WE ARE MPLP-ING
IN A DIGITAL WORLD

Confronting the Realities
of More Product, Less Process
in Digital Archives

Rachel Gattermeyer
BPE Conference | April 30, 2019
Outline

- MPLP
  - What it is?
  - How does it fit into digital archives?
- MPLP at University of Wyoming's American Heritage Center
  - What it looks like in action
- Strengths
- Weaknesses
- Discussion

Outline for session
Focus on born digital, not digitized
1. What is MPLP? This is kind of a contentious topic, but at its root, it is a tool/framework for a style of archival processing. MPLP comes out of a concern over growing backlogs and the consequential slowdown of getting records into the hands of users.

2. In the 2005 article, authors Mark Greene and Dennis Meissner cite an Association of Research Libraries study that found that of the many archives surveyed, they had an average of 1/3 of their holdings in a backlog (1998). This is a very stark, mind blowing amount.

3. Mark Greene and Dennis Meissner point to the struggle for traditional processing practices to keep up with the pace of new acquisitions and made a call for a change of archival processing to address the backlog. The goal of MPLP “aims to provide sufficient physical and intellectual access to collections for research to be possible”.

4. It introduces the topic of the Golden Minimum...of finding “good enough”. It’s asking the question of “what is the least we can do to get the job done in a way that is adequate to user needs, now and in the future?” The Golden Minimum is open to interpretation, which is a good thing. What is necessary at an archives with temperature and humidity control is different from an archive without that.
The Golden Minimum is also flexible/moving target
1. MPLP is also a time saver. It moves boxes from receiving to shelves quickly to hands of users relatively quickly...

2. It is a resource management tool...it asks archivists to weigh personnel, time, and budgetary resources to make decisions about what and how to process collections. When you have a large deluge of digital records waiting at your door and limited resources, you’re going to have to make hard decisions.

3. It’s not one-size fits all. MPLP is flexible and non-prescriptive. Each collection has unique needs, and minimum can be enough for one, but not another. Some collections need heavy processing and minimum isn’t always the answer.

4. Because the ultimate goal of MPLP is to provide users with access as soon as possible.
So how do we take a framework about processing collections and adapt it to digital preservation?

1. We broaden the idea of MPLP. We take the specific recommendations and situations unique to processing and move it up to an abstract and applicable level. Find the digital equivalent of not removing paperclips.
2. We keep the goal the same: Get digital records into the hands of researchers as soon as possible.
3. We need to understand MPLP as a tool of resource management, where we can use it to make decisions about where to allocate personnel, time, and money when preserving digital records. We can use MPLP to address our unique situational realities to meet the goals of digital preservation.
1. We ask: “What is ‘good enough’ to preserve digital records”? What does ‘good enough’ mean?
This is a challenging question. Digital files are weirdly fragile, and digital preservation is more complex and challenging than analog objects. Because there are so many types of digital records (text, image, audio, blockchain, databases, virtual reality), and within a certain type, there are hundreds of file formats. We can’t just stick them on digital shelves on servers or hard drives; digital records need additional care to ensure that they are usable in the future.
1A. So we look at standards - we look at the NDSA’s levels of preservation, 1B. we look at the TRAC audit requirements, 1C. we look at the OAIS reference model. We look at the standards built by the community to ensure that our records are available into the future.
2. Then, we look back at the previous point: figure out how to balance preservation needs with available resources to hit that golden minimum. This takes some time, so don’t worry if it doesn’t come right away.
3. And know that you will have to adjust the meaning of “adequate” when getting the job done as technology evolves and as new tech arrives into our archival collections.
MPLP at the American Heritage Center: Digital Archives

- Do no harm
- Our framework
  - Authenticity
  - Integrity
  - Interoperability
  - Reliability

So how does applying MPLP to the Digital Archives at the AHC look like? What do we find is “good enough”? How do we balance our resources with preservation needs?

1. It’s looked different over the years, but when we set our minimum, the guiding principle is: Do no harm.

2. We’ve done the research, took classes, looked around, and decided that for us, the essentials for a strong preservation program is based on the principles of: authenticity, integrity, interoperability, and reliability.

Within these essential elements, we determine what is absolutely necessary for digital preservation and what would be considered extra. We do do the extra, but only as necessary.
So what does this look like in action?
1. We don’t do disk images -> if we’re already doing 3 copies in 3 location, it ends up taking up redundant space. And by using bagger, we’re creating a bit-for-bit copy of the originals, so we aren’t leaving anything behind.
2. Metadata -> it’s pretty minimal; Metadata is definitely a critical part of digital preservation, but we’ve made decisions to pare it down to the essentials. We use FITS to grab technical metadata and use Data Accessioner to add Dublin core elements that hit on rights; provenance; admin info (e.g. identifiers, collection numbers) and a really brief description. We have the information in other places; why waste time to write it again to just have it in one location? If you take more than 5 minutes to add Dublin Core metadata, you’re taking too long.
MPLP at the American Heritage Center:
Digital Archives

- Migration
  - *It depends*
- Renaming
  - *Minimal*
- Duplicate renaming
  - *Rarely done*
- Arrangement
  - *It depends*
- “Good enough”

1. Migration -> we don’t always do this, but it’s always an option. One criteria we consider is what can a user open on their computer (or our computer in the reading room)? Can they open Word Perfect? Probably not. So we convert it to PDF/A. Do we migrate .doc? Yep. Do we migrate .docx? Nope. That’s going to hold on for a bit longer. And we have other criteria, of course, like file format obsolescence migration. Migration takes time, but it’s the minimum that we believe is necessary to open and read files in the future. And with the help of automated software programs, it can go a lot more quickly than doing the tasks manually.

2. Renaming -> This is pretty minimal. We look out for special characters, spaces…anything that might cause a computer to freak out and replace it with a dash or an underscore. We’re looking at renaming in a technical light, rather than a human-readable efficient light, so that the files will be openable and readable in the future. This can get really time-involved, but with the help of Bulk Rename Utility, it goes past pretty quickly.

3. Duplication -> this is rarely done. If we find software programs, we’ll remove them because of heightened security/virus risks and it’s outside of our collecting scope. And we’ll fill out a form that looks like this. We don’t look for duplicates or near duplicates…it’s time consuming and we haven’t found a software program that’ll do this well enough or is financially viable enough.
4. Arrangement -> we tend to leave it alone. Unless it’s a large external hard drive, then we will break it up into smaller, more manageable chunks and make a note that the smaller components are part of a larger whole. If there needs to be greater arrangement, the collection will be picked up by our Arrangement and Description department.

5. This is a lot of the “good enough” that we’ve found where we still hit high standards, are secure that digital records are preserved according to community standards, and maximize our time and resources.
All of what I’ve discussed touches a lot of the ways that MPLP shapes digital preservation at the AHC.

And after some reflection on the workflow and practices for the Digital Archives at the AHC, I’ve come away seeing some strong benefits to using this tool.

1. First, it makes my job manageable. As a lone arranger, my reality is that I am one person and my task is dauntingly huge. Preserving digital records is one part of my job – a big one, for sure – but by no means the only part of my work. I can hit that Golden Minimum and feel confident that the digital records are well taken care of. Is it perfect? No, but it’s good enough. I also know that if it’s needed, the Processing department can step in and do fuller processing as needed.

2. Second, users get their hands on the material fairly quickly. I’m a little behind analog and their 6-month goal for new collections, but I think that’s a pretty fast turn around.

3. Third, MPLP is a wonderful tool that helps me meet the realities of my situation: rate of new materials that need preservation, personnel, time, priorities of high risk or researcher requests.

4. Fourth, I still meet best practices. I still can confidently say that I am hitting authenticity, integrity, interoperability, and reliability.
Now, MPLP is by no means perfect. There are drawbacks to using this tool.
1. First, minimal metadata makes for poorer searchability on online databases/platform than more in-depth metadata description.
2. Second, it places a burden on users to delve into the digital folders. For the casual user, this might be a barrier. When searching for the Edith K.O. Diary collection, we don’t add metadata about the contents of the diaries, just main subject terms that we’ve gathered. This might not be a big deal for dedicated researcher, but for casual users, it might be too much for them to look deeper.
3. Third, if the framework is too minimal, there’s a real danger that I could do harm. Digital records are weirdly fragile, and we’re still figuring out how best to preserve files, what file formats are ideal, and how important look and feel is compared to content.
4. Fourth, not removing duplicative files adds space...so much space! At this point this isn’t a high priority, but it will be. It absolutely will be.
So now that I’ve talked about MPLP in my digital archives, I’d love to hear about your experiences with MPLP. I have some questions to get the conversations started, but you’re not limited to them.
Bibliography

Thank you!

Rachel Gattermeyer
Digital Archivist
University of Wyoming’s American Heritage Center
rgatterm@uwyo.edu
307-766-5614