4: Hire an archivist 11

Job Description 11

Step 5: Establish policies 12

Step 6: Undertake a small-scale or pilot project 13

Step 7: Go after grant money 14

Step 8: Hold a symposium 14

Can this be done? 16

Is the backstory worth saving? 16

Further reading 17

Introduction

1908-2008: The *Missourian* 'backstory'

When MU's School of Journalism opened its doors in 1908, Associated Press was already 54 years old—a venerable history. However, one of the things AP didn't pay much attention to until very recently was its archives.

As a result, its news legacy, in pieces, only goes back to about 1937. That is a gap of more than 90 years.

Fast forward to right now—June, 2008. The 162-year-old AP just released a fascinating report on a topic of wide interest in publishing: How are Gen Y'ers and the Millennials, those coveted 18- to 34-year-olds, consuming news in this mobile age? How do they get it? What do they read? What do they want? How do you advertise to them? (http://www.ap.org/pages/about/pressreleases/pr_060208c.html)

To get the answers, AP hired a team of Baltimore-based anthropologists to look at the digital news culture. They studied groups in the United States, India and the U.K. What they found surprised AP's strategic planners. After concentrating on delivering news in as many media and diverse formats as possible, AP found that young news consumers were overloaded with headlines but hungry for context, the background to the headline, the history behind the news, or what AP calls in the report the backstory.

How a news organization delivers the backstory to digital consumers is a technical challenge that we'll leave to Associated Press to figure out for now. More to the point, there has to be a backstory to deliver. The archives at AP are getting a lot more attention these days.

Like AP's a few years ago, the current mindset that the Missourian's news archives are "nice to have" but very low priority sets it up a situation of inevitable loss.

We needed another word for archives and AP gave us one. Newspaper backstories are in peril all over the world. As print ad revenues shrink and cost controls widen, the libraries and archives and their staffs make attractive targets. Naively, the web is viewed by many publishers and editors as the big database in the sky, so standard approaches like internal databases and microfilm start looking impossibly old school. Trusting the backstory to digital formats is an experiment on a civilizational scale.

The Missourian has already lost part of its 100-year backstory, a digital part. As we all know, the system crash in 2002 wiped out 15 years of text and several of photography. That digital backstory is now inaccessible to digital news consumers. True, the short-term backstory is on the website but that model is unsustainable. Websites are not archives. There has to be a systematic approach to capturing the backstory currently parked on the web servers, or it, too, will be lost.

Above all, if the University of Missouri-Columbia School of Journalism, the Reynolds Journalism Institute (RJI), the administration, faculty, students and staff don't think archives are important, there is no need to proceed further into this report.

But if the backstory is worth having around for the bicentennial in 2108, then the School is going to have to learn how to manage it as part of everything else it does.

Associated Press discovered in the mid-1990s, at the dawn of news on the internet, that they'd relied way too long on benign neglect to carry their news coverage into the future. This is a lesson that news organizations have been waking up to for a decade now, but, ironically, the alternative–aggressive, continuous intervention in the technology to keep it alive–comes at a cost that few are able to meet these days.

When I visited Columbia in April, our discussions focused on two areas – the prospects for rescuing the Missourian's digital archives, and how the Reynolds Journalism Institute might contribute to research in news preservation, especially in the context of hosting the development of a formal repository.

The following proposes steps toward that goal.

'SWOT' analysis

SWOT analyses are useful in laying out the major issues in a proposed project. Whether the J-School and RJI are successful in addressing the backstory problem depends on working through some of these considerations and concerns.

Table 1.

Area	Issue	Observation
Strengths	Visibility	J-School Centennial, opening of RJI, current controversy over Missourian print product
	Scope of university resources	Cross-campus collaboration with J-School, University Libraries, SISLT, College of Engineering
Weaknesses	Infrastructure	There is no infrastructure in place to undertake digital archiving
	Commitment	Among competing priorities, Missourian's archives are expendable
	Track record	Missourian has already lost 15 years of its digital backstory
Opportunities	Funding	Grants are available for digital preservation infrastructure projects
	Industry and research initiatives	New tools for web archiving, automatic categorization, permanent repositories
	Institutional interest	New willingness to tackle the Missourian's backstory.
Threats	Isolation	J-School's reputation as an independent entity on campus
	Status quo	Failure to train students, faculty and especially IT staff in new procedures
	Funding	There is no funding for any project having to do with the archives. The Missourian is already losing a lot of money.
	Bureaucracy, internal divisions	Librarians are routinely out of the loop on J-School technology developments and decision making
	Potential loss of donor material	The J-School is taking in donations of historic collections with no concrete plans or predictable funds to manage them.

The project 'backstory'

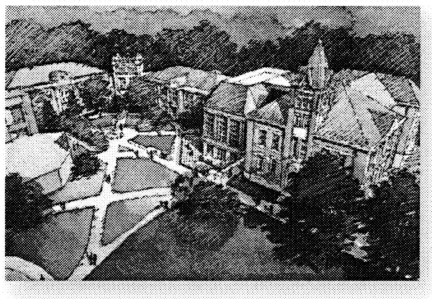
Let's review. What brought me to Columbia at the end of April?

- Pick up where we left off in 2006 regarding rescue of Missourian database, lost in a system crash in 2002.
- Check on condition of Pictures of the Year originals and progress with physical stabilization and digital access
- Hear from RJI team about plans for creating and maintaining a digital record of the institute's future projects
- Advance the discussion of repository model for news and how Mizzou, the J-School and the Institute could play a role.

When I first visited Columbia in August, 2006, I made a number of observations (report dated Aug. 26) about problems I saw at the J-School that put historic, probably irreplaceable, content at risk.

- I suggested that the Reynolds Institute planning team add to its agenda a discussion of how it would make available to future scholars and historians the products of its projects and programs—in other words, was it planning to create a new journalism archive.
- I recommended that the PoY collection be looked at professionally so that, if nothing else, it could be benchmarked to measure against future deterioration while more suitable, permanent storage was arranged.
- I asked that IT support at the Missourian look for the backup disks of the database, which some people thought might still be in existence. Without them, obviously, discussing a rescue is moot, and an alternative solution to restoring digital access (such as scanning the microfilm) would be required.

About six months later, the Missourian disks were located but not further action was taken until my return to Columbia.



Two years later

I don't think it surprised anyone that, with the exception of PoY and the University Libraries' campus-wide repository activities, little had advanced in 18 months. The situation that we encountered in 2006 prevails today, explained by the virtual mantra among the J-School stakeholders: No time, no people, no money.

The *Missourian*

The situation has grown more acute with the announcement in early June by Missourian Executive Editor Tom Warhover that the paper's losses, approaching $1 million a year, had reached a point where print publication was no longer feasible without help from outside, possibly a partnership with the Columbia Daily Tribune. Another alternative on the table is all-digital or mostly-digital publishing, relying primarily on the web to present the news.

Both of these prospects have a potential significant impact on the way the Missourian keeps the backstory, particularly publishing to the web. All-digital publishing is a growing phenomenon in newspapers but the processes for capturing it and retaining web pages permanently (and their all-important hyperlinks) are in their infancy as well. One of the discussions was fruitful for alerting the Missourian IT staff that the columbiamissourian.com website was not a permanent archives, even if "search the archives" is a labeled function on the site.

Meanwhile, the missing database disks were presented at the final meeting on May 1, consisting of several boxes of 3.5-inch floppies and a batch of CD-ROMs. It was decided then that the University Libraries and RJI would contribute several thousand dollars apiece toward a rescue.

Two weeks later, Missourian IT reported that the vendor of the defunct system, did, in fact, still exist—contrary to earlier reports—and offered to contact the vendor. That action is still pending.

Reynolds Journalism Institute

While everyone agrees that archives are "nice to have," they are pretty low on the list of the Missourian's and J-School's priorities. At the higher-level discussions with the Reynolds Journalism Institute team, however, there was enthusiasm for taking a leadership role in a news repository venture. Not only would it be high profile, but it would be part of a national effort to build up a permanent digital infrastructure, and would serve RJI's mission.

The repository model has been around for several years and is no longer the rocket science that prevailed with startup projects a few years ago. Criteria for trusted repositories have been tested against real-world systems, and they reveal the characteristics that make an institution a good bet for long-term digital preservation.

What I find challenging, though, is trying to understand how an organization currently so ill-equipped to deal with its own relatively simple archives at the Missourian can ever reach the point where it meets those criteria; put another way, what would be required to position the J-School and RJI to take on something as sophisticated and complex as a trusted national repository for news? That is the question we need to answer.

A hefty responsibility

Ready or not, the School of Journalism and by extension RJI are already acting as a repository in some ways. As I pointed out in my 2006 report, the J-School is the physical home of the archives of the FOIA Center and Pictures of the Year, both collections on the level of national treasures. Yet the PoY photos are stored in the university's records management warehouse on the outskirts of Columbia, not an optimum location (although not as bad as it could be). Myriad rights concerns preclude a safer haven at the State Archives. While the photography exists on black-and white microfilm and has been scanned and catalogued up to a certain point, it's a hope-for-the-best situation with the originals and not a long-term plan.

Meanwhile, the photojournalism faculty has been actively soliciting the personal archives of one of the lions of the discipline—former professor Angus McDougall. There is, however, no formal plan for how the material will be safely stored and accessed, only some uncertain prospects for a future museum near St. Louis or building on campus. I know everyone involved has the best of intentions—and I'm the first to argue that McDougall's archives should be in Columbia—but to be negotiating that kind of gift without a program in place to care for it is flirting with a breach of ethics, at least in the archives profession. It would be better to see those photographs go to another repository than put them at risk.

Then there is the prospect for future born-digital collections. When the next Angus McDougall arrives at Mizzou, where will his or her pictures end up? The Missourian photo database is already a couple versions out of date. CD-ROMs are not archival. A third-party quasi-commercial repository like ArtStor—proposed for PoY—is a good solution for access but not for preservation. What can you tell a photojournalist about the future of his digital work?

The gap that I am hoping to close here is the one between intention and reality, between institutional willingness and practicality. In the next section, "Toward a News Repository at Mizzou," I discuss some steps for starting small and building. We're starting from virtually zero.

Toward a news repository at MU

I realized after I left Columbia that what the J-School and RJI would really like to have happen is to bring in somebody else to worry about this, so everybody can go back to their real work.

Unfortunately, it's not that simple. Building infrastructure is painful, and, like re-piping a house, the rewards are pretty invisible. Building the backstory cannot be done from outside; you can't hire a dozen temps to build a system who then disappear from the scene. Continuity is critical.

Digital preservation, in the form of a repository, has to happen organically within the university as part of the creation of knowledge. Creators have to take ownership there will be no commitment to the backstory. Left to one person or a temporary group, it will always be someone else's project, and this can never be about one or two people—it's about an entire institution's willingness.

The intellectual considerations of publishing online—information channels, user-applied tagsets, packaging, web formats, linking to related elements—all apply equally to the art and science of keeping the backstory, and need to be integral to the teaching of online journalism. I would argue that the web-oriented multimedia journalist who graduates from Mizzou needs to understand some information science.

The infrastructure problem

As I mentioned above, we're starting more or less at zero. The Missourian has a web site and a MerlinOne database that captures some of the text and images coming out of the printed newspaper. Other publications—Vox, the website—may be archived on an ad-hoc basis. The database, for financial reasons, is not running the current version of software (the same situation that existed when the previous system crashed in 2002). The IT team considers the website to be "the archives" and puts its resources there, meaning an already strapped crew is supporting two systems doing more or less the same thing.

Can we build on this thin foundation? Think the answer is yes—at the moment there is no alternative. When RJI opens the day will quickly arrive when simple server architecture won't function for storage and retrieval and a database solution will be needed. These steps, I hope, will take us up to that point and beyond. The main thing is wanting to do it. If the institutional willingness is there, here are my suggestions for how to proceed.

Eight steps

Before it can wade into the big repository model, the J-School and RJI need to understand how to keep digital content on a small scale. I freely admit that I am blowing right past all the institutional barriers and ambiguous lines of authority to get you started; those are for you to figure out.

Step 1: Designate a project owner

The vacuum here is intense. I got the unmistakable impression at our meetings that no one wanted to touch this with a 10-foot pica pole. Well, expectations have to change.

True, you could bring in a project manager at a pretty high level, but that would have to come with a fat budget and a great deal of clout. That scenario is highly unlikely, so we'll work within the situation as it is: people and resources on hand today. If you hire the right archivist (Step 4) that person can take over, but you're not ready to hire an archivist until some other decisions have been made. Also, in the kind of cross-domain team I'm proposing, leadership has to be on a fairly high level. Does this person report up through RJI? The libraries? Both?

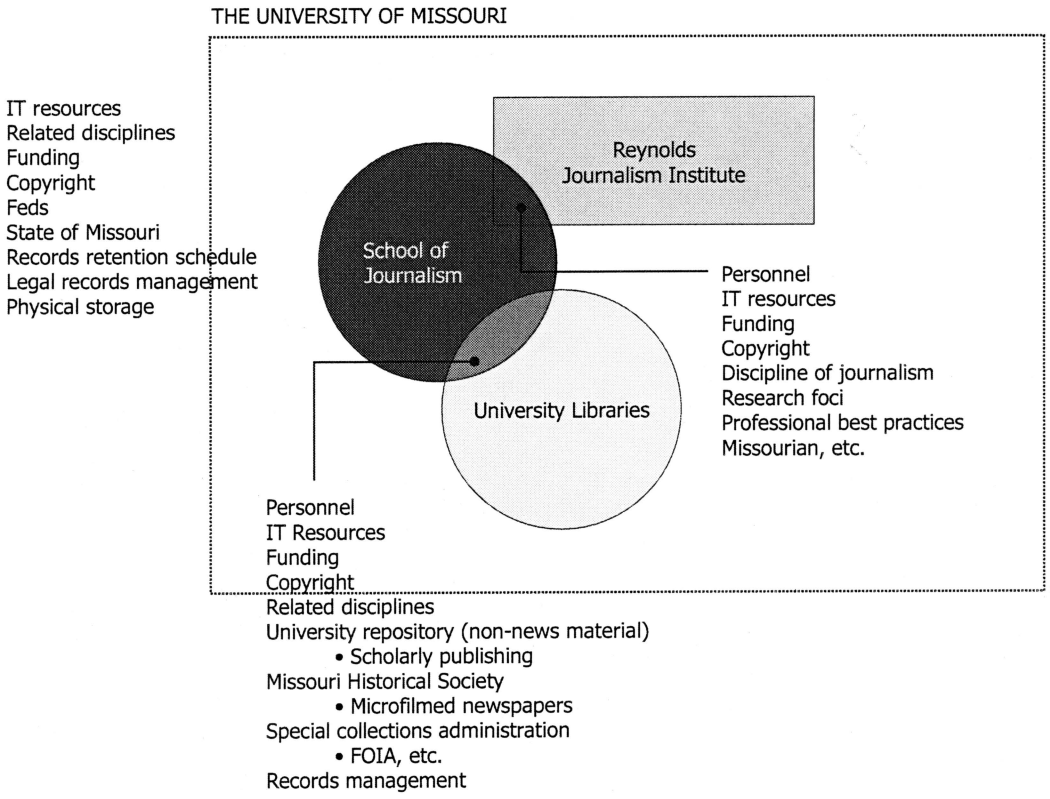
Step 2: Form a project team

With a leader in place, move to start building your team. Digital archives and repositories embrace all of these

- Content
- Research needs
- Metadata and data standards
- Policy
- Hardware and software
- Rights management
- Money

You can't overlook any of those aspects, so from within the university you will need someone from RJI, the J-School, the University Libraries and Archives, representatives from computer science, library school, a copyright expert. As you move ahead, find a grantwriting guru and someone familiar with budgeting.

Chart 1. Domain map. Some of the overlapping considerations in a repository.



Build in accountability. That means some kind of identifiable reporting structure, regular milestones and reports to administration. Each of these areas will probably have its own sub-group. At every meeting, you'll be needing to check in on the progress of the others. There are a lot of show-stoppers embedded in this project, many of the same ones that stop the show today—lack of communication, competing priorities, turf issues.

Treat it as a business. A formal business plan should be a goal. Your funding agencies at the federal level may require one, in fact. That means agendas, minutes of meetings, documentation of steps, spreadsheets.

The team should share its findings. Since this is a test bed, it means publishing results along the way. This has the signal virtue of being a digital preservation project that will generate interest outside of libraries of archives. The availability of the news to future generations should be taken seriously by the academic community.

Form an advisory board of experts involved in ongoing projects. Like a good board of directors, these will be invaluable sources of knowledge, industry support, conduits to funding and people, and provide reality checks. (I've met all of the following except Katie Hafner:)

Table 2. Advisory board prospects.

Name	Current post	Relevant work
Bernard F. Reilly, Jr.	President, Center for Research Libraries	Lead investigator on TRAC and other projects. CRL is a major analog news repository (among other things)
Brian Lavoie	Economist, OCLC	Expert in the economics of repositories and frequent research project participant
Paul Wheatley	Head of digital preservation projects, British Library	Researcher
Valerie S. Komor	Corporate archivist,	Associated Press
Jim Kennedy	Vice president for strategic planning	Associated Press
AP's digital future	Katie Hafner	Technology reporter, *New York Times*, writes about digital preservation

Step 3 (concurrent): Inform yourselves

There are listservs, study groups, innumerable projects documented on the Web. There is a large body of literature about digital preservation, but almost none of it treats news; it is virgin territory for research.

Here are some places to go to learn the basics of digital preservation:

Projects

- Open Archives Information Systems (OAIS). Although it is a NASA report, it amounts to Preservation 101 because it provides a framework for understanding digital preservation. It is cited in thousands of reports, but to start with the document itself, go to

http://public.ccsds.org/publications/archive/650x0b1.pdf

- Trusted Repositories Audit and Certification work. This would be another good starting point. Look at the TRAC criteria and start getting a handle on what is required. Item No. 1 is institutional willingness—the ability to commit the resources to undertaking digital preservation. The point of TRAC is assure depositors that you can be trusted with their irreplaceable material.

(http://www.crl.edu/content.asp?l1=13&l2=58&l3=162&l4=91)

- Look at the Joint Information & Systems Council (JISC) in the U.K. The Europeans are way ahead of the U.S. in this area. JISC has an interesting repository listserv that keeps track of standards and developments in Europe and elsewhere—good for understanding what actual practitioners are dealing with day to day.

http://www.jiscmail.ac.uk/lists/DIGITAL-PRESERVATION.html

Conferences

Part of investing in a repository is continuing education. This in this realm update continuously.

- Society of American Archivists. The movie stars of the field show up here. SAA is in need of a News Backstory Roundtable if you're interested. It's in San Francisco in late August this year.
- IFLA—the International Federation of Library Associations—used to have a once-every-10-years newspaper conference. They're starting to get worried so now it's oftener. It's in New York City this October.
- The Special Libraries Assn. has the respected News Division, a dedicated group of news librarians and archivists, desperate for solutions. These are the people who can really be helped by your work.

This part is easy—it only takes curiosity, interest and dedication. If you have those, the team will find the time. Move on to step 4.

Step 4: Hire an archivist

Now that you've got a team in place, one that has schooled itself in some of the basics of digital preservation, you're equipped to find a professional to help take it to the next level. Needless to say, this won't be entry level or part time.

Job descriptions are quite variable and often ask the impossible—such as, "this person shall be responsible for ensuring that all electronic formats are preserved in perpetuity." Anyone who understands digital preservation knows that this is neither possible nor desirable. Any candidate who promises it is underqualified.

If you decided to go for full project leadership, any candidate with hard repository management experience will come with a high price tag—including the expectation of a tenure track faculty position. Under the circumstances it seems preferable to go after someone who has come out of a graduate library science program since 2000, has a strong interest in digital preservation and has at least two or three years' experience with managing digital assets. Here is some suggested language, a composite of several recent job postings I collected from the academic, nonprofit and for-profit sectors

Job Description

> The Reynolds Journalism Institute at the University of Missouri School of Journalism is undertaking the creation of multimedia, born-digital journalism archives with a data horizon of 100 years. To that end, RJI and the School are seeking a full-time digital preservationist to guide, manage and help implement the system. The successful hire will
>
> Participate in the development of policies, practices and standards for digital archiving of news stories, photography, video, audio and graphics.
>
> Be the pivotal member of a cross-domain team including journalists, information scientists and engineers in envisioning the system, articulating requirements and determining the optimal technological solution.
>
> Develop practical and policy-driven procedures and workflows for accessioning the daily output of the Columbia Missourian as well as other digital objects produced by the school in the course of its teaching and research.
>
> Over the next 3-5 years, work with the team to guide development of a permanent repository for news within the general criteria set forth by the Trusted Repositories Audit and Certification project.
>
> Qualifications include a master's degree in library and information science with an emphasis in management of digital archives; at least three years' experience managing digital assets, knowledge of metadata standards,

including those for long-term preservation. It is expected that the successful hire will enter the position with a high level of knowledge about the state of digital preservation research projects and be committed to staying current. The visibility of the position and importance of the project to news preservation also suggest an opportunity for contributing to the literature about preservation, repositories and issues unique to news archives. A journalism background and experience with grantwriting are pluses.

I'm not equipped to talk about competitive salaries in mid-Missouri but a comparable position on the West Coast would probably command at least $60,000 a year, more in the private sector (as much as $100 an hour). Archives jobs traditionally pay less than librarians' but that is predicated on the old paper records management model. The job is more akin to a university librarian with responsibility for acquiring, managing and budgeting for digital resources.

Once your archivist is on board, you are ready for the fifth step.

Step 5: Establish policies

You must describe in a formal way what you're going to go after and sustain in your archives—what part of the backstory you are going to keep. The shorthand for this is "Choose what to lose": with finite resources, decide what is worth saving. Written policies and a mission statement are worth their virtual weight in virtual gold, because you can design systems around them, adjudicate disputes with depositors; delete marginal material with a clear conscience. Some thoughts:

- In April we spent a lot of time talking about the "news repository." What is news? This isn't an easy question to answer, and deciding what of all possible news is worthy of repository treatment isn't as simple as you think. It will be a fascinating discussion. You might want to join it.
- What goes into a news repository? News is not like journalism research. Does scholarly belong there?
- Similarly, the administrative records emanating from the J-School and Reynolds institute are not news and they're not scholarly research. They have a separate treatment as well in the University Archives. There are important reasons to make decisions based on existing policies, which is where the University is going to have a say.
- Be careful with the kind of formats you commit to preserving, and keep that list short. If an artist insists on drawing in exciting new software, ask him to stop, or at least to submit his final work in a program you use.
- A graduate student's one-off project in a home-built database is impossible to support for very long. The whole subject of multimedia master's theses and dissertations is an ongoing discussion that impacts University, Library and J-School policy. If a culminating project is a news package, where does it belong?
- Within Step 5, this is probably one of the most important tasks: Start identifying and resolving copyright issues that have come up in all these discussions. What rights does the School of Journalism have to deposit students' Missourian work in a repository? What actions can Pictures of the Year take to digitized, store and preserve the collection? What are the restrictions on public access to FOIA content? Angus McDougall's donation? Will material have to be embargoed (placed in digital "dark" archives)?
- I kept hearing about various copyright unknowns like PoY, and, while, they may not be knowable, we should develop a running list and get answers to each. I'm on the side of archiving now and apologizing later. It is highly doubtful anyone will sue you for trying not to lose something.

While you're talking about policy, you can also be moving on to Step 6.

Step 6: Undertake a small-scale or pilot project

Learn by doing. It's how every other institution is doing it in this realm. Recall from the SWOT table that there is no infrastructure within the J-School and, moreover, your current IT staff isn't up to speed yet. You'll have to build it.

Can you use the Missourian's existing MerlinOne database? It's hardware, software, a decent search interface and accepts a variety of formats. I don't know the logistics of partitioning some resources for a pilot "news repository" but it is worth investigating with the vendor. This will undoubtedly require an upgrade, which is already a risky situation, a couple versions behind what they're actively supporting. Budget for the upgrade one way or the other.

> Note on new developments: The transition to web publishing of the Missourian may force the issue. MerlinOne does have a tool for archiving HTML, and if the existing system is outdated the latest functionality isn't going to be available. Upgrading MerlinOne would go a long way toward ensuring that the Missourian doesn't lose any more digital content than it already has.

But is MerlinOne the best spot for the pilot project? Possibly, but are there better solutions? This is where your cross-discipline team will be crucial. It makes more sense to me to look for University resources and get some help from computer engineering. This project is about institutional learning. I can't imagine you wouldn't garner interest on campus. Look for solutions you haven't thought about.

When you've found a system for your material—MerlinOne, DSpace or something else—practice with the Missourian output. You'll need programmers (from the Engineering School), metadata analysts (from SISLT) and newsroom workflow best practices to make it happen—all potential research projects in three domains—computer science, library science, and journalism.

Perfect your text feeds before you move on to photos. By the time you get pictures figured out, maybe video formats will have stabilized. Your team will have to decide whether it's feasible even to take on more than text and image archives, weighing cost and complication against future use and research needs.

If you're good at text, it may be time to branch out to other newspapers. The Missouri Press Association is a logical start. Here you'll learn about forcing your contributors to adhere to your standards for submission, and they won't like it at first. But there are benefits for them too.

> Another note on new developments: Since April, the Associated Press is offering members a really intriguing deal. If you submit your content to the AP, they'll tag it in their new automated categorization system and return it to you, or they'll even store it. In return, you get a break on your AP fees. I don't have all the details, but the benefits are uniformly structured text according to AP's standard and ready integration with AP's web utilities if you're using them, including web video. This might be well worth looking at; contact your AP reps.

As for the Missouri Press Association, they'll need to know that this isn't a free service. There is nothing wrong with expecting newspapers to contribute nominally to the maintenance of the repository. The profession of journalism and the publishing business together need to buy into this notion of news preservation, and here is a way for them to do it in a tangible way. Their reaction will be a research project unto itself. (Again, the AP tagging project means that at least some of the MPA material can be uniformly structured, and when AP figures out how to link automatically to the backstory, you'll all be in a position to take advantage of it. Content uniformity is basic to a repository.)

So we've got a pilot project going. Who is writing the checks?

Step 7: Go after grant money

Your grant-writing guru on the team will have been on the case all along. There is money out there for repositories and digital preservation projects. Here are the largest:

- NDIIPP—The National Digital Information Infrastructure Preservation Program. (http://www.digitalpreservation.gov/library/about.html) Congress has funded about $100 million through the Library of Congress to develop partnerships—like matching funds—that will result in usable models and systems toward a national infrastructure of repositories and programs.
- NDNP—Another Library of Congress project to digitize backfiles of newspapers state by state. The copyright has to be clear, meaning before about 1920; funded by the National Endowment for the Humanities, among others. These people would be interested in what Missouri's doing. New guidelines due this summer. (http://www.neh.gov/grants/guidelines/ndnp.html)
- Mellon Foundation—They've funded the "Trusted Repository Audit and Certification" project, TRAC, and a number of repository projects as well. They're very interested in digital archiving projects. (http://www.mellon.org/internet/grant_programs/programs/rit)

Be very sure of one thing: This is project with a beginning, a middle, but no end. Digital archives are essentially patients on life support. You tend a system, manage the life cycles of its objects, back it up, migrate it when formats become obsolete (refer to the TRAC criteria) but you never shut it down.

One of the big funding issues, not surprisingly, is "sustainability," which refers to the problem of trying to operate a system "forever" on a grant-to-grant/hand-to-mouth basis. That warrants a firm understanding with depositors that there are entry and membership fees to at least bridge the gaps between grants.

It also means looking at the backstory, as AP is, as a source of revenue. There are no "knowns" in this area. Newspapers have accustomed their readers over the last 10 years to expect free content, and the public's tolerance for viewing and downloading fees has yet to be tested (although the preliminary results aren't promising). The economics of a repository are the subject of a high level of inquiry, but it is safe to say that grants alone are not going to be enough. There are no answers today, but your team is there to stay abreast of the issues.

If this works, you've formed a cohesive group across campus, are tapping regularly into the experts in the field; you're ingesting the nightly Missourian feeds and confident about copyright. You're publishing your research results, and the beleaguered newspaper publishing community around the world is starting to take notice. Time to go public.

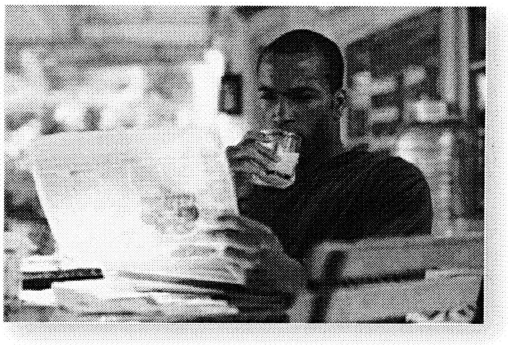
Step 8: Hold a symposium

There was some some discussion in Columbia about putting on a symposium in 2009 that would cover the topic of news preservation and the role of the repository model. Within a short time, you will have assembled a team, project management, the bases for digital policy and a small pilot project, with enough initial funding to at least see some results from the pilot. A year and a half might be enough time to have something initial to show for all of it.

Even if you haven't reached the pilot project stage, I can think of some topics from the library, journalism, engineering, business, legal, etc., domains:

- What is news, and how do you decide what to keep? (journalism, library science)
- How can algorithms help? Is there a "Google" solution for recognizing news? (library science, engineering)
- What are the differences between news and scholarly publishing when you're talking about the repository model? (journalism, library science)

- How have we finally conquered some of the difficult copyright issues emanating from the 2001 Tasini decision? (journalism, law, library science)
- What is the tolerance for a fee structure to support the repository among Missouri newspapers? (business, journalism)
- What does really ideal news metadata look like? How can it work to link the news with the backstory? (library science, journalism)
- Incorporating the backtory in journalism education—tagging, packaging and designing in online publishing (journalism, library science)



Can this be done?

Is the backstory worth saving?

The potential of the project pivots on those two questions.

The more I mused on my week at the J-School, the more I was struck by how much there is under one roof, or if not one roof, at least roofs within an easy walk of 9th and Elm. What I've heard all week—not to mention two years ago—is that there is no time, no money, no people. I'm sorry, but there will never be time, people, or money, and if that's the only response you've got, perhaps we should let some other outfit take this on.

Will the backstory fade away? There are a lot of entrenched ways of doing things that have vanished in the new reality of digital access. One of them is consigning the "old" material to a morgue (even if the morgue is a database) and forgetting about it. If you believe that putting things in a database is the end of the process, yes; the backstory will fade away—even if people continue to read the paper on newsprint.

But as the new AP report found, digital news consumers expect a seamless interface between the new and the old, and when they want the backstory, they don't want to have to go looking for it. Even if a national news repository seems too big an undertaking for the J-School and RJI, you'll still have to contend somehow with the Missourian backstory.

To draw this full circle, let's return to the defunct Missourian database. There was a string of emails the middle of May when the vendor of the defunct database surfaced with the good news that the system was still supported in Oracle—a "modern" system.

Assuming the bulk of back data would still be irretrievable, the issue of a Plan B was quickly raised—digitizing the microfilm. The Missourian team enthusiastically offered to get a scanning quote from a small vendor. This is good. Action at last.

Unfortunately, scanning and columnar news page geometry and making it searchable is much more complicated than ordinary text, by a factor of magnitude. This was duly reported back to the Missourian by the library team. There, one more time, the communication stopped.

It is a telling episode, because it reveals one of the basic impediments to keeping the backstory viable: the discontinuous borders of the School of Journalism with the rest of the university. This is not a journalism problem, and it's not a problem that the IT department of a single small newspaper has the wherewithal to solve. Maybe the time, people and money that no one has are really across campus.

Let's just try to solve the Missourian database problem. It will be a good way to find out whether the J-School and RJI will be able to take on more, like archiving the website. Even if the rescue fails, a lot will have been gained in the attempt. You will have learned a lot about institutional willingness.

Further reading

In addition to the websites mentioned above, here are some articles specific to news preservation:

- Reilly, Bernard F., Jr. "Knowledge Biodiversity" The Perilous Economics of World News Heritage Materials." ACRL Twelfth National Conference, Minneapolis, MN. April 7-10, 2005.
- "Kiss Your Assets Goodbye: Best Practices in Digital Archiving in the Publishing Industry," Seybold Report, August 23, 2007
- "Heart of Darkness: A Look Inside Aging JPEGs," Seybold Report, February 22, 2006. (With Alexander Mikhalevitch.)
- "Preserving History's First Draft," California Publisher (Vol. 94, No. 3), Fall 2005
- "No Pain, No Metadata," Seybold Report, June 22, 2005
- "Following the Trail of the Disappearing Data," Seybold Report, February 15, 2005
- "News That Moves: Accessioning Video for Newspaper Archives," The Moving Image: Journal of the Society of Moving Image Archivists, 4 (2) Fall, 2004

University Libraries
University of Missouri

Digitization Information Page

Local identifier	Jschool-backstory

Source information

Format	Book
Content type	Text / Text with images
Source ID	personal copy not added to collection
Notes	

Capture information

Date captured	April 2019
Scanner manufacturer	Fujitsu
Scanner model	fi-7460
Scanning system software	ScandAll Pro v. 2.1.5 Premium
Optical resolution	600 dpi
Color settings	8 bit grayscale
File types	tiff

Derivatives - Access copy

Compression	Tiff: LZW compression
Editing software	Adobe Photoshop CC
Resolution	600 dpi
Color	grayscale
File types	tiff
Notes	Images cleaned up, and darkened