Discovering the Present, Preserving the Past: the Development of a Digital Archive at the University of Maryland

NA LIN, PATRICIA G. HINEGARDNER

University of Maryland Health Sciences and Human Services Library, Baltimore, Maryland, USA

ABSTRACT

Collecting, preserving and distributing intellectual works has always been an important mission of libraries. The digital age has presented new challenges including maintaining active links and providing formats that can be adapted over time. It has also brought new opportunities for sharing information globally and providing access to "hidden" collections. Providing access to local resources both current and historic through repositories/archives is a service many libraries are providing. In 2010, the Health Sciences and Human Services Library (HS/HSL) at the University of Maryland (UM) began the implementation of the UM Digital Archive. This article describes the experiences and lessons learned through this endeavor.

KEYWORDS

Institutional Repository, Digital Archives, Dspace, Metadata Management, Digitization, Digital Permission

INTRODUCTION

Collecting, preserving and distributing intellectual works has always been an important mission of libraries. The digital age has presented new challenges and new opportunities. It is important that the enormous amount of information being produced in electronic format not be lost if the format in which it resides is replaced with the next latest and greatest technical advance. Providing an environment that will house formats

that can be adapted over time and providing a persistent URL so that access is not lost due to an infrastructure change are of vital importance. A challenge of equal importance is the need to digitize and provide access to print material, especially the grey literature (materials that cannot be found easily through conventional systems of publication). The dissemination of these works via the web provides access to otherwise hidden materials and enhances information sharing across the scientific community. Materials that provide insight into the history of a university are also important to preserve and distribute so that future generations can appreciate and build on a university's legacy. The Health Sciences and Human Services Library (HS/HSL) at the University of Maryland (UM) is meeting these challenges with the development of the UM Digital Archive.

BACKGROUND

The HS/HSL is located on the Baltimore campus of the University of Maryland. It serves the Schools of Dentistry, Medicine, Nursing, Pharmacy, and Social Work as well as the Graduate School. As the founding campus (1807) of the University System of Maryland, the campus has a rich history and a dynamic environment for education, research and clinical care. The mission of the Archive is to collect, preserve, and distribute via the web the academic works of this population.

The development of the Archive began with the writing of a white paper on Institutional Repositories. The Executive Director of the HS/HSL established the Repository Information Project Task Group (RITPG) in 2008 to explore the feasibility of developing a repository. The group was lead by the Associate Director for Resources and consisted of 4 faculty librarians and 3 IT staff. The group reviewed the current literature, talked to experts in the field, and discussed the ramifications for the HS/HSL and the University. This research culminated in a white paper which recommended the creation of a repository for the University of Maryland in Baltimore. ¹

The HS/HSL Executive Director made the decision to commit staff and resources to the development of a digital place to collect, preserve and distribute the academic works of the university with an initial focus on grey literature including research reports, annual reports, newsletters, white papers, meeting presentations and posters. Another important aspect of the project was to capture the historical record of the campus by digitizing collections from the Historical and Special Collections of the HS/HSL. Instead of referring to the product as a repository, it was decided to use the term archive.

The word archive seemed to put the project into a more understandable context for the campus community. It also reflected the initial focus of the project. The Repository Information Project Task Group was now given the responsibility to implement the UM Digital Archive https://archive.hshsl.umaryland.edu.

During the investigation phase of the project, DSpace (http://www.dspace.org/) was chosen to be the platform for the Digital Archive. DSpace is a stable, well established open-source software product with a large community of users and developers worldwide. It is free, and allows for customization to achieve desired branding and functionalities. The HS/HSL IT department felt confident to deploy the product in-house.

New members were added to the original group including the recently hired Head of Resource Sharing and the Digital Archive and a programmer. The group recognized that three fundamental elements needed to be addressed: infrastructure, content and access, all closely interrelated. To that end, the group was divided into three corresponding teams: Technology, Content, and Metadata Management. The Technology team was tasked to build the DSpace environment and customize the Archive site; the Content team was to prioritize and recruit content; and the Metadata Management team was to determine the best practice for metadata implementation to optimize content access. Each team was also responsible for establishing policies and guidelines relating to their areas. This included copyright and digital permission, content management, submission, digital formats, metadata requirements, and withdraw procedures. The project leaders, the Associate Director and the Head of the Digital

Archive sat on each team as contributing members and also to coordinate the efforts among the teams.

DELIVERABLES AND TIMELINE

To realize the vision of the Digital Archive, the deliverables of the project had to align with its goals and objectives, be specific, achievable, and measurable, and needed to be driven by a rigorous timeline to ensure on time completion. The deliverables were:

- archive built using DSpace on Redhat Linux
- user interface customized and tested to optimize browsing, searching and site navigation
- policy and guidelines
- submission protocols including content curation process
- content prioritized, populated and curated
- help website created
- RSS enabled
- operational process created: resource and procedures
- marketing plan developed
- a plan for systematic content recruitment created

The project timeline extended from May 2010 to May 2011 when the launch was scheduled, during which milestones were set, checked and adjusted as needed to keep track of progress. See Figure 1 for the project timeline.

Mapped to these deliverables, each team member's responsibilities were defined and documented corresponding to various stages in the timeline. The teams worked in parallel to fulfill their responsibilities, convening biweekly to report back progress, discuss issues, and reach decisions for next steps. Meeting minutes then followed to reaffirm the action items and identify the team or individual responsible for completing each task.

A SharePoint site was created to house all the documents generated from the project including meeting minutes, policy and procedures, test results, and project schedules. Paper permission forms received from content providers were digitized and stored in this central location. A wiki site was created for team members to submit customization requests, which were color-coded once completed. This documentation will be followed for re-customization in future DSpace upgrades.

TECHNOLOGY

The Technology team consisted of four IT staff providing network engineering, web development and software programming. The team was responsible for infrastructure development and website design. The network engineer went through due diligence to get the system ready. Initially, the Dspace v1.6 XML User Interface (UI) environment was deployed. However, since the programmer position was still in the recruitment process, it was decided to switch to JSPUI so that the web developers could begin customizing the look and feel of the site.

To make the site both appealing and functional, the task group reviewed many institutional repository sites and sought feedback from users. Once the programmer was hired, the team started backend customization. The guiding principle was to customize as little as possible yet achieve maximum usability. DSpace as an open source software will continue to evolve with great agility to allow broader configurability in lieu of extensive customization, which poses enormous challenges of re-customization during each version upgrade. In fact, DuraSpace, a not-for-profit for DSpace solutions, took part in the Google Summer of Code < http://code.google.com/soc/> that resulted in a number of Dspace enhancements on the submission process and JSPUI design.

Using the principle of minimal customization, only features critical to usability such as search and display were customized. Nice-to-have features were evaluated and tried to measure the importance of their added value versus the effort to implement. The focus was on enhancing end user experiences while accommodating Archive staff needs for

operational efficiency. Some good concepts were eventually put on the backburner because the return did not justify the amount of work required. It is hoped that the technology will catch up to enable the implementation of these with less effort. An example of this was the attempt to dynamically populate appropriate digital permission form(s) based on the type of copyright owner, individual or organization, during the submission workflow. For now the solution is to put both in one form since the submission staff would be able to determine whether to move forward with the submission according to the content and permission given.

Based on user feedback and the team's experience, a number of customizations made it to the final implementation:

- A banner was added to the homepage to highlight featured collections as Figure 2 illustrates.
- 2. An auto-complete feature was applied to Basic Search
- Advanced Search was revamped for more effective searching such as providing the ability to select multiple material Types
- 4. The attachment (bitstream) was repositioned at the top of the page as Figure 3 illustrates. This placement provides quicker access to an item.
- 5. An URI field was incorporated into the display to link to externally hosted content that renders in a new window keeping the Archive site in sight
- A workflow step option was added to the batch import function so that
 metadata librarians can decide whether or not to go through the metadata
 approval process.
- 7. SFX link resolver (FindIt) was integrated for full text linking to published articles and electronic theses and dissertations (ETDs) at ProQuest.

As the Archive site developed, it was decided to move out of the DSpace environment to create the Archive Help site using Wordpress. This approach allowed for complete freedom in designing the site. Wordpress was easy to use so each team could populate

the site with their policies, guidelines, FAQs and glossaries. Google Analytics was set up to capture usage, in addition to DSpace's native statistical function.

CONTENT

The task group realized that to get all schools of the campus on board was essential to the success of the Archive, but it would require an extensive campaign. With limited resources at the implementation stage of the Archive, this was not feasible. Instead, the team decided to start small, working with several schools with which the Library had close relationships and digging into the Historical and Special Collections held by the Library. After several brainstorming sessions, with the consultation of the library executive director, the team composed a list of prioritized content to be added in the initial phase.

Through collaboration with the liaison for School of Social Work, a demonstration of the Archive was conducted for the Associate Dean for Research. Although the version of the Archive at that stage was crude, she immediately saw the benefits that open access to her team's research reports could bring to her funded programs. She granted permission to distribute the reports via the Archive, and selected a Creative Commons license for these reports.

At that time, the School of Pharmacy was celebrating a new building addition. The event signified a milestone for the School and the materials developed for the celebration evidenced a historical moment for the campus and needed to be preserved in the Archive. Working with the liaison librarian, the Director of Communications of the School was contacted. She sent an email agreeing to send the materials. The team capitalized on this agreement and requested a face-to-face meeting with the director to promote the benefits that the Archive could bring to the School, and to explore the School's potential content. By the time the meeting was over, the School's content

including course catalogs, yearbooks, annual reports, newsletters and more had been secured.

Theses/dissertations and the Archive are, not surprisingly, best suited for each other. A decision was made to include only electronic versions of theses and dissertations (ETDs) from ProQuest. The Library and Graduate School share a reporting structure making it easy to initiate the ETD submission process.

Simultaneously, the team was picking the "low hanging fruits" from the historical collections at the Library. With campus history dating back to the 19th century, the Library possesses a wealth of literature and images about the campus and schools: 19th century theses of the School of Medicine, historical art collections in dentistry, manuscripts by a past professor who pioneered and influenced state legislature on physical therapy education, oral history of a renowned professor in the School of Medicine, biographies of historical figures of the campus, and 19th and early 20th century publications by the Schools.

Once the initial content lined up, materials from the School of Social Work, the School of Pharmacy, the Graduate School and the Library, a content structure in the Archive was created.

METADATA MANAGEMENT

The charge for the Metadata Management team was to develop and implement metadata strategies for discovery, access, and management of content in the Archive, and to formulate a curation process that would be applied in future operations.

It was first decided to use and maintain DSpace's native Dublin Core schema. The team examined and selected the elements to describe various types of content. While author provided keywords were included, for enhanced findability, the team also decided to add controlled vocabulary from the Medical Subject Headings (MeSH) and Library of Congress Subject Headings (LCSH) as a way of allowing the user to find similar works

together in a search. For the same reason, additional metadata elements, displayed in Table 1, were added to the central index. Individual and entity names are normalized against the campus directory.

Metadata presentation is of vital importance. Metadata should be as comprehensive as possible to adequately describe the content, yet cannot be so overwhelming that it drives users away. Once metadata is selected, how should it be ordered on the page? What are the critical pieces that should stay on the top section of the page so that scrolling is not necessary? How should it be labeled to be meaningful to users? It was important to get user feedback to these questions. Usability testing included feedback from the Library Executive Director, faculty and staff, and an "information location" session with staff from various backgrounds who had never seen the Archive. Based on this information and after reviewing several mock ups, a final page display, as illustrated in Figure 3, was completed.

To create operational procedures, the team streamlined content submission and approval processes by customizing the workflow. For example, adding a metadata curation step if ETDs are batch imported, and skipping the step if the content has been pre-curated, as in the case of Lyrasis digitization (see Digitization). A metadata group of two metadata librarians was created in DSpace to execute the submission approval process. Metadata librarians started training support staff in submitting content and created templates for the submitters to follow when necessary. Through the department heads, communication methods were established between the submitters and metadata group to ensure efficiency.

With the system ready, content selected and the curation process in place, the task group was ready to enter content, only to face a set of new challenges of staff and financial resources. At the time of the Archive implementation, staff resources for content submission, the budget, and expertise to convert large collections from print to electronic versions were all limited. In-house scanning was limited to small pieces on a

flatbed scanner. For large amount of content that was being targeted, a better way was needed to handle it.

DIGITIZATION

Converting campus historical publications dating as far back as the early 19th century from print to digital would require an investment in more personnel, equipment, and training. This was not realistic given budget constraints. In order to meet these demands, a search began for a strategic partnership with vendors and/or organizations that could more economically handle the situation.

After evaluating several vendors based on technical capability and pricing, it was decided to join the Mass Digitization Collaborative at Lyrasis. Internet Archive http://www.archive.org, a non-profit internet company who partners with Lyrasis in this initiative, provides digitization and free hosting of digitized content.

Several awards were received from the National Library of Medicine, National Network of Libraries of Medicine (NN/LM) Southeastern/Atlantic Region (SE/A). The amount was limited so collections were prioritized based on perceived impact, costs, physical conditions, copyright considerations, etc. A process was then established to ensure the quality of and access to the digitized content:

- 1. Metadata creation The Metadata Management team led a group of staff to review and assign metadata for each piece of material, placing the information in a spreadsheet.
- 2. Digitization Upon completion of the metadata, materials along with the spreadsheets of metadata were shipped to the designated scanning center. The turnaround time between the shipment and digitization completion was about two to four weeks.

- 3. Quality control Although quality control was embedded within the digitization process, an in-house quality check was performed after the digitization. The Metadata Management team verified the metadata and noted corrections to be made. A technician went through a number of volumes to ensure all pages were scanned according to specs, and then spot checked the rest.
- 4. Final publishing The Internet Archive uploaded the changes made through the quality control for the last time before publishing.
- 5. Link to the Internet Archive Metadata librarians tweaked the Lyrasis metadata format to make it conform to the Archive format. The data were then imported into the Archive to allow for search and browse as well as linking to the full text at the Internet Archive site.

To date, historical yearbooks, magazines, manuscripts and course catalogs of over 200,000 pages had been or are queued to be digitized. Digitized materials are freely accessible through the Archive and Internet Archive. Joining Lyrasis' Mass Digitization Collaborative program allowed us to take advantage of its low cost, digitization expertise and free hosting at the Internet Archive. It has yielded greater value for the money and productivity than if it had been done in-house.

DIGITAL PERMISSION

To allow the Archive to distribute and preserve scholarly works in perpetuity, the authors have to declare their rights ownership and grant the Archive permission. The general practice for enforcing these requirements is to create a Non-Exclusive Distribution and Preservation License (License) for authors.

To form policies on digital permission for the content to be submitted in the Archive, a review of the literature was conducted as well as the review of practices of many institutional repositories. It was found that there were standard elements, but also situations unique to the HS/HSL that warranted a close assessment of each element and

a tailored policy as a result. From the beginning of the project, the task group worked closely with the University Counsel in drafting the License to make it comprehensive yet concise and easy to understand. When the License was finally approved, it became a part of the submission workflow in the Archive.

However, as schools were approached for content, the suitability of the License was questioned for school owned publications since the language in it was geared toward individual authors. To address the need, the University Counsel created a separate document, the UM Digital Archive Contribution of Work Agreement - for materials in which UMB controls the copyright (Agreement), and it was then incorporated in the Archive submission workflow. With this Agreement, schools provided permission for archiving and distributing their publications including course catalogs, newsletters, yearbooks, annual reports, etc.

Another form was developed as a result of entering electronic theses and dissertations (ETDs) from the Graduate School. The School worked with the University Counsel and introduced a form, the Thesis and Dissertation Distribution Authorization that specifies embargo periods. The document is embedded within the ETD process at the Graduate School. Signed forms are digitized and stored together with other permissions in the Archive documentation space on SharePoint.

Another situation arose with the content on the Internet Archive. Content digitized through the Lyrasis' Mass Digitization Collaborative program is hosted on the Internet Archive site. The advantage of this arrangement is that it saves in-house storage. However, the schools had signed the original Agreement which stated that the archived content would be stored in-house at the University. The University Counsel advised that an explanation be made to the schools about the rationale of this arrangement and get their consent. Some schools gave permission immediately while others went through their own due process before eventually granting permission to proceed. The original Agreement was modified to incorporate the external hosting clause.

Digital permission needs to be specific but broad enough to be applicable to different circumstances. While it's impossible to predict all scenarios, it is important to identify issues and act upon them. However, in times when things are unclear, a good relationship with legal counsel is critical.

Although the Archive concentrates on grey literature, published scholarly works are accepted as well. For those whose copyright is owned by a publisher, license restrictions are checked on the SHERPA/RoMEO site http://www.sherpa.ac.uk/romeo as well as working directly with the publishers, who usually grant the permission for Archive distribution.

Despite all the anticipated scenarios and encountered circumstances, there are still times when a copyright owner of an unpublished work can not be located, or the copyright ownership can not be verified. A take-down policy was implemented and publicized on the Archive. With that, a calculated risk was taken to include "orphan" works in the Archive, but only if they are hosted in-house. If the copyright owners are identified and contest inclusion, the works can be taken down immediately.

STAFFING

The Archive was conceived as a new service provided by the Library, however, without a budget and additional staff. It was to be instituted in the Resource Sharing Department within the Resource division. The Archive service was modeled to consist of both an end user self-service submission option and a full service submission option provided by Library staff. It was realized that at least for the introductory stage, in order to market it and get faculty on board, offering the full service option was needed.

This required that Resource Sharing staff, in addition to their interlibrary loan (ILL) and documents delivery responsibilities, take on the duties of the Archive service that include copyright clearance, submission, digitization, withdrawal and collection of usage metrics. Since the volume of ILL requests had been decreasing over the years, this was

an opportunity to redefine staff roles and responsibilities. There was a re-allocation of staff resources, the building of new skill sets, and the creation of procedures. From December 2010 till the Archive launch in May 2011, Resource Sharing staff went through training and experimented with different approaches to build a process that would be able to sustain the service in the long run.

The initial approach was to have every staff member submit content. Staff received a crash course on scanning, using DSpace, and entering appropriate metadata. This approach turned out to be counterproductive due to staff interruptions to handle ILL duties. It was decided to have one or two staff members dedicate part time to the Archive service. Three quarters of one FTE time was allocated to the Archive.

As content piled up and more digitization projects emerged with deadlines mandated by the funding agency, the staff allocated for the Archive was challenged to get all the work done in time. It became clear that there was a need to take a holistic approach to strategically develop a pool of expertise and resources, both in the Division and at the Library level to sustain the Archive operation. Staff from the Collection Management department was assigned to work on content submission and digitization for materials not suitable for the Lyrasis digitization program. Staff from the Services Division were recruited to assist in metadata curation for historical collections. This collaboration not only enabled us to meet the digitization and Archive launch schedules, but also functioned as a test for establishing operational processes post launch.

Staff training will continue to be needed so that they become facile in all aspects of the Archive service. The goal is to empower staff to not just follow the procedures, but to make decisions throughout the process with the skills and the understanding of the practice. There is also the potential for staff to assume some of the metadata management responsibilities to ease bottlenecks during content curation.

PROMOTION AND CONTENT RECRUITMENT

The official launch of the Digital Archive which included a reception and demonstration

was held on May 4, 2011 at the Health Sciences and Human Services Library. The 30 people who attended were enthusiastic about the project and contacts were made that proved fruitful for obtaining content.

Several approaches have been used to promote and recruit content. Presentations have been made to campus-wide committees including the Communications Council and the Research Council as well as faculty meetings at various schools.

Partnering with library liaisons has also proven successful. There is a library liaison for each school on campus. The liaisons have promoted the Archive through their access to school email lists, sending e-flyers about the Archive. They also promote it through their interactions with faculty and have provided names of faculty who they thought would be interested in contributing to the Archive.

Identifying possible contributors through the review of campus/school websites and news outlets has also proved fruitful. Through this method the director of a key research center on campus was contacted and is now a contributor to the Archive. Direct contact with each school's communications office has been the way that newsletters, annual reports, event materials, etc have been obtained. Contacting the Graduate School started a conversation that eventually led to the Archive being included as part of the dissertation/thesis process. These are now sent automatically to the Archive.

Content from the HS/HSL Historical and Special Collections was easy to obtain but had the added step of needing to be digitized. Some items in the collection are very fragile and require extra care in handling. The digitization of collections of dental illustrations as well as 19th century School of Medicine theses has given the Historical and Special Collections new exposure. The Archive is helping to promote many hidden collections.

The Archive is promoted in a variety of ways including posters and digital displays in each school. A column, *Digital Archive Highlights*, in the HS/HSL's newsletter, *Connective Issues*, brings individual collections or items to the readerships' attention. A brochure is in the process of being finalized for a mailing to faculty and staff in the fall 2012. The strategy is to keep the Archive in constant view using various methods.

Content recruitment is an ongoing challenge. People are enthusiastic about having their works in the Archive but it requires work on their part to locate the content they want to contribute. Often it requires tactful repeat reminders before items are obtained. Another difficulty is the lack of knowledge as to who owns the copyright of an item. In the case of school publications, it often took several tries before the person was found who could sign the permission agreement. Content recruitment requires patience, creativity and tenacity.

CONCLUSION

As the Archive continues to evolve, challenges continue with a new DSpace version next year, the need for more staffing and training, the need for funding for digitization projects, and the never ending task of promotion and content recruitment. All these challenges are seen as opportunities for better content discovery, more staff development, creative funding and the showcasing of a project that will benefit the University now and in the future.

References

1. Gresehover, Beverly; Behles, Richard; Douglas, C. Steven; Pinkas, Maria M.; Hinegardner, Patricia; Pinho, Thom An Institutional Repository for the UMB Campus: a white paper (2010) http://hdl.handle.net/10713/7

Figure 1 Project Timeline

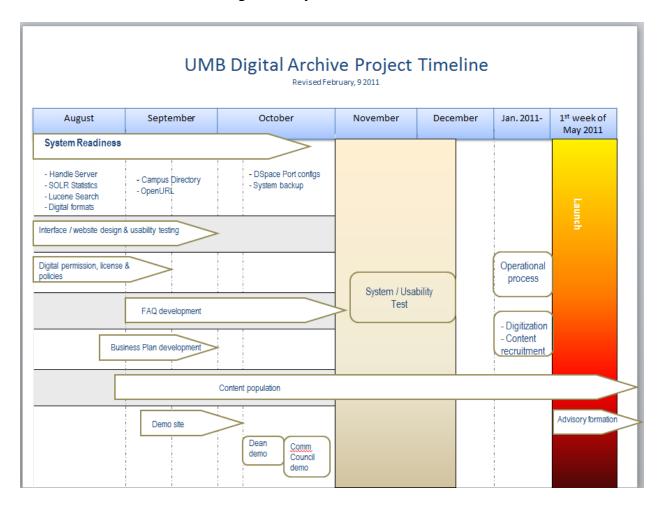


Figure 2 Sliding Banner to show featured collections

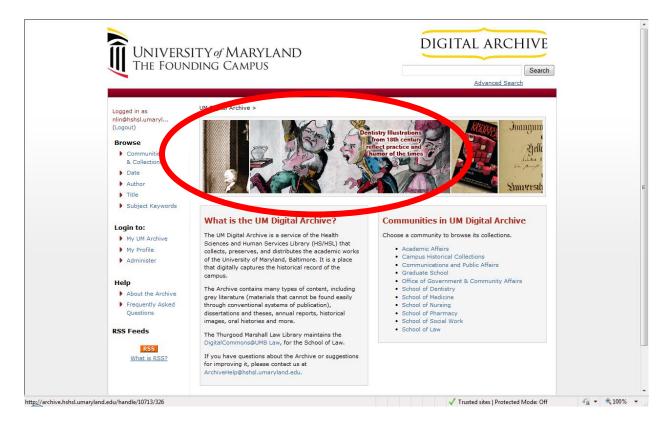


Figure 3 Metadata Display & Bitstream Position

UM Digital Archive > Graduate School > Theses and Dissertations Graduate School >

Please use this identifier to cite or link to this item: http://hdl.handle.net/10713/609

Title: Basolateral Amygdala Stimulation Evokes Feed-Forward Inhibition in the Medial Prefrontal Cortex

Dilgen_umaryland_0373D_10218.pdf (3.16 MB)

Authors: Dilgen, Jonathan E. Advisors: O'Donnell, Patricio

Date: 2011-01 Embargoed Until: 2012-08-01

Abstract: The amygdala sends a strong projection to the prefrontal cortex (PFC) and both brain regions are

critical for high order cognitive processes. This pathway may convey emotional and incentive-value related information to the PFC where it could be used to guide behaviors such as decision making and working memory. Behavioral studies have shown that both the amygdala and PFC are important for emotional processing and learning, but the underlying neural correlates remain unclear. Anatomy studies show that the basolateral amygdala (BLA) sends an excitatory projection to the medial PFC (mPFC). Paradoxically, electrophysiologic evidence indicates that most mPFC neurons are inhibited by BLA stimulation. This apparent incongruence could be explained if mPFC interneurons are considered mediators of BLA to mPFC neurotransmission. In this thesis, I tested the hypothesis that the BLA-mPFC projection includes a feed-forward

inhibition mechanism with several experiments using electrophysiological approaches.

Subject Keywords: Amygdala

Electrophysiology Intracellular Prefrontal

Description: University of Maryland in Baltimore. Neuroscience. Ph.D. 2011

Type: Thesis

Appears in Collections: Theses and Dissertations School of Medicine

Table 1 Additional Metadata Elements

Element	Qualifier	Note
description	embargo	Embargo ending date
description	uri	URL of external content
description	uriname	URL Name
identifier	doi	DOI
identifier	pmid	Pubmed ID
identifier	spage	Starting Page
identifier	epage	End Page
identifier	issue	Publication issue number
identifier	volume	Publication Volume number
identifier	ispublished	Flag if content is published