



NATIONAL ENDOWMENT FOR THE

Humanities

OFFICE OF DIGITAL HUMANITIES

## **Narrative Section of a Successful Application**

The attached document contains the grant narrative and selected portions of a previously funded grant application. It is not intended to serve as a model, but to give you a sense of how a successful application may be crafted. Every successful application is different, and each applicant is urged to prepare a proposal that reflects its unique project and aspirations. Program guidelines also change and the samples may not match exactly what is now required. Please use the current set of application instructions to prepare your application.

Prospective applicants should consult the current Office of Digital Humanities program application guidelines at <https://www.neh.gov/grants/odh/digital-humanities-advancement-grants> for instructions.

Applicants are also strongly encouraged to consult with the NEH Office of Digital Humanities staff well before a grant deadline.

Note: The attachment only contains the grant narrative and selected portions, not the entire funded application. In addition, certain portions may have been redacted to protect the privacy interests of an individual and/or to protect confidential commercial and financial information and/or to protect copyrighted materials.

Project Title: Ensuring Access to Endangered and Inaccessible Manuscripts

Institution: St. John's University, Collegeville

Project Director: Columba Stewart

Grant Program: Digital Humanities Advancement Grants, Level III

# vHMML 3.0 Table of Contents

A 1	List of project participants
B 1	Abstract
C 1-10	Narrative
D 1-3	Biographies
E 1-2	Budget and budget notes
F 1-9	Appendices Appendix 1: HMML's Digital Projects with Eastern Christian and Islamic Manuscripts Appendix 2: Detailed Work Plan
G 1-9	Letters of commitment and support Prof. Fatemeh Keshavarz, Roshan Institute for Persian Studies, University of Maryland, College Park Prof. Sabine Schmidtke, School of Historical Studies, Institute for Advanced Study, Princeton, New Jersey Prof. Jack Tannous, Department of History, Princeton University
H 1-2	Data management plan
I 1-3	Sustainability Plan

# vHMML 3.0 Project Participants

## HMML Staff

- Columba Stewart OSB, D.Phil., HMML Executive Director and project director
- Daniel Gullo, Ph.D., Joseph S. Micallef Curator of the Malta Study Center, HMML Assistant Director, and assistant project director
- William Straub, M.L.I.S., Systems Librarian
- Eileen Smith, M.L.I.S., Metadata Librarian
- Wayne Torborg, Director of Digital Collections and Imaging

## Software Developer

- Chad LaVigne, B.S., software developer, Solution Design Group, Golden Valley, MN

## Advisory Board

- Scott Fitzgerald Johnson, D.Phil., Assistant Professor of Classics and Letters, University of Oklahoma, Norman, Oklahoma
- Matthew Thomas Miller, Ph.D., Roshan Institute for Persian Studies, University of Maryland, College Park, Maryland
- Carin Ruff, Ph.D., Independent Scholar, Washington, D.C.
- Caroline T. Schroeder, Ph.D., Professor of Religious Studies, University of the Pacific, Stockton, California

## Letters of Commitment:

- Fatemeh Keshavarz, Ph.D., Roshan Institute Chair in Persian Studies, University of Maryland, College Park, MD
- Sabine Schmidtke, D.Phil., Professor in Historical Studies, Institute for Advanced Study, Princeton, NJ

## Letter of Support:

- Jack Tannous, Ph.D., Assistant Professor of History, Princeton University, Princeton, NJ

## **vHMML 3.0 Project Abstract**

The Hill Museum & Manuscript Library (HMML) at Saint John's University, Collegeville, Minnesota, seeks support for the next phase of development for its recently-launched vHMML platform for manuscript studies ([www.vhmml.org](http://www.vhmml.org)). vHMML 1.0 was released in October 2015 with resources to support the use of manuscripts in research; vHMML 2.0 launched in August 2016 with an online Reading Room that is making tens of thousands of otherwise inaccessible and often endangered manuscript books and archival documents available to users around the world free of charge. NEH funding will make it possible to create vHMML 3.0, with greatly increased discoverability of manuscripts and metadata, and much richer data sharing with other digital humanities projects. vHMML 3.0 will add features requested by partner projects and researchers, and NEH support in both outright and matching funds will sustain the human resources needed to guarantee best-practice administration and continued development of vHMML.

# HILL MUSEUM & MANUSCRIPT LIBRARY

## Proposal for a Level III Digital Humanities Advancement Grant

### Introduction

The Hill Museum & Manuscript Library (HMML) at Saint John's University, Collegeville, Minnesota, seeks support for the next phase of development for its recently-launched vHMML platform for manuscript studies ([www.vhmml.org](http://www.vhmml.org)). vHMML 1.0 was released in October 2015 with resources to support the use of manuscripts in research; vHMML 2.0 was launched in August 2016 with an online Reading Room that is making tens of thousands of otherwise inaccessible and often endangered manuscript books and archival documents available to users around the world free of charge. NEH funding will make it possible to create vHMML 3.0, with greatly increased discoverability of manuscripts and metadata, and much richer data sharing with other digital humanities projects. vHMML 3.0 will add features requested by partner projects and researchers, and both outright and matching funds will sustain the human resources needed to guarantee best-practice administration and continued development of vHMML.

### Enhancing the Humanities and Protecting Our Cultural Heritage

#### Background

Since 1965, HMML has been photographing manuscripts around the world to ensure that their contents are preserved in case of loss and available to researchers. The project began in Austria to microfilm monastic libraries threatened by potential conflict during the Cold War and then expanded to other European locations, photographing tens of thousands of manuscripts in libraries across the Continent. In the early 1970s, HMML began a groundbreaking project in Ethiopia (the Ethiopian Manuscript Microfilm Library, EMLL) to film manuscripts threatened by modernization, insecure guardianship, and illegal trafficking in cultural heritage. Thousands of Ethiopian manuscripts were photographed against the backdrop of a revolution and enormously disruptive social and political forces. Many of those manuscripts have disappeared into private collections abroad, beyond the reach of scholars. The EMLL microfilms and the catalogs published by HMML describing their contents have transformed the understanding of Ethiopian literature both within Ethiopia itself and among western scholars. HMML has also worked systematically to photograph manuscript archival materials related to Malta, crossroads of the Mediterranean, tracing the history of the Knights of Saint John of Jerusalem (the Order of Malta) and their commercial and political ties throughout the region, as well as documenting the social life of the early modern Mediterranean. That work continues to this day, supporting the research of those interested in the legacy of the Crusades as well as economic and social historians of the region. Other recent efforts in Europe have focused on Slavonic-tradition manuscripts in Ukraine, Romania, Hungary, and the Balkans.

In 2003, HMML adopted high-quality digital technology and launched a new initiative to photograph the manuscripts of Christian communities in the Middle East. HMML provides equipment, training, salaries, and technical support to local workers selected by the project partners. The owning institutions receive a copy of all images and retain publication rights; HMML manages data archiving, cataloging, and scholarly access, which is provided most recently through

## vHMML Reading Room.

In April 2003, digitization work began in Lebanon, soon spreading to Syria, Turkey, Iraq, the Old City of Jerusalem, and Egypt, working with Armenian, Syriac, Christian Arabic, and Coptic communities. There have been several new projects in Ethiopia, including the digitization of the famed Abba Garima Gospels, considered by many experts to be the oldest Ethiopian manuscripts in existence, and the Islamic manuscripts of Harar. The work with Syriac manuscripts led to several projects with the historic Syriac-tradition communities in Kerala State, southwest India. A project with a Christian Syriac Orthodox monastery in the Old City of Jerusalem prompted the decision to expand the scope of HMML's work to manuscripts held by prominent Muslim families in Jerusalem. This opening to Islamic manuscripts has resulted in HMML's leadership of the effort to digitize the manuscripts of Timbuktu in Mali, and new partnerships focused on Persian and Yemenite-tradition Shiite manuscripts, complementing the Sunni-tradition projects in Harar and Jerusalem.

The vast majority of the manuscripts digitized in recent projects are accessible to researchers only through HMML's digital surrogates. For example, the important family libraries of Islamic manuscripts in Jerusalem are now physically inaccessible to their primary audience of scholars in the West Bank who are no longer able to visit Jerusalem. Through HMML's work the resources of these libraries are becoming available online. Dozens of other manuscript collections are held in remote or otherwise inaccessible locations, kept by communities that have no infrastructure for serving visiting scholars. In other cases, the original manuscripts have been destroyed or moved to unknown locations. Several hundred manuscripts digitized in Iraq under HMML sponsorship were destroyed during the conquest of Mosul and surrounding villages in the summer of 2014. The future security of many other collections digitized by HMML in the Middle East, South Asia, and Africa is uncertain. The contents of all of these libraries digitized since 2003—already totaling more than 50,000 manuscript books and archival documents, and growing rapidly—complete the partial view previously obtainable from manuscripts taken to libraries in Europe and North America.

## vHMML

To make these resources available and to prepare a new generation of scholars to work with manuscripts, HMML has spent the last several years developing vHMML (*virtual* HMML). vHMML provides researchers with an environment for learning about manuscripts, discovering new texts, comparing versions of known texts in several languages of transmission, and tracing the circulation and use of manuscripts across time and cultures. Underlying vHMML Reading Room is a sophisticated catalog database with tools for original cataloging, standardization of metadata, and export of metadata for use in other projects. These resources and tools are attracting leaders of other digital manuscript projects to choose vHMML as the home for their materials, e.g., the Roshan Initiative for Persian Digital Humanities (University of Maryland) and the Zaydi Manuscript Tradition project (Institute for Advanced Study in Princeton, New Jersey). Their Letters of Commitment are part of this application.

As of January 1, 2017, vHMML Reading Room contains more than 10,000 objects, with more collections being added weekly. Current information on featured collections is available at <https://www.vhmml.org/readingRoom/>. All of HMML's digitized manuscripts as well as scans of much of the legacy microfilm from Europe and Ethiopia will ultimately be in vHMML Reading Room, making it by far the world's largest online manuscript collection. Eventually the holdings

could be greater than 100,000 digital objects.

Study of these little-known materials has the potential to change how the history and literature of ancient Christian and Muslim communities are understood both in their particular contexts and in their interactions with each other. As the online home of manuscripts “left behind” or never discovered by western manuscript hunters, vHMML complements the famous manuscript collections of former colonial powers such as the British Library, the Bibliothèque Nationale de France, the Berlin Staatsbibliothek (formerly the Prussian Royal Library), or the Vatican Library. The attached list of recently digitized collections being made available in vHMML Reading Room provides a sense of the breadth and depth of these resources (see Appendix 1). One of the great advantages of vHMML Reading Room is the ability to search across collections, shifting the traditional collection-based approach to manuscript studies to one freed from the historical contingencies and accidents of acquisition and displacement. At the same time, those interested in the history of particular libraries can focus their research accordingly.

Developing vHMML 3.0 will add important features—identified by both internal and external users—that could not be included in the initial versions because of limitations of time and funding. Improvements will be both outward and inward facing. Images and metadata in vHMML will become discoverable to search engines (currently the database is not exposed to web-crawling bots). Researchers who depend on the site for access to the manuscripts they need will find an enhanced search experience that makes exploring across collections, cultures, and languages much easier. They will be able to download images and email metadata. Additional tools for vHMML administrators and catalogers will improve the efficiency of creation, curation, and export of data. Building upon the metadata standardization and authority control of vHMML 2.0, greatly improved metadata export workflows and formats will facilitate richer sharing with other digital humanities projects.

vHMML has been designed to be as open and accessible as possible. The key software components are available on Github for review and reuse by other projects (see below). The educational resources on vHMML (School, Folio, Lexicon, Reference) and all metadata for vHMML Reading Room are freely available to anyone who visits [www.vhmml.org](http://www.vhmml.org). Access to images of complete manuscripts is available to users who complete a one-time, no-cost, registration process. Further information can be found in the vHMML Terms of Use (<https://www.vhmml.org/terms>) and Privacy Policy (<https://www.vhmml.org/privacy>). All metadata in vHMML is available for reuse according to Creative Commons licenses (CC BY-NC 4.0 and CC BY 4.0) and can be exported for use by other projects upon request. A goal of vHMML 3.0 is to make this process simpler by allowing data harvesting in encoded formats.

## Environmental Scan

Access to digitized manuscripts through online galleries has become commonplace, even if important collections are not yet, or only partially, available. HMML considers “e-codices: Virtual Manuscript Library of Switzerland” (<http://www.e-codices.unifr.ch/en>) to be a benchmark in terms of technical standards and the significance of its holdings, which include very early and important Latin manuscripts. Like vHMML Reading Room, e-codices is a “library of libraries,” though with a more restricted geographical scope. This points to one of the challenges of the current online landscape: individual libraries, or in the case of e-codices, specific regions, develop digital collections hosted on separate platforms. An example of well-designed single-library platform is

the Digital Bodleian <http://digital.bodleian.ox.ac.uk/> from Oxford. There are tentative efforts to create aggregating access portals to make it easier for researchers to find the texts they need (e.g., a new initiative by the Center for the Study of Manuscript Cultures at the University of Hamburg, Germany, with which HMML is sharing data), but these efforts are hampered by varying standards for metadata. Progress is being made on unifying standards and establishing protocols for shared metadata, but there are still many barriers to discoverability.

vHMML Reading Room was designed to conform to best practices for describing both manuscript books and archival materials, and has separate catalog branches for those two kinds of objects. The metadata structure for books implements the [Descriptive Cataloging of Rare Materials](#) (DCRM) standards developed by the Bibliographic Standards Committee of the Rare Books and Manuscripts Section (RBMS) of the American Council of Research Libraries. The schema for archival materials is based on [Describing Archives: A Content Standard](#) (DACS) of the Society of American Archivists, which is the implementation of international standards in the United States. vHMML Reading Room has implemented authority controls linked to the Library of Congress (LC) and Virtual International Authority File (VIAF) records for places, institutions, personal names, and titles. Controlled vocabularies for Genre/Form and Features use terms from the standard references used by special collections librarians and archivists ([The Art & Architecture Thesaurus](#) maintained by the Getty Research Institute and the RBMS Controlled Vocabularies). Genre/Form authorities have an associated PURL linking to the *Art & Architecture Thesaurus*. A user can click on any of the authority controlled names or terms and find all records containing the same name or term. Each object in vHMML Reading Room has a permanent URL for reference and linking from other projects.

vHMML Reading Room uses the Mirador IIIF image viewer, chosen for its superior viewing experience, potential for interoperability, and robust user community. Mirador was developed at Stanford University in partnership with Harvard, Yale, the University of Oxford, the National Gallery of Art, and other partners. Mirador allows multiple objects from within a collection, as well as objects from other libraries that use the IIIF standard, to be viewed in multiple viewing frames within a single workspace. (Both e-codices and the Digital Bodleian use the earlier Open Sea Dragon IIIF image viewer as their default, with the option of invoking a more sophisticated IIIF environment, Mirador in the case of e-codices, and Universal Viewer in the case of the Digital Bodleian.) Although most of the images in vHMML Reading Room are not immediately available for use in non-vHMML Mirador environments because of restrictions by the owners of the manuscripts, upon request HMML can supply trusted partners with the IIIF manifest URLs needed to invoke images from vHMML for use in their projects.

## History of vHMML

vHMML is the latest stage of HMML's evolution from a microfilm-based archive that produced printed handlists and catalogs, to being a pioneer in the electronic cataloging of manuscripts as a partner in the EAMMS (Electronic Access to Medieval Manuscripts) project in the mid-1990s, to the introduction in 2005 of HMML's first image database, *Vivarium*, as an adjunct to the online catalog (renamed OLIVER in honor of HMML's founder, Fr. Oliver Kapsner, OSB, and now called the Legacy Catalog). *Vivarium* contained sample images from each born-digital object as a complement to the metadata.

HMML's staff and Board of Overseers saw that HMML needed to establish a stronger presence in

the digital space and to make its digital manuscript surrogates more easily accessible. vHMML began as a project funded by the Institute for Museum and Library Services in 2012 to create an online environment for manuscript studies, focusing on instruction in paleography, annotated images, and supporting resources. A grant of \$350,903 was awarded in August 2012. vHMML 1.0 with those research tools launched in October 2015. A vHMML Council, consisting of experts in various fields of manuscript study, met annually from 2012 to 2014 to help plan and refine the project. Initially an external institutional partner provided software development; the project had to be rebooted with a contracted software developer. He has done excellent work, and who remained with the project for vHMML 2.0. He is the proposed developer for vHMML 3.0.

In the meantime, the project scope was broadened in 2013 to include an online Reading Room that would provide no-cost access HMML's digital collections. As planning continued, HMML staff determined that the Reading Room would require a completely redesigned cataloging schema and database to support effective searching by users. As a result, development of the project proceeded along two tracks: 1) the vHMML Reading Room platform largely funded by the Henry Luce Foundation (a planning grant in 2014 for \$40,000, followed by a three-year project grant of \$500,000 in 2015); 2) creation of the new catalog database and continued original cataloging of recently-digitized manuscripts, supported by a project grant from the Andrew W. Mellon Foundation (three-year project grant of \$651,000 in 2015, including \$72,000 for software development, with rest for cataloging). vHMML 2.0 with Reading Room and the new catalog database launched in August 2016. A final 2.1.9 release in December 2016 completed the development cycle.

As of January 8, 2017, vHMML Reading Room contains 10,277 objects and has almost 500 registered users. As collections are added weekly to vHMML Reading Room, the pace of registration is accelerating. The Legacy Catalog (formerly OLIVER) will be deactivated once all of its records have been corrected and migrated to vHMML Reading Room.

Funding from the Henry Luce Foundation continues through December 2017, supporting the Systems Librarian and other expenses associated with adding objects to vHMML Reading Room. The Mellon grant continues through June 2018, funding the Metadata Librarian and catalogers. All software development for vHMML ended in December 2016 with the release of vHMML 2.1.9. No further funding to develop vHMML 3.0 is available from these sources.

Throughout the development process vHMML was rigorously tested by both internal and external users. The team used agile development techniques for maximum responsiveness to bugs, evolving requirements, and user feedback. Every element of the project continues to be tracked using Atlassian's JIRA software, which has contributed greatly to the project's success. During the last development cycle vHMML Reading Room was demonstrated at numerous conferences and workshops, including meetings of the Mediterranean Seminar at Fresno State University and at HMML; the International Congress on Medieval Studies, Kalamazoo; staff of the National Archives of Malta; two sessions for leaders of the African/Middle East Division and the World Digital Library at the Library of Congress; two workshops at the Center for the Study of Manuscript Cultures and the Hiob Ludolf Center for Manuscript Studies at the University of Hamburg, Germany; the Roshan Institute for Persian Studies at the University of Maryland, College Park; members of the School of Historical Studies, Institute for Advanced Study, Princeton; demonstrations for partner libraries in Malta, Jerusalem, Lebanon, Iraq, and Egypt. Throughout the development cycle the Systems Librarian has participated in bi-weekly meetings with the IIF Discussion Group and the Mirador

Discussion Group. Members of the vHMML Council provided counsel, served as testers, and contributed content to vHMML School and Folio.

The four-year process (2012-2016) included many challenges and lessons learned. By the time vHMML 1.0 was launched in October 2015, the team was seasoned and the workflow was clear and robust. This smoothed the path for developing vHMML 2.0 and positions vHMML 3.0 for success.

## Work Plan

**Project Timeline:** 24 months, September 2017-August 2019

### Project Phases

- Phase 1: September 2017— Mid-January 2018 (18 weeks)

This phase will focus on business analysis to outline the elements of the project in detail, including step-by-step descriptions of functionality, user scenarios, and mock-ups. This work is time-consuming, but previous experience with vHMML 1.0 and 2.0 has proven the importance of thorough description done before the software developer starts work.

Responsibility for this phase: Columba Stewart (Project Director), Daniel Gullo (Project Assistant Director), William Straub (Systems Librarian), Eileen Smith (Metadata Librarian), Wayne Torborg (Director of Digital Collections and Imaging).

Requested funding: none. HMML will contribute staff time as cost-sharing.

- Phase 2: Mid-January—August 2018 (34 weeks) [see detailed work plan in Appendix 2]

Software development for vHMML will be done in 17 two-week sprints as described in the detailed work plan. There will be three kinds of development: 1) major system upgrades and features; 2) additional features to enrich the user experience; 3) features to improve the cataloging workflow. The added features either could not be accomplished within the constrained time and funding cycles of the earlier versions, or emerged as needs in later phases of development or after the launch of vHMML 2.0 with Reading Room.

- **Major system upgrades and features.** vHMML requires periodic upgrades of its underlying software to ensure both security and performance. New systems features will include making metadata discoverable by search engines (a major improvement of crucial importance for making the holdings more widely known and useable) and developing the ability to export metadata in encoded form (e.g., METS, OAI, and EAD) for easier sharing with other digital humanities projects and institutions. This feature has been requested by several partner projects.
- **Additional features to enrich the user experience.** These will enrich the search experience, make possible the downloading of images (where allowed by the owning library), allow email export of metadata by users, and improve communication within vHMML to keep users better informed about additions to the holdings, new features, and other news.

- **Improvements to Cataloging Workflow.** Several HMML digitization projects have been focused on collections that have not yet been cataloged and are unlikely to be fully cataloged in the near future (e.g., the tens of thousands of manuscripts from Timbuktu, or palm-leaf archives from India). Provisional short-form metadata records are created so that such materials can be housed in vHMML without waiting for full cataloging. The present system for doing so is cumbersome and cannot be done directly by those at project sites. The plan is to create a new form within the vHMML Reading Room catalog database system for direct input of basic metadata by those at digitization sites.

Responsibility for this phase: HMML project staff as for Phase 1, particularly the Systems Librarian (William Straub) and Metadata Librarian (Eileen Smith), working with a contracted software developer, Chad LaVigne of Solution Design Group (<https://solutiondesign.com/>).

Requested funds: 34 weeks of contracted development work @\$150/hour; support for the Systems Librarian from January 2018; NEH Sustainability Match and Third-party Funding to support the Metadata Librarian from July 2018. HMML will contribute other staff time as cost-sharing.

- Phase 3: September 2018-August 2019 (52 weeks)

This phase is focused on the launch and marketing of vHMML 3.0, post-production tasks of documentation and code sharing, optimizing the new functionality for discoverability and sharing of data, continued work on the content of vHMML Reading Room, and planning for the future.

- **Marketing and Community Involvement.** vHMML 3.0 will be marketed through HMML's various communications channels, as well as in presentations to scholarly audiences at conferences and workshops. The Systems Librarian will continue to participate in the IIF technology group. (see "Final Product and Dissemination" below).
- **Post-Production Documentation.** The Systems Librarian will expand the software documentation provided by the programmer to make it more accessible and comprehensive, including preparation of documents complying with accessibility standards and removing internal or private information. He will create a second working copy of the vHMML 3.0 codebase scrubbed of internal or private information and upload it to the Github repository for access by other projects.
- **Optimizing New Functionality.** Building on the new search engine discoverability developed in Phase 2, the Systems Librarian will use SEO (Search Engine Optimization) technology and analytics to maximize the impact of discoverability. He and the Metadata Librarian will implement workflows for harvesting and sharing vHMML metadata using the new tools for export of encoded data.
- **Continued Uploading to vHMML Reading Room and Metadata Issues.** Under the direction of the Metadata Librarian, newly-cataloged objects will be regularly uploaded to vHMML Reading Room, and legacy cataloging will be conformed to new standards. The new authorities (particularly author names) being generated by HMML cataloging will be shared with other projects (e.g., syri.ac, syriaca.org, the Center for the Study of Manuscript Cultures) and submitted to VIAF /LC for

- permanent registration.
- **Looking Ahead.** The Systems Librarian will work with the Library Information Technology Association (LITA) to understand technologies associated with linked data and the semantic web, such as RDF (Resource Description Format) and Sparql (RDF query language and protocols).

Responsibility for this phase: As for Phase 2, these efforts will involve all HMML project staff to some degree, especially for the marketing of vHMML 3.0, but most of the tasks will be the particular responsibility of the Systems Librarian and Metadata Librarian.

Requested funds: support for the Systems Librarian; NEH Sustainability Match and Third-party Funding to support the Metadata Librarian. HMML will contribute other staff time as cost-sharing.

### **Risks and Challenges**

vHMML 3.0 builds upon the solid foundation of the previous versions. The parameters and time estimates for the project are based on two years of experience working with the developer. They include time for testing the new features and for regression testing to ensure that the new features do not adversely affect the existing software. This previous experience will mitigate the risks of unreasonable delay or cost overruns. Experience also suggests that some features may be implemented faster than planned, allowing resources to be redeployed. HMML runs its digital projects with agile software development techniques using the Atlassian JIRA platform, which has proven very useful both for software development and for ongoing metadata preparation and correction for vHMML Reading Room. HMML also uses PB Works, a wiki software, to consolidate resources related to software development, plans, goals, logic, and documentation. The whole project benefits from CSB/SJU IT Services infrastructure and expertise for server administration, data back-up, and security (see Data Management Plan).

### **Evaluation of Project Activities**

Planning for vHMML 1.0 was done with the help of an external advisory group. A new, longer-term, advisory board was established in December 2016 and will play an important role in developing vHMML 3.0 and strengthening ties with other digital humanities projects working with manuscripts and/or the manuscript cultures represented in vHMML. vHMML has a growing and active user base that has already provided guidance for the proposed additional features for vHMML 3.0. The project team is well connected with users through a range of scholarly channels (HMML's fellowship programs and summer language schools, professional societies, etc.) as well as through vHMML itself, which allows user-submitted comments and corrections. Users will be surveyed both during the development phase and after the launch of vHMML 3.0 for feedback about their experience of the platform and about how they use it for their research. vHMML administrators can track usage of the various components of vHMML through Google Analytics and usage of the objects and collections in vHMML Reading Room through reporting tools developed for vHMML 2.0. These are vital indicators of success in reaching new users and more deeply engaging established ones.

## Project Team

### HMML Staff

- Columba Stewart OSB, D. Phil., Project Director. Executive Director of HMML since 2003; Professor of Theology at Saint John's University; P.I. of several grants for digital humanities projects. He will provide overall project direction and management, and brings scholarly expertise in several of the manuscript cultures represented in vHMML Reading Room. Anticipated time commitment: 25% as cost-sharing contributed by HMML.
- Daniel Gullo, Ph.D., Assistant Project Director. Joseph S. Micallef Curator of the Malta Study Center at HMML since 2014; Assistant Director of HMML since 2016; project manager for vHMML Reading Room, with extensive experience of business analysis and interaction with the software developer using JIRA. He brings expertise in medieval western history, Mediterranean studies, and archival materials. Anticipated time commitment: 25% as cost-sharing contributed by HMML.
- William Straub, M.A., M.L.I.S., Systems Librarian at HMML since 2014; principal systems administrator of vHMML and point of contact with both CSB/SJU IT Services and the software developer. Manages the VMs for vHMML, the MySQL database, and user administration. Anticipated time commitment: 100%, services contributed by HMML as cost-sharing for 4 months, then grant-funded for 20 months.
- Eileen Smith, M.A., M.L.I.S., Metadata Librarian at HMML since 2015; manager of cataloging workflow and administrator of authority controls for vHMML Reading Room. Anticipated time commitment: 100%, services contributed by HMML as cost-sharing for 10 months, then funded by Sustainability Match and Third-party funding for 14 months.
- Wayne Torborg, Director of Digital Collections and Imaging at HMML since 2004. Manages image creation, archiving, and preparation for uploading to vHMML Reading Room. Provides back-up support to the Systems Librarian for database management. Anticipated time commitment: 20% as cost-sharing contributed by HMML.

### Software Developer

- Chad LaVigne (Solution Design Group). Coded vHMML 1.0 and 2.0. Anticipated time commitment: 100% during 34-week development cycle, grant-funded.

### vHMML Advisory Board

A new vHMML Advisory Board was established in December 2016 as a successor to the vHMML Council. The initial members bring expertise in both manuscript studies and digital humanities projects. Periodic virtual meetings will be supplemented by face-to-face meetings at conferences or at HMML.

- Scott Fitzgerald Johnson (University of Oklahoma): Coptic, Greek, and Syriac manuscripts; co-founder of "syri.ac: An annotated bibliography of Syriac resources online" (<http://syri.ac/>); former member of the vHMML Council.
- Matthew Thomas Miller (Roshan Institute, University of Maryland): Persian and Arabic Manuscripts, co-director of the Roshan Initiative in Persian Digital Humanities (<http://persdig.umd.edu/>).
- Carin Ruff (Washington, DC): creator of Latin manuscript resources in vHMML School and Folio; former member of the vHMML Council.
- Caroline Schroeder (University of the Pacific): Coptic manuscripts; founder of the Coptic Scriptorium (<http://copticcriptorium.org/>).

## Final Product and Dissemination

### Marketing vHMML

HMML regularly promotes vHMML on the [www.hmml.org](http://www.hmml.org) website, on Facebook, and through Twitter, including weekly updates about new collections added to vHMML Reading Room. Registered users of vHMML are notified about new features through system-wide emails. HMML curators are active in professional societies such as the annual International Conference on Medieval Studies in Kalamazoo, the Medieval Academy of America, the North American Patristics Society, the Society for Biblical Literature, the Mediterranean Seminar, and specialist societies for Syriac, Armenian, and Islamic studies. HMML and vHMML Reading Room news is announced on listservs. The Systems Librarian participates in the regular online meetings of the IIF and Mirador working groups. The Metadata Librarian will continue to present seminar papers at RBMS (Rare Books and Manuscript Section of the Association of College and Research Libraries) and at the ALA (American Library Association) about metadata issues associated with unknown collections and understudied literary cultures.

### Access to Users with Disabilities

vHMML is committed to fulfilling US Federal Section 508 for web users with disabilities and follows the W3C's Web Content Accessibility Guidelines (WCAG). Almost without exception the features are keyboard (rather than mouse or trackpad) accessible. Sans-serif fonts were chosen for the sake of clarity. HTML tags have been embedded at key points to assist readers in recognizing the presence of links. vHMML also uses checkboxes, dropdowns, and other text fields that have descriptive HTML tags. As a web-based project, all vHMML content can be viewed with browser magnification, and all manuscript images in vHMML can be viewed in full-screen mode.

### Plans for the White Paper

The white paper for vHMML 3.0 will provide an honest narrative of the project's history and challenges. It will emphasize the importance of collaboration with other institutions and experienced individuals for successful attainment of project goals. The white paper will describe in detail the activities of the project planning cycle, the development workflows, tracking tools used by the project team, and interactions with testers and users.

### Sharing of Software and Content

- The software created for vHMML 2.0 is being uploaded to a Github repository. This includes all of the Java, HTML, and CSS code. A useful IIF manifest generator will be placed in a separate Github repository. This software is available under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/) (CC BY-NC 4.0). The same will be done for software created for vHMML 3.0.
- The textual content of vHMML School and Folio are available for reuse under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/) (CC BY-NC 4.0).
- The manuscript metadata in vHMML Reading Room is available for reuse under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/) (CC BY 4.0).

All metadata currently in vHMML Reading Room is easily exportable in JSON and CSV formats; vHMML 3.0 will add encoded forms such as METS, OAI, and EAD.

# vHMML Biographies

## HMML Staff

- **Columba Stewart, OSB, Project Director** (A.B., Harvard College, 1979; M.A., Yale University, 1982; D.Phil., University of Oxford, 1989) is Professor of Theology at the Saint John's School of Theology, specializing in the history of early Christian monasticism east and west, and has published widely in his field of research. Since 2003 he has been Executive Director of the Hill Museum & Manuscript Library at Saint John's University, where he has guided the transition to digital media for manuscript preservation and access, and expanded HMML's field work to several concurrent projects in Europe, the Middle East, Africa, and South Asia. He developed the original concepts for vHMML 1.0 and for the Reading Room component of vHMML 2.0. He has been successful as a grant writer for HMML (IMLS, Mellon Foundation, Luce Foundation, Arcadia Fund) and in securing fellowships for his own research (NEH Fellowships in 1994-95 and 2009-10; a Dumbarton Oaks Fellowship for 2009-10; Membership in the School for Historical Studies at the Institute for Advanced Study in Princeton for 2016-17; a Guggenheim Fellowship awarded in 2016). He is currently writing a new history of the origins and early development of Christian monasticism, to be published by Oxford University Press. More information can be found at <https://ias.academia.edu/ColumbaStewart> .
- **Daniel Gullo, Assistant Project Director** (B.A., University of Wisconsin, Madison, 1991; M.A., University of Toronto, 1996; M.A., Saint John's University School of Theology, 1998; Ph.D., University of Chicago, 2009) has been the Joseph S. Micallef Curator of the Malta Study Center at the Hill Museum & Manuscript Library since 2014 and the Assistant Director of HMML since 2016. His area of research is late medieval and early modern Spain, with a particular interest in the use of printing technology in religious reform movements. Daniel currently serves on the board of the Mediterranean Seminar, and is a continuing advisor to the project, "Digital Observance: Visualizing Observant Reform in the Middle Ages." He held fellowships at the Newberry and Huntington Libraries (2002), and has supported HMML's mission through successful grants with the Gladys Kriebel Delmas and B.H. Breslauer Foundations (2014 and 2015). Daniel was heavily engaged in the development of vHMML 1.0 and managed the development of vHMML 2.0, working closely with the software developer. More information can be found at <http://www.hmml.org/malta-study-center.html> or <https://hmml.academia.edu/DanielGullo>.
- **William Straub, Systems Librarian** (B.A., Saint John's University, 1988; M.L.I.S., University of Wisconsin, 2010) joined the HMML staff in 2014 to help develop vHMML Reading Room but was also deeply involved with the creation of the vHMML 1.0 platform. He brings expertise in digital services management, digital pedagogy, and accessibility issues. William manages the complex server array used for vHMML in coordination with CSB/SJU IT Services, is the main contact with the software developers, does basic coding and CSS updates, and is the principal point of contact for users, processing user registrations and administering system-wide messaging for updates, technical alerts, and planned maintenance. Before coming to HMML, William was systems librarian for the University of

Minnesota-Morris, specializing in patron-user access requests and proxy issues for electronic databases.

- **Eileen Smith, Metadata Librarian** (B.A., College of Saint Catherine, Saint Paul, Minnesota, 1980; M.A. in Library and Information Science, University of Arizona, 1999) joined the HMML staff in 2015 to lead the effort to impose authority controls and standardization of metadata for the new catalog database behind vHMML Reading Room. She brings expertise in cataloging rare books at several libraries, including the Huntington Library, and more than a decade at the Beinecke Library, Yale University, before coming to HMML. Eileen has also served as a NACO contributor at the Beinecke Library, submitting authority files to the Library of Congress. As Metadata Librarian she reviews all records before they are uploaded to vHMML Reading Room. She also works with catalogers at HMML and off-site to maintain updated lists of controlled vocabulary; to establish preferred LC and VIAF authorities for places, personal names, and titles; and to record instances in which no such authority exists so that these can be submitted to LC and VIAF.
- **Wayne Torborg, Director of Digital Collections and Imaging** (B.A., Saint Cloud State University, 1984), joined the HMML staff in 2003 to lead the move into digital imaging and archiving. He designed and set up the first digital studio in northern Lebanon, and has continued to refine the technology and workflow that have supported dozens of projects in Europe, the Middle East, Africa, and South Asia. He archives all digital images, prepares images for uploading to vHMML Reading Room, administers the uploading workflow for images, and supports the Systems Librarian in work with the MySQL database.

## Software Developer

- **Chad LaVigne** (B.S., Saint Cloud State University, 2000) has been a consultant at Solution Design Group (<https://solutiondesign.com/>) in Golden Valley, Minnesota, since 2013. Before coming to SDG, Chad spent several years as a consultant for TEK Systems. In both positions he served as enterprise architect, technical lead, and senior developer for complex database projects and web applications for the educational, corporate, and non-profit sectors. Chad works in all major programming languages and technologies, and is fully conversant with current tools and methodologies in software development. He is a Sun Certified Java Programmer, Sun Certified Web Component Developer, and Liferay Certified Developer. Chad took over development of vHMML 1.0 at a crucial point, and brought it to successful launch in October 2015. He was retained as developer for vHMML 2.0 with Reading Room, which launched in August 2016. He works in a deliberate, efficient, and responsive manner with a close eye on timelines and costs.

## Advisory Board

- **Scott Fitzgerald Johnson** (B.A., Vanderbilt University, 1999; M.Phil., University of Oxford, 2001; D.Phil., University of Oxford, 2005) is Assistant Professor of Classics and Letters at the University of Oklahoma. His research interests are multilingualism in the Late Antique East, Syriac language and literature, and the organization of knowledge in the Ancient and

medieval worlds. He has received several research fellowships, including the Harvard Society of Fellows (2004-06), a Dumbarton Oaks Fellowship (2009-10), and a Dumbarton Oaks Teaching Fellowship (2011-15). He has published monographs on Late Antique hagiography and cartography, translated Greek and Syriac hagiography, and edited several multi-author volumes, including *The Oxford Handbook of Late Antiquity*. With Jack Tannous (Princeton University) he is co-creator and co-editor of “syriac.ac—An annotated bibliography of Syriac resources online” (<http://syriac.ac>). He served on the original vHMML Council. Further information can be found at <http://scottjohnson.com/cv/>.

- **Matthew Thomas Miller** (B.A., Washington University in Saint Louis, 2007; M.A., Washington University in Saint Louis, 2009; Ph.D., Washington University in Saint Louis, 2016) is Associate Director, Roshan Initiative in Persian Digital Humanities, and Instructor in Islamic Studies at the University of Maryland, College Park. His research interests are Persian literature of all periods, digital humanities, literary and cultural theory. He has published articles on Persian poetry and is co-PI for both the Persian Manuscript Initiative (<http://www.persianmanuscript.org/>) and the Islamicate Texts Initiative (<http://iti-corpus.github.io/>). He brings expertise in Islamic texts and manuscripts, both Persian and Arabic, and a dedication to making both digitized manuscripts and textual corpora more widely available. Through the Roshan Institute, he is working with HMML to digitize important collections of Persian manuscripts in India and Afghanistan, and to catalog the Persian manuscripts of the Library of Congress. All of these resources will be available in vHMML Reading Room. Further information can be found at <http://www.matthewthomasmiller.com/>.
- **Carin Ruff** (B.A., Swarthmore College, 1987; M.Phil., University of Oxford, 1990; M.S.L.S., The Catholic University of America, 1992; Ph.D., University of Toronto, 2001) was Associate Professor of English at John Carroll and Cornell Universities before leaving full-time teaching to work on projects related to historical preservation. She has published on medieval Latin literature, particularly in an Anglo-Saxon context, and continues to teach Classics and Latin paleography part-time at the University of Maryland. Her pedagogical skills and subject expertise are evident in the Latin components of vHMML School and Folio, which she created. She served on the original vHMML Council.
- **Caroline T. Schroeder** (A.B., Brown University, 1993; Ph.D., Duke University, 2002) is Professor of Religious Studies at the University of the Pacific, where she has taught since 2007 and was Director of the Humanities Center 2012-2014. Her research primarily concerns asceticism and monasticism in early Christianity, with a particular focus on monasticism in Egypt. Her book *Monastic Bodies: Discipline and Salvation in Shenoute of Atripe*, was published by the University of Pennsylvania Press in 2007, and she is completing a monograph entitled *Monks and Their Children: Family and Childhood in Early Egyptian Monasticism*. With collaborator Dr. Amir Zeldes of Georgetown University, Dr. Schroeder is a co-founder of an online platform Coptic Scriptorium, dedicated to open-source and open-access texts and technology for the study of Coptic language and literature. The project can be accessed at <http://copticcriptorium.org>. More information is available on her website at <http://www.carrieschroeder.com>.

## **vHMML 3.0 Appendices**

**Appendix 1 (pp. F-2—F-5): List of Recent HMML Projects with Eastern Christian and Islamic Manuscripts (note: does not include recent Malta-related projects)**

**Appendix 2 (pp. F-6—F-9): Detailed Work Plan for Phase 2 and Technical Notes**

# Hill Museum & Manuscript Library

Collection Status: January 2017

*NOTE: these manuscripts will be made available in the new vHMML online Reading Room*

## Arabic Manuscripts (Christian)

### Lebanon

Al-Harf	Monastery of St. George	17	
Balamand	Our Lady of Balamand Monastery	210	
	Institute for Antiochian Orthodox Heritage	41	
Beirut	Near East School of Theology	293	
	Université Saint-Joseph (includes Islamic manuscripts)	1895*	
Brommana	Church of the Prophet Isaiah	15	
Brommana	Orthodox Archbishopric of Mt. Lebanon	9	
Bzommar	Clergy Institute of Bzommar (Arabic collection)	250	
Gharzouz	Church of the Theotokos	7	
Hamatoura	Monastery of Our Lady	47	
Harissa	Paulist Order	300	7574
Jounieh	Aleppan Basilian Order	1358	
Khonchara	Chouerite Basilian Order	930	
Kosba	Monastery of St. Dimitri/Church of SS. Sergius and Bacchus	15	
Mhaidseh	Church of the Theotokos	12	
Monsef	Church of SS. Sergius and Bacchus	28	
Shwayya	Monastery of St. Elias	63	
Zouk	Basilian Chouerite Nuns	74	

### Syria

Aleppo	Greek-Catholic Archdiocese, Archbishopric Collection	428	
	Greek-Catholic Archdiocese, Sbath/ Salem Collection	550	
	Syriac Catholic Archdiocese	628	
Damascus	Greek-Catholic Patriarchate	170	
Homs	St. George Al-Hmayra Monastery	46	
	St. Elian Antiochian Orthodox Church	42	
	Antiochian Orthodox Bishopric	53	

## Arabic Manuscripts (Islamic)

### Ethiopia

Harar	Sherif Harar City Museum	800*	
-------	--------------------------	------	--

### Jerusalem

Old City	Al-Budeiry Family Library	1150	
Old City	Al-Khalidiyya Library	1263*	
Old City	Sheikh Abdul Aziz Al-Bukhari Library	50*	
East Jerusalem	Dar Issaf Nashashibi Library	294	

**Lebanon**

Beirut	Bibliothèque Orientale, Université Saint-Joseph	see above
	Near East School of Theology	see above

**Mali**

	<b>(with SAVAMA-DCI; both literary and archival documents)</b>	
Timbuktu	Mamma Haidara Library	37000*
	Abu Bakr Bin Said Library	7574

**Armenian Manuscripts****Austria**

Vienna	Mechitarist Fathers (microfilm)	1200
--------	---------------------------------	------

**Iraq**

Baghdad	Armenian Orthodox Diocese of Baghdad	15
---------	--------------------------------------	----

**Lebanon**

Antelias	Catholicosate of Cilicia	242
Bzommar	Clergy Institute of Bzommar (Bzommar, Antonine, and New collections)	1890

**Syria**

Aleppo	Armenian Orthodox Diocese	220
--------	---------------------------	-----

**Turkey**

Istanbul	Armenian Patriarchate	786
----------	-----------------------	-----

**Coptic Manuscripts** (may include Arabic)**Egypt**

Wadi Natrun	Deir Abu Maqar	550*
-------------	----------------	------

**Ethiopic Manuscripts** (may include Amharic)**Ethiopia**

	EMML project (microfilm)	8000
	Gunda Gunde Monastery (with Gervers and Witakowska)	219
	Aba Garima Monastery (Gospels project)	2
	Institute of Ethiopian Studies (with EMIP)	1450

**Slavonic Tradition Manuscripts** (may include Greek)**Croatia**

Pakrac	Serbian Orthodox Eparchy of Slavonia	40
<b>Hungary</b>		
Szentendre	Serbian Orthodox Archdiocese	150
<b>Romania</b>		
Iași	Museum of Romanian Literature	30
<b>Ukraine</b>		
L'viv	Andrei Sheptytsky National Museum	2000*
	L'viv Historical Museum	160
	Vasili Stefanyk Library, L'viv, Ukraine	842

## Syriac Tradition Manuscripts (may include Garshuni and some Arabic manuscripts)

<b>India</b> (with the Association for the Preservation of the St. Thomas Christian Heritage; may include Malayalam)		
Bangalore	Dharmaram College	68
Ernakulam	Syro-Malabar Major Archbishopric	55
Kottayam	Saint Ephrem Ecumenical Research Institute (SEERI)	51
	Mar Thoma Seminary	36
Mannanam	Monastery of St Joseph	113
Pampakuda	Konat Collection	244
Piramadam	Gethsame Dayro	36
Thrissur	Chaldean Syrian Church (Assyrian Church of the East)	130
Trivandrum	Syro-Malankara Major Archbishopric	41
<b>Iraq</b> (with the Centre Numérique des Manuscrits Orientaux)		
Alqosh	Chaldean Diocese	145
Ankawa	Bishop Jacques Isac	48
	Pontifical Babel College Library	89
	Museum of Syrian Heritage	26
	Chaldean Sisters Daughters of Maris	46
Baghdad	Chaldean Patriarchate	500
	Syrian Catholic Archdiocese	132
	Dominican Fathers	2
	Saint Peter Seminary for Chaldean Patriarchate in Iraq	6
Baqofa	Saint George Chaldean Church	7
Bartelli	St George Syrian Catholic Church	84
Batnaya	Chaldean Church	56
Dawdiya	Eglise Mar Youhanna Al_ Mamdane	2
Duhok	Chaldean Diocese	32
Erbil	Chaldean Archdiocese	176
Kirkuk	Chaldean Archdiocese	137
Komani	Eglise Mar Qyriacos	3
Mangesh	Mar Gorges Chaldean Church	16
Mosul	Syrian Orthodox Archdiocese	299
	Dominican Friars	775
	Mar Behnam Monastery	529
	Chaldean Archdiocese	93

	Congregation of the Daughters of the Sacred Heart	93
Qaraqosh	Syrian Catholic Diocese	135
	Eglise Mar Yaqub	4
	Eglise Mar Yuhana al_Mamdane	18
	Mar Sarkis and Bakhos Syrian Orthodox Church	81
	Saint Ephrem Institute	5
Tell Kaif	Qalb Al-Aqdas Chaldean Church of Tell-kaif	227
Tellesqof	Saint George Chaldean Church	59
Zakho	Chaldean Archdiocese	76
	Virgin Mary Syrian Catholic Church	2
Various	Private libraries in Alqosh, Baghdad, Kanda-Kossa,	141
<b>Jerusalem</b>		
Old City	Saint Mark's Monastery	476
<b>Lebanon</b>		
Beirut	Université Saint-Joseph	58
Hrash	Sisters of St. John the Baptist	200
Jounieh	Lebanese Maronite Missionary Order	541
Kaslik	Lebanese Maronite Order	1185
Louaize	Notre Dame University	500
<b>Syria</b>		
Aleppo	Syriac Orthodox Archdiocese (Urfa Collection)	250
	Syriac Orthodox Archdiocese (St. Ephrem Collection)	202
	Syriac Catholic Archdiocese (also Arabic)	190
Homs	Syriac Orthodox Archdiocese (includes Sadad, etc.)	465
<b>Turkey</b>		
Ain Wardo	Mor Hadbshabo Church	54
Diyarbakir	Meryem Ana Syriac Orthodox Church	355
Enhil	Churches	84
Mardin	Chaldean Cathedral (Diyarbakir+Mardin collections)	588
	Church of the Forty Martyrs	1150
	Monastery of Deir Ulzaferan	266
Midyat	Monastery of Mor Gabriel and villages of Tur 'Abdin	320
	Midyat churches	275
		41827

\*=in progress

#### Contact

[www.hmml.org](http://www.hmml.org)

[www.vhmm.org](http://www.vhmm.org)

Columba Stewart OSB: cstewart@hmml.org

Anton Pritula: apritula001@hmml.org

## Appendix 2: Detailed Work Plan for Phase 2

### Phase 2: Mid-January—August 2018 (34 weeks)

The project includes software development of three kinds:

- **Major System Upgrades and Features**
  - Bring MySQL (currently 5.5.50), Tomcat (currently 7.0.54), JAVA (currently 1.8.0\_111-b14), and Elasticsearch (currently 1.7.5) to the latest versions (estimated time: 5 weeks)
    - Rationale: essential system updating for performance and security
  - Make vHMML Reading Room metadata discoverable by search engines (estimated time: 8 weeks)
    - Rationale: metadata is currently contained within the MySQL database hosted on Saint John's University servers, unexposed to web crawling bots, greatly limiting discoverability of the objects in vHMML Reading Room
  - Build tools to allow export of vHMML Reading Room metadata in encoded form (e.g., METS, OAI, and EAD) for easier sharing with other institutions and projects (estimated time: 8 weeks)
    - Rationale: HMML is committed to sharing metadata with other digital humanities projects to maximize use of the materials in vHMML Reading Room. Version 2.0 allows the export of metadata for individual collections or for the entire database in JSON format, and of permalinks for each object in CSV format. Partners have asked for data exports in encoded formats (e.g., METS, OAI, and EAD) for easier ingestion into their databases for portals and other aggregating efforts. Working closely with partner projects, the Metadata Librarian (Eileen Smith) will assess their metadata standards and work with the HMML Systems Librarian (William Straub) to crosswalk vHMML data into appropriate formats.
- **Enriched User Experience**
  - Ability to download images from vHMML Reading Room where allowed by owning institution (estimated time: 1 week)
    - Rationale: the Mirador viewer used in vHMML Reading Room allows for image download but this feature has not yet been implemented in vHMML Reading Room. The catalog database already incorporates a field to indicate if an object is available for download, and, if so, at what size. Although many of HMML's international partners do not yet permit download of images of manuscripts from their collections, an increasing number of them do. Some already permit download of small reference images. HMML will allow download of digital surrogates of its own manuscripts and rare books, and hopes that partner libraries will come to see the advantages of such access.
  - Ability to email formatted metadata to oneself or other users from within the platform (estimated time: 1 week)
    - Rationale: users now expect such functionality

- Additional hotlinks in vHMML Reading Room metadata display for uniform title, subject, and author (estimated time: 1 week)
  - Rationale: At present, the display of metadata in vHMML Reading Room includes hotlinks for system-wide searching for Country/ City/ Repository, Language, Genre, Features, Writing System, and Script (see an example at <https://www.vhmml.org/readingRoom/view/125944>). This upgrade would add hotlinks to other fields used extensively by scholars, consistent with modern library databases. The introduction of strict LC/VIAF authority control during the planning for vHMML Reading Room makes this possible.
- Improved searching, including advanced wildcard searching (estimated time: 4 weeks)
  - Rationale: Version 2.0 received several improvements in the Elasticsearch algorithm during its development cycle, and prolonged user experience is providing feedback that will be helpful for further smoothing and ranking of search results. Prolonged user experience is also providing feedback about further smoothing and ranking of search results. vHMML Reading Room currently allows for "\*" wildcard searching for the final letter or letters of keywords (e.g., searching for "Inquis\*" returns Inquisition, Inquisición, Inquisitionis). We were unable to implement advanced wildcard searching with "?" to represent single letters within keywords (e.g., searching for "advis?r" would return both "adviser" and "advisor"). This type of search is extremely useful for finding variant forms of names and titles, especially when no uniform title yet exists. Version 3.0 would also make it possible to do multi-match wildcard searching that combines the two forms.
- Newsfeed incorporated into vHMML Reading Room landing page and other landing page upgrades (estimated time: 4 weeks)
  - Rationale: at present, communication with users occurs through system-wide emails, which are useful for notifications about system upgrades, downtime, etc. but are not an appropriate medium for updates about new collections, tips for better searching, revised support pages, and other important information for users. At present, information about particular collections (i.e., whether they are in vHMML Reading Room, their cataloging status, etc.) is provided by dynamic spreadsheets accessible from the landing page ("What You'll Find in Reading Room," <https://www.vhmml.org/readingRoom/>). The information is up-to-date, but is not being pushed to users in an effective manner. The Reading Room landing page itself is the least visually compelling part of vHMML and would benefit from further design work to present its resources in a more effective manner.
- **Improvements to Cataloging Workflow**
  - Create new metadata form with limited fields for very basic, Cataloger Level 4 data entry (estimated time: 2 weeks)
    - Rationale: the catalog database supporting vHMML Reading Room has a sophisticated data entry system with elaborate forms for both manuscript books and archival materials. Access to the cataloging tools is granted by vHMML Reading Room administrators at three levels: Level 1 catalogers can create, revise, and approve any record, as well as create authority

records. Level 2 and 3 catalogers can edit only those collections to which they have been given access. For many recently-digitized collections (e.g., the Timbuktu manuscripts from Mali) detailed cataloging is a distant prospect. For such collections HMML creates very basic records that make it possible to house the object in vHMML Reading Room and available to scholars (see, e.g., <https://w3id.org/vhmmml/readingRoom/view/138014>). At present these are created in locally-hosted MS Access databases, with the .mdb file then converted to an MS Excel spreadsheet for easier review by HMML metadata librarians, then converted back into an MS Access file for import into the vHMML MySQL database. It would be very useful to have a simple form built into the cataloging system for the direct creation of such records by external catalogers who would be given a new Level 4 access. Adding this fourth tier of cataloger to enter basic metadata will improve access and discoverability for endangered collections.

Software development to be done by Chad LaVigne of Solution Design Group (SDG, <https://solutiondesign.com/>), who wrote the software for vHMML 1.0 and 2.0. After two years of full-time collaboration, HMML has great confidence in his ability to estimate the time needed for a particular task and to deliver a solid product. We have confirmed with SDG his potential availability, but if for any reason he would become unavailable, they have other developers who could be deployed for the project. The time estimates were established by HMML staff in consultation with Chad LaVigne and should be regarded as informed but not exact forecasts. Should these estimates prove to be overly generous, HMML staff will work with the developer to determine the best use of available development hours.

### **Technical Notes for vHMML 1.0 and 2.0**

This information was provided by the software developer to expand the basic information available at <https://www.vhmmml.org/about>.

- vHMML is a distributed system consisting of two Spring-based Java web applications, a search server, image server, and database.
- Scrum methodology was used with JIRA as the tool to manage sprints & tasks.
- System was developed using Eclipse IDE and deployed to Tomcat 7.
- MySQL 5.5 is used for data persistence.
- System was architected using Spring MVC controllers, dependency injected Service Façade objects and dependency injected Spring JPA Data Access Objects.
- Keyword and faceted search capability using Elasticsearch 1.7.
- Application security using Spring Security and ESAPI.
- Permanent link functionality via w3id.org.
- Diglib Image Server for IIIF compliant image manipulation and retrieval.
- Custom web services to provide IIIF compliant image manifest data for IIIF image viewers.
- External integration with Zotero reference citation software.
- Mirador IIIF compliant image viewer used to deliver high-resolution image viewing with zooming and comparison capability.

- Presentation layer was developed using JSP, Tiles, JavaScript, jQuery, Twitter Bootstrap, Handlebars and CSS.
- Flyway was used for automated schema management; Maven and Jenkins used for build management and continuous integration; JUnit, Mockito, Spring testing and JMeter used for automated unit, integration and load testing.
- Subversion source code control.

# vHMML Data Management Plan

## Data Types and Backup Procedures

The core of HMML's data preservation plan is to have backup copies of critical data in multiple formats and in multiple locations. Since the vHMML system is hosted by CSB/SJU IT Services (CSB/SJU=College of Saint Benedict/Saint John's University), HMML relies upon their workflows and processes for data loss prevention and restoration, as well as for system software backup and restoration. The vHMML system is tiered on multiple virtual machines (VMs) with the MySQL database and the image data as separate VMs. In addition, there are separate tiers for a vHMML Test platform as well as for the Production platform that is exposed to the public. Since the Test platform is regularly deleted and re-created, sometimes with bogus data for testing purposes, the data management policy outlined here deals only with the Production platform.

Two main types of electronic data are an intrinsic part of the vHMML system: (1) alphanumeric data such as catalog records and user data (registration details and access privileges), and (2) image data and associated image transfer metadata (IIIF JSON manifests). The first is stored in tables in a relational MySQL database; the latter is stored on a VM in a Linux directory structure organized by HMML sub-collection.

## Data Type 1: MySQL Database

### Backup Procedures

Currently, the vHMML MySQL database is backed up in four ways: 1) nightly through a MySQL dump to disk in distinct directories on the VM labeled with the date, as well as an off-campus dump to the data center at our partner institution, the College of Saint Benedict; 2) nightly taped incremental backup kept in the Saint John's University IT Services office; 3) weekly taped full backup (each Tuesday) kept in the Saint John's University IT Services office; 4) monthly taped full backup kept in a fireproof safe in another building on the Saint John's University campus. HMML plans to add a copy of the MySQL database to the backup tapes of our digital image collections sent periodically to off-site storage in Utah (see below). Every two hours throughout the workday, the MySQL database is dumped as a file on the VM, so that at most two hours of user and catalog data could be lost in case of a system failure. These bi-hourly data dumps are overwritten on a daily basis since they are redundant with the nightly taped backups.

### Sharing and Dissemination

All catalog metadata is publically exposed in vHMML Reading Room and can be viewed without registration. This metadata is available for reuse according to a Creative Commons license (CC BY 4.0). Currently, a vHMML Administrator can export catalog metadata as JSON or comma delimited CSV files for specific collections or for the entire database. vHMML 3.0 would add tools for export and harvesting of metadata in encoded format, as requested by project partners. Sharing of user data is done under the limited conditions outlined in the Privacy Policy:

<https://www.vhmml.org/privacy>.

### Intellectual Property

All metadata creators agree in writing to make their work available according to the CC BY 4.0 license.

## Data Type 2: Digital Images

### Image Formats and Backups

The cameras used by HMML produce two types of image files simultaneously—a RAW digital image and a JPEG image of extremely high quality. Both types of files are retained after capture. A complete set of both RAW and JPEG images is recorded to a hard disk drive for the use of and retention by the owning library. Another hard drive is prepared for shipment to HMML.

Once the hard drive arrives at HMML, all of the JPEG images are copied to an on-site file server operating on an internal local area network (LAN). This storage area network (SAN) system makes use of redundant hard disk arrays to safeguard the data. The JPEGs on the SAN are backed up by CSB/SJU IT Services using LTO-5 tape cartridges stored in a vault on campus. The RAW images are copied to LTO-5 backup tapes that are shipped to a secure storage facility in Utah (Perpetual Storage: <http://perpetualstorage.com/>). The original hard drive is retained and stored in the secure, climate-controlled, and fire-protected HMML microfilm vault.

An additional set of JPEG images is produced for use in vHMML Reading Room. These high-resolution JPEG images, slightly more compressed than the “maximum quality” JPEG images produced by HMML’s cameras, are far more usable for internet delivery but still have excellent viewing qualities. Images are rotated for the proper display orientation and JSON manifests are generated to allow them to be viewed in a IIIF environment. After preparation at HMML, these files are copied via secure FTP to the vHMML Reading Room Production image server. The original set of derivative images is retained on hard drives kept at HMML. All of the files on the Production server are backed up on tape by CSB/SJU IT Services as described above for the MySQL database. These images will also be taped for shipment to Perpetual Storage in Utah.

Throughout the process of making and storing backups, careful recordkeeping ensures that the status and location of all digital assets is known at all times. These records are likewise backed up and stored in redundant copies.

### Sharing and Dissemination

Image files are being made available in vHMML Reading Room. Viewing images for most of the collections requires a one-time, no-cost registration that allows HMML to monitor usage and address violations of the terms of access.

## Period of Data Retention/ Future Compatibility

HMML is a permanent institution with an endowment sufficient to guarantee ongoing operations at a level that would ensure continuing availability of data. The images and metadata in vHMML Reading Room will be available permanently even if in the future they are migrated to new systems. HMML staff are in continuous communication with CSB/SJU IT Services to ensure that hardware and software are available to retrieve data as needed. As storage formats or media used by HMML become deprecated or superseded, data will be copied to the preferred newer system, while in some cases retaining the data in older formats or media.

# vHMML Sustainability Plan

vHMML Reading Room is rapidly becoming the core resource of the Hill Museum & Manuscript Library, serving as both the point of discovery of resources through cataloging metadata and the environment for research through digital images of complete manuscripts. As a result, sustaining vHMML is inseparable from sustaining HMML itself.

The sustainability plan has three aspects: **institutional**, **financial**, and **human**.

## **Institutional Framework**

HMML was founded in 1965 as a sponsored program of Saint John's University. As such, HMML benefits from the infrastructure of CSB/SJU IT Services for server systems (including creation and maintenance of the VMs), data protection and retention, and other essential services. HMML pays an annual fee for servers, data storage, and backup services. In FY 2017 the fee will be approximately \$55,000.

HMML has its own Board of Overseers, which has governance authority delegated by the Saint John's University Board of Trustees. The Board of Overseers consists of 30 members, meets quarterly, and has a full array of active committees providing oversight for every aspect of HMML's work. The committees relevant for vHMML are the Preservation Committee (digitization projects and data management) and the Program Committee (vHMML and other forms of service to scholars and outreach to the general public).

In 2014 the HMML Board approved a 2020 Strategic Plan with three pillars:

- Continue to support HMML's leadership in the digitization of manuscripts.
- Become a leader in Digital Humanities.
- Expand opportunities to engage with HMML-based research.

The second pillar expresses the Board's commitment to vHMML and vHMML Reading Room, and their support for the fundraising required for its creation and support.

## **Financial Framework**

Context. HMML has responsibility for its own budget within Saint John's University. The FY2017 budget is \$2.3 million. Revenue sources are: 1) endowment, both restricted and unrestricted, providing approximately 20% of revenue; 2) annual fundraising, providing approximately 25%; 3) grants, currently providing approximately 40%, more than normal because of recent activity associated with vHMML Reading Room; 4) the remaining 15% comes from institutional support provided by Saint John's University and Saint John's Abbey, royalties, interest income, etc. HMML has an active fundraising program with a full-time director of development and external relations, assisted by two part-time support staff. HMML recently completed a successful capital campaign that focused on renovating HMML's 40-year-old facility (renovation completed in 2014), continuing digitization efforts at sites around the world, developing vHMML, and continuing to build endowment. The last effort is being addressed primarily through a planned giving initiative.

vHMML Reading Room. Maintaining vHMML Reading Room requires the full-time attention of two staff members: 1) the Systems Librarian and 2) the Metadata Librarian. The Systems Librarian manages the servers and databases, is the liaison with CSB/SJU IT Services, the principal contact

with the software developers, and does some coding. This new position was created in January 2014 at the beginning of the vHMML 2.0/Reading Room project, funded by the Luce grant for Reading Room (annual expense: \$70,000 for salary and FB). Funding from Luce ends in December 2017. This proposal requests funding to continue the position for an additional 20 months, through August 2019.

The Metadata Librarian is responsible for content preparation for vHMML Reading Room. She controls authority lists for the cataloging schema and vets all records created by subject-expert catalogers. This position is currently funded by a grant from the Andrew W. Mellon Foundation that ends in June 2018. Continued funding for this position is the focus of our request for a Sustainability Match (see below).

Other HMML staff play important roles in sustaining vHMML and vHMML Reading Room. The Director of Digital Collections and Imaging manages the image content of vHMML Reading Room, complementing his management of image workflow from the point of capture at fieldwork sites to data archiving at HMML (see the Data Retention Plan). HMML curators and catalogers are involved in the creation of metadata and in conceptual development of vHMML as part of their general HMML duties.

Software development is provided by an outside contractor, Solution Design Group (SDG) of Golden Valley, Minnesota (<https://solutiondesign.com/>). Software development for vHMML and then vHMML Reading Room was funded by grants (IMLS, Luce, and Mellon), with additional expenses covered by gifts that had been internally designated for the HMML renovation in 2014 but were not needed for that project. Those funds have now been exhausted. Further development of vHMML Reading Room will require new grant support, as the previous grants are not renewable.

Most financial needs for sustaining vHMML are complementary to other HMML activities and thus already built into the HMML budget and funded by the revenue streams described above. The exceptions are the positions of Systems Librarian and Metadata Librarian. Long-term sustainability of vHMML Reading Room will require that these positions become permanent. In the interim, grant funding and the Sustainability Match will provide HMML time to anchor this position in long-term budgetary and fundraising models. The challenge provided by the matching requirement will help HMML staff highlight for the Board of Overseers and other HMML donors the ongoing need for support of this essential aspect of HMML's mission.

### **Human Resources**

HMML is developing a long-term staffing model to integrate the digital and curatorial/cataloging aspects of its work more effectively, and to deepen in-house skills in database and systems administration. Like most research libraries, HMML no longer considers technology to be auxiliary to, or separable from, core curatorial and research activities. Identifying needed skills, refining position descriptions, and succession planning are part of a forward-facing approach to sustaining vHMML and the digital collections which populate it. For example, the Systems Librarian (a position that is acquiring more of the traits of a System Administrator) needs a range of skills to be able to liaise effectively with the contracted software developers and with CSB/SJU IT Services staff. The Metadata Librarian is assuming more leadership over the entire cataloging workflow. Other key staff must be conversant with the conceptual framework of vHMML as well as with

practical tasks that must be done on a regular basis. Workflows are being documented in case key staff members become unavailable.

### **Sustainability Match**

HMML is requesting \$ (b) (6) in matching funds to sustain the Metadata Librarian position. This position is currently funded by a grant from the Andrew W. Mellon Foundation that ends June 30, 2018. The Metadata Librarian (Eileen Smith) is the crucial staff member for the uploading of items to vHMML Reading Room, as she controls authority lists for key aspects of the cataloging schema, vets all records created by subject-expert catalogers, and participates fully in the planning and development of vHMML.

HMML has had previous success with matching opportunities, including the successful completion of fundraising to match an NEH Challenge Grant awarded in 2003 to endow HMML's Malta Study Center. The prominence of digital humanities in the 2020 Strategic Plan and the successful launch of vHMML 1.0 and 2.0 give us confidence that we will be successful in meeting the match requirement. The matching effort will also be an important demonstration to HMML's Board of the need to transition funding streams for technology-related positions from grants to other revenue sources.

### **Summary**

Funding from the NEH for vHMML 3.0 is essential for immediate needs to keep the platform active and responsive to users and partner projects, as well as for bridging project funding to a model that will combine predictable revenue streams with periodic extraordinary support from foundations and individual donors for major upgrades.