Privacy in Public Archives: Managing Personally Identifiable Information in Special Collections

Archivists aim to make research and manuscripts accessible to the public. However, accessibility becomes tricky when donors or institutions enforce limitations. Sometimes limitations need to be enforced, especially when dealing with sensitive information such as personally identifiable information (PII), unpublished works, and student records. Redactions and restrictions may be necessary in these situations, but archives find this difficult to accomplish because of the size of the collections, results from previous accession practices, and the lack of staff and resources. The Special Collections department at Edith Garland Dupré Library, University of Louisiana at Lafayette is addressing this problem and has put forward methods to confront its backlog of PII while staying true to its accessibility mission. This article examines the challenges of handling PII in physical archival materials, the impact of More Product, Less Process (MPLP) on sensitive information, and how the Special Collections department, while using some MPLP processing methods, adopted a slightly more meticulous and efficient approach to protect privacy while still providing access.

There are two primary duties an archival institution provides its patrons: preservation and accessibility. While the preservation of archival materials allows for the conservation of the past, accessibility ensures patrons’ remembrance of that past. However, what happens when sensitive items not meant to be known to the public are made available with no checks? This can lead to ramifications ranging from private citizens expressing irritation to possible legal actions. No matter how serious the ramifications, archival institutions are placed in tricky situations, as there are no easy ways to maintain privacy yet to remain true to the mission of accessibility.

Of course, archives are not immune when it comes to privacy issues. Academic libraries in particular are caught in a bind as they manage not only personal collections but also university archives and records. Archives certainly need to respect donors’ rights for privacy, but it becomes complicated when donors do not place restrictions on items that should include them. This is especially the case when

© 2021 by Zachary G. Stein
these items contain personally identifiable information (PII). According to the U.S. Department of Labor, PII is defined this way:

Any representation of information that permits the identity of an individual to whom the information applies to be reasonably inferred by either direct or indirect means. Further, PII is defined as information: (i) that directly identifies an individual (e.g., name, address, social security number or other identifying number or code, telephone number, email address, etc.) or (ii) by which an agency intends to identify specific individuals in conjunction with other data elements, i.e., indirect identification. (These data elements may include a combination of gender, race, birth date, geographic indicator, and other descriptors). Additionally, information permitting the physical or online contacting of a specific individual is the same as personally identifiable information. This information can be maintained in either paper, electronic or other media.¹

In the case of archives, PII may include, but are not limited to, Social Security numbers, email addresses, credit card numbers, and bank account information. The release of PII not only invades individuals’ privacy but can also place their well-being at risk if the information falls into the wrong hands. PII can appear anywhere in the collection, requiring archivists to take careful notice.

In some cases, it may be relatively simple to detect PII and confidential records through processing. However, the task becomes overwhelming due to the size of the collections and the backlogs of unchecked materials. In addition, archives with small staff and reduced resources are dealt a heavy burden when managing an immense backlog of sensitive information. This leaves the archive’s mission of accessibility in check, as sensitive information exposing privacy needs to be protected. It may be next to impossible to tackle this problem all at once, but there are small steps archives can take to protect sensitive information and still serve researchers’ needs. The Special Collections department at Edith Garland Dupré Library (Dupré Library), University of Louisiana at Lafayette (UL Lafayette) exercised some of these steps to combat sensitive information; while there is still a long way to go, addressing and taking action on the issue has provided the department an opportunity to refine its policies and processes. The processes described in this article pertain to physical items, as Special Collections is still in the early stages of handling sensitive information in born-digital formats. These processes may be reflected in a future article.

**Literature Review**

To make informed decisions on combating PII in archives, it is important to understand the privacy laws that impact archives and how archivists have dealt with them. Archivists can collect a wide variety of materials, including personal papers, correspondence, financial records, and personnel files. Each of these materials may be subject to sensitive information that may be either illegal or embarrassing to divulge. While this article is mainly focused on PII, much of the literature reviewed here focuses on potentially embarrassing writings and passages, which is still important for understanding the difficult path archivists must tread to protect privacy.

Managing sensitive information has long been an issue in the archival field, though there are conflicting views as to how to appropriately move forward. Samuel Warren and Louis Brandeis’s 1890 *Harvard Law Review* article “The Right to Privacy” is often cited in the recent literature on this topic. Warren and Brandeis write about the “right to be let alone,”2 and that unpublished works may only be published by another party if the original creator provided consent.3 In fact, their view entails that the work belongs to the creator and the creator alone. It is up to the creator to decide if a work gets published,4 and, once it is, the privacy on that document is forfeit.5 While Warren and Brandeis’s argument may be broad and restrictive, it has helped identify the expectations and constitutions of privacy, laying the groundwork for current archival practices.

The Warren and Brandeis argument has also become something of a double-edged sword. Since one of the key missions of archives is access, restrictions can lead to serious ethical problems, especially if the restriction is made against the wishes of the donor.6 The Society of American Archivists (SAA) sustains a Code of Ethics as a guide for archivists. In the section under Access and Use, archivists are charged to “actively promote open and equitable access to records” while “striving to minimize restrictions and maximize ease of access.”7 However, the code also contains a section for Privacy, in which archivists must “place access restrictions on collections to ensure that privacy and confidentiality are maintained,” as long as the restriction is justified by law and transparency on the restrictions and the lengths of embargo.

---


periods are clearly documented.8 This particular section of the code makes clear the rules archivists must follow when juggling access and privacy, but it still does not answer the question of how to properly maintain sensitive information. Hodson accurately states that interpreting legal statutes on confidential archival records can get murky and that SAA’s Code of Ethics only provides general advice sans specific guidelines.9 Once again, archivists are left to their own devices to determine what is appropriate to make available.

There are ways, however, for archivists to make clear for donors what can be made available in writing. Deeds of gift are certainly helpful for mitigating what is and is not accessible. UL Lafayette’s deed of gift, for example, contains a section allowing donors to identify any limitations or restrictions they wish to place on aspects of the collection. In fact, Warren and Brandeis discussed contracts and the descriptions of protections within,10 a precursor of sorts to archival donor forms and listing limitations. The deed of gift ultimately decides the path archivists must take for collection maintenance, but even these forms can create privacy traps. Greene mentions that donors often do not check their collections for sensitive information, especially if donating on behalf of others.11 Hodson further explains that these third-party donors, even if descendants of the creators of a collection, cannot necessarily be considered reliable liaisons, as they may not have the right to speak and act on behalf of the creator regarding the disposition of the collection.12 The responsibility usually ends up falling onto the archivists, forcing them to decide on appropriate measures for balancing access and privacy. This can lead to much inconsistency due to the lack of clear guidelines,13 not to mention burdening the archivists with the size of the collections and individual judgments that may not be appropriate to make.14

The lack of guidelines certainly works against archivists, though actual laws put in place could also complicate matters. There are several laws that archivists commonly confront when handling sensitive information. The Family Educational Rights and Privacy Act (FERPA) is one particular law that university archives often face. This law aims to keep student records, such as grades and transcripts, private. Specifically, Part b, Paragraph 1 of FERPA states that “no funds shall be made available” to an educational institution if there is a “policy or practice” in place “per-

---

12. Hodson, “To Reveal or Conceal,” 159.
mitting the release of education records” or “personally identifiable information contained therein.” While FERPA is certainly well-meaning in protecting privacy, it complicates the situation when archives keep records in perpetuity. FERPA does not address the “archival life or historical value” of these records, so archivists need to decide for themselves if the records are worth keeping. Generally, records can be disposed of with records retention schedules, but it is difficult for archivists to make that decision. This is especially tricky if academic units wish to keep these records, even after the death of a student.

Modern archival practices have been used to acknowledge the issues regarding sensitive information, though it is still uncertain how effective they are. A key practice involves Mark Greene and Dennis Meissner’s influential More Product, Less Process (MPLP) framework. Basically, MPLP calls for minimal processing to “maximize the accessibility of collection materials to users.” Traditionally, archival processing that mostly embodies perfectionist practices takes too long to complete and can be overwhelming to archivists who need to process tens, if not hundreds, of feet of collections. MPLP is meant to be a more efficient and fluid method, where arrangement, description, and preservation are treated with the same level of focus. The overall policy for MPLP calls for unprocessed collections to be accessible with the exception of certain legal, physical, and valuable concerns.

While MPLP is effective, it involves a somewhat insouciant approach to dealing with sensitive information, despite the exception mentioned above. Greene and Meissner refer to the act of restricting “embarrassing material” as “absurd over-cautiousness” and that item-level security is a waste of archivists’ time. In this sense, they are specifically referring to materials such as private letters or correspondence that reveal gossipy secrets. There is an argument to be made that researchers have a right to see these kinds of materials, but MPLP does not appear to offer much of a solution when it comes to PII or legally confidential records. Van Ness brings this up in his criticism of MPLP, stating that minimal processing can lead to archivists overlooking sensitive information, which may consequently result in legal challenges. Cox, while not discussing sensitive information per se, did infer a related consequence to minimal pro-

---

processing. “out of scope material” and unrealized preservation issues may be missed.23 This is especially concerning if these out-of-scope materials contain sensitive information. Even in a follow-up article defending MPLP from critics, Meissner and Greene continue to explicate their lenient approach to privacy. They state that archivists “are not prescient” and should be cautious about removing items that may be perceived as presently problematic.24 Removing items, Meissner and Greene believe, can open archivists up to legal challenges if they set too high a standard for themselves when protecting privacy.25 Basically, if they make a single mistake, the burden falls on the archivists when, in reality, there should be a three-way collaboration among the donor, researcher, and archivist.26 This three-way collaboration is certainly an important way to combat sensitive information, though it provides a gray area when one of the parties is unknown or not present, a common occurrence for older collections.

There is no argument here that MPLP has merit. However, sensitive information and PII should not simply be dismissed as an unnecessary nuisance. As Cox explains, archivists need to care for the collections with the utmost high standards while still acknowledging their limits.27 This inspired Cox’s maximal processing model. Maximal processing is less about fast and easy accessibility and more about massaging and perfecting. Described in three major stages (predescription, description, and postdescription),28 maximal processing is a sort of modified version of MPLP, where accessibility is still the major driving force. Collections are minimally processed with available descriptions as a starting point but are then set aside for more detailed processing depending on such factors as “availability of external support, political considerations, requests from researchers, anticipated use, or potential for marketing.”29 This may appear as a compromise, but by balancing the act of thorough processing and accessibility, a more circumspect approach replaces the expeditious approach suggested in MPLP, making archivists more mindful of their holdings.

In addition to the processing methods mentioned, archivists can navigate through PII by institutional policy. Hodson states that policy can “protect the archivist or repository should anyone step forward and claim a privacy violation,”30 especially if it is clearly transparent and consistent. Meissner and

30. Hodson, “To Reveal or Conceal,” 162.
Greene agree that clear policies can save archivists much grief. In particular, they push for educating donors on problems resulting from sensitive information being accessioned and educating users on what to do when coming across sensitive information and the laws that protect them.31 These elements and more ended up helping the Special Collections department at Dupré Library figure out a plan for dealing with its backlog of PII and confidential records and apply it for future practices.

**Situation at UL Lafayette**

The Special Collections department (referred to as Special Collections from this point) at Dupré Library comprises the University Archives and Acadiana Manuscripts Collection (UAAMC) and the Louisiana Room. UAAMC, as the name suggests, contains the historical records of UL Lafayette and separate manuscript collections related to the Acadiana region. The University Archives section of UAAMC comprises more than 2,000 feet of materials. These materials come from various units, including the Office of the President, vice presidents’ offices, individual academic departments, athletics, administration and finance, and student organizations. The Acadiana Manuscripts Collection section of UAAMC comprises around 700 collections, all varying in size. These collections mostly contain the personal papers of people and organizations that make up the Acadiana community, such as the Jefferson Caffery Papers, Rice Millers’ Association Records, Edwin E. Willis Papers, and the Council for the Development of French in Louisiana (CODOFIL) Records. Sizes can range from a single folder to several hundred boxes. The Louisiana Room contains materials specifically affiliated with Louisiana, including published books, maps, genealogy, newspapers, and vertical files.

Academic and administrative units will often transfer materials for inclusion in the University Archives. While a great majority of items include announcements, lesson plans, photographs, and ephemera, these units will often send student records and personnel files as well. When transferred to the archives, these files can arrive in large loads. The current practice for collections such as the Office of the President papers involves leaving the papers in their original folders (assuming the folders are not damaged), placing them in alphabetical order by year, and cataloging the inventory into the finding aid. This practice pretty much follows MPLP by making these papers accessible as quickly as possible. If the collections are small to medium size (around 10 boxes), then the papers get transferred to acid-free folders with any metal fasteners removed.

Perusing every processed collection to search for PII is not an option. Special Collections is made up of only three faculty (Head of Special Collections, Reference Archivist/Louisiana Room Librarian, and Digitization Archivist), three full-time staff (Archives Assistant, processing assistant, and Louisiana Room assistant), and part-time student workers and scholarship students who may only work up to 12 hours a week. With more than 2,000 feet of University Archives and 700 collections of personal papers to sift through, it is counterproductive to actively search for PII in already processed materials. However, it does become a necessity when PII is identified or cited in finding aids. A single manuscript collection, the working files of a former university employee who had passed away a few years earlier, acted as a catalyst to push for change. While the majority of the papers were benign, the collection contained personnel files, including letters referring to active employees in a perceived negative context. This required an aggressive screening of the collection to weed out the confidential materials meant for administrative use only.

Another catalyst for change in privacy screening involved the vertical files in the Louisiana Room. Vertical files are simply newspaper clippings organized by categories related to the state as a whole (i.e., churches, food, hurricanes). The clippings were attached to scrap paper and placed in folders based on their subjects. These files amount to possibly hundreds of thousands of individual clippings. Unfortunately, a great number of these articles are attached to scrap paper containing PII such as social security numbers, student names, addresses, and even course grades. Not only do these contain private information of people who are most likely still alive, but these could also act as a FERPA violation. With all of this in mind, Special Collections determined it needed to make combating this PII backlog a high priority.

**Methods for Privacy Protection**

Despite the large backlog and limited staff and resources, Special Collections has actively changed its practices to accommodate for privacy issues. While still using basic MPLP for collections, the staff has adopted methods for screening PII, one partially inspired by Cox’s maximal processing model. The succeeding sections detail the actions being taken in each of the major sections of Special Collections.

**Acadiana Manuscripts Collection**

When it comes to the archival collections, new policies have been put in place. A whole section based on privacy has been added to both the Special Collections and

---

Reading Room policies. The language informs patrons of their responsibilities and the consequences of finding sensitive information. The policy also specifies Special Collections responsibilities, such as reviewing materials prior to access and removing information from the collection if necessary. While it is Special Collections’ duty to screen for this information, it is also the duty of the patron to identify any anomalies he or she comes across. According to the policy, patrons must refrain from making records or notes of sensitive information and notify Special Collections staff immediately. Much like copyright, the responsibility for privacy infringement falls on patrons.

When PII is found in collections, it must be dealt with in a way consistent with archival practices. Taking a page from Cox’s maximal processing model, the papers in a collection are briefly screened for PII. If a folder contains 20 to 50 papers full of PII, a note is made in the finding aid, and the folder is restricted until it can be revisited at a later date. While it may seem antithetical to special collections’ missions, restriction is only meant to be a temporary response. This brings to mind the postdescription step of Cox’s maximal processing model. The idea here is to revisit a collection after it has been processed and make changes that were not acknowledged in the original processing.33 A similar method needs to be used for restricted collections. This could include redacting, reprocessing, or deaccessioning. Time and resources hinder the ability to revisit these collections, though Special Collections has been able to go through certain collections and remove items such as canceled checks. While not extensive, this acts as a small step forward in dealing with restricted material.

If a manageable amount of PII is identified, the information is redacted during processing. While redacting PII is not a new practice, Special Collections decided to take a somewhat different kind of approach. Basing the approach on common digitization practices, the Head of Special Collections created guidelines that used different copies of items, since potential research value is difficult to decipher. If PII is found on an item, that original item (or the master copy) is photocopied in color. Private data in this mezzanine copy is redacted with a black Sharpie pen or white-out. Once the redaction is complete, the mezzanine copy itself is photocopied and then shredded. The second photocopy, the access copy, replaces the original item in the collection since the redactions cannot be removed. The original document is placed in a restricted folder within the collection’s control file, which contains printed finding aids, donor forms, correspondence, or any materials relevant to the collection. Control folders are kept in filing cabinets inaccessible to anyone except Special Collections staff. This method is labor intensive, but it is an efficient way

to conserve an archival item, protect the sensitive information it contains, and still make the document available to patrons all at the same time.

Cox states that the goal of the full processing stage of maximal processing is to “maximize” archival processing “with respect to appraisal, arrangement, and description, while always keeping a clear eye on costs.” Special Collections redaction method requires an adaptation of this kind of processing to maintain focus on appraising the contents of folders with PII. It is up to the processor to decide how much PII is manageable, as time always plays a factor in processing. Nevertheless, this method is necessary for ensuring the protection of PII while also allowing patrons to view the document, even if it is a photocopy.

**University Archives**

The redaction method outlined above works well for individual manuscript collections, but it is trickier when dealing with the University Archives. Certain collections are filled to the brim with student and faculty records, and even if they did not include PII, student records are still protected under FERPA and it is highly unlikely that past students can be contacted to give permission for access. Some departments have given consent to have records disposed of, but that in itself causes other problems. While some of these records may not seem to be of tremendous importance, they still document the history of the institution. If UAAMC were allowed to dispose of the records, there are still the protocols of records management to deal with, which include working with the Louisiana State Archives and Secretary of State office to draft and approve records retention schedules. The Head of Special Collections acts as the records manager for the entire university, which means he is the keeper of all records retention schedules for all university departments. If a department needs to dispose of records, it will need to have an up-to-date schedule on file. The department, through the records manager, also needs to send a disposal request to the Secretary of State’s office to receive approval. It ends up becoming an incredibly long and arduous process.

To help mitigate the act of screening PII in University Archives records, the Head of Special Collections initially proposed that patrons be required to make appointments at least two days in advance. Appointments are a common practice in archives, as they give archivists time to search for the requested collection and have it available right when the patron needs it. Furthermore, the two-day window would allow for Special Collections staff to comb through the requested boxes and folders to isolate any confidential materials.

---

34. Cox, “Maximal Processing.” 144.
Initially, the idea of appointments, while acknowledged as a well-intentioned and thorough way to eliminate risk, was rejected due to the potential strain they could place on the patrons and the staff. As part of an academic library, UAAMC is expected to be open to the public. Patrons frequently come to the archives unannounced, and it would be a nuisance to force them to make appointments and come back another time. This is especially problematic for students who may have important projects with looming deadlines.

Despite the above concerns, Special Collections did end up adopting the process in response to the COVID-19 pandemic, and the reticence subsided. In fact, the concerns ended up not posing much of a problem, as patrons have accepted the appointment process and the staff is able to gather collections in an efficient way. Patrons have the ability to schedule appointments by submitting an online form, emailing, or calling the Special Collections reference desk. The appointments have prepared staff for incoming requests and even give them time to review the finding aid for any possible restricted items. It remains to be seen whether appointments will continue after the pandemic has passed; but, due to the efficiency they have provided, it is likely Special Collections will continue to require scheduled appointments.

**Louisiana Room**

The central items that needed attention for PII screening were the vertical files in the Louisiana Room. In the fall of 2017, Special Collections hired a new Reference Archivist/Louisiana Room Librarian (referred to as Reference Archivist from this point). Part of the job description included maintenance of the vertical files. Noting the problems mentioned in the previous section and realizing that the current practices were not sustainable, the Reference Archivist organized a long-term project to digitize the vertical files. This would simultaneously allow for better preservation of the clippings and elimination of the PII stuck on the paper. Each article is scanned at around 300dpi in PDF form; student aides conduct much of the scanning, and the Reference Archivist is in charge of quality control. Once the scans are approved, the physical articles are disposed. The Reference Archivist and Louisiana Room assistant take care to flag files with PII, which are specifically put aside for proper shredding.

For the most part, digitizing the physical items removes the risk of exposing the sensitive information on the back of the files. Currently, if sensitive information bleeds through the paper, the scan is not used and deleted. The tricky part is making these files available, since copyright would restrict online accessibility for newspaper clippings. The digitized copies are currently located in a folder on a hard drive on one of the two Special Collections Reading Room computers; the
folder is shared on the second computer. The Reference Archivist and Library Information Technology (IT) Systems Specialist have access to this folder from their work computers, which allows them to continue adding and editing files. The articles themselves are only accessible through unique logins for the Reading Room computers. Patrons can navigate to the library website and the vertical files page. Outside Special Collections, the webpage only provides an index of the subjects and categories. On the Special Collections computers, patrons can click these subjects, which open PDFs of the articles via file paths. This method is fairly consistent with legal advice Meissner and Greene sought out regarding a lack of distinction between making records available in a reading room and making them available online, though Special Collections is still protecting itself from potential copyright infringement complaints from these newspaper outlets.

Of course, this project is expected to last a very long time. As of this writing, Special Collections is three years into this project, having only made it through the H entries. In addition to the long process, Special Collections also needs to be wary of the storage space required. The files themselves may be small, but the large number of files accrued can use up much digital space, which may be needed for other digital projects. Therefore, the Library needs to be conscious of budget necessities.

In the meantime, patrons request vertical files, and it is not certain how many of them contain PII. Originally, Special Collections staff simply went through each folder and pulled red flags just before handing off to patrons. However, this has proven to be a tiring process. Some patrons may request to look at more than 50 folders at one time, which can become too exhausting for a small staff. Another suggested idea was to restrict all of the vertical files until the digitization project was completed. While this certainly takes care of protecting the sensitive information, it also undermines the whole purpose of archives and libraries: making information accessible. Vertical files tend to be popular items for patrons, especially students, so making them unavailable would turn these patrons away. Ultimately, a compromise of sorts was suggested: if a patron requests vertical files that have not been vetted, the Special Collections staff will put the folders on hold. These folders are prioritized for digitization, and the patron is notified when they are available. Additionally, scheduling appointments have made screening vertical files more manageable, as staff has more time to look through the vertical files and pull red flags.

**Conclusion**

Because of the backlog of sensitive information in archival collections, Special Collections staff must take a more careful approach when screening for PII. While still

---

using modified versions of MPLP to process, the screening of PII requires a more observant method to ensure that sensitive information is not being glossed over. The redaction method adopted by Special Collections at Dupré Library is a practical way to ensure that PII is being protected without sacrificing the original item. Cox’s maximal processing has also helped put screening for PII into perspective, especially when using initial screenings, detailed appraisal, and revisiting restricted items. Redacting is by no means a perfect method, however, as it does take time to complete. It can also become difficult if the photocopies are not produced the same as the original. Nevertheless, the redaction method does help maintain the integrity of the collections, while simultaneously being mindful of privacy.

The methods described in this article act as a first step in combating PII. Importantly, they help Special Collections acknowledge this serious subject and take action for a more responsible and secure repository. At the same time, Special Collections is also transparent with patrons through policy and staying true to the mission of accessibility. These policies give staff a blueprint for how to navigate their archival materials and how patrons can help if they come across PII. Items with sensitive information will always pose problems for archival institutions, but these methods can help soften the burden and make archivists more aware of what they hold. With a plan to protect privacy, archivists can perhaps become more comfortable with their collections and continue to exercise their duties for conserving and disseminating their holdings.