

Optimizing Data Collection for Heritage Preservation

Fenella G. France, Preservation Research and Testing Division, Library of Congress, Washington, DC, USA

Meghan Hill, Preservation Research and Testing Division, Library of Congress, Washington, DC, USA

Abstract

As institutions face reduced staffing with the same number of requests for research and services, being able to optimize the collection of data and workflows becomes critical. Many workflows in analyzing cultural heritage objects become standard practice, but we often accept this as “the new process” without fully documenting and reviewing the effectiveness. Scientists in the Preservation Research and Testing Division (PRTD) have developed efficient workflows that utilize non-invasive optical imaging and spectroscopies and have added another component that greatly improves accuracy to the activities, fully characterized reference samples. Within PRTD, the Center for Heritage Analytical Reference Materials (CHARM), is a collection of heritage scientific non-collection samples that have been characterized by a range of imaging modalities and spectroscopies, so that identification of inks, pigments, various substrates and additives can be readily identified from the reference data that has been captured. Working together in teams, pre-identifying regions of interest (ROIs) and having pre-characterized samples available to confirm composition of heritage materials being analyzed has led to a productive outcome.

Introduction

During the pandemic, while still trying to understand the implications of the virus remaining on surfaces, it was clear that the previous approach to analyzing collection materials in the lab, where one person worked on a project and then brought in others as needed, was not going to be sustainable for projects and research to move forward. Since 2020, our staffing has continually decreased with positions not being refilled, so this was and remains a challenge. Further, at this time PRTD was in the midst of a large externally funded paper research project that had deadlines, and, to add another demand, the Library was dramatically increasing its exhibition spaces, so the number of demands for testing of exhibit materials, especially micro-fade testing, was expanding.

The problem PRTD faced was how to maintain services, demands and juggle priorities while still being effective and completing the research needed for all the research and testing areas PRTD staff were involved in. And, given the pandemic, what workflow would keep people safe, while still meeting deadlines. The areas of PRTD services included collections research, materials testing, researcher requests, and an expanding exhibition program, as well as work on an ongoing paper research project. PRTD had technician staff who were still becoming proficient in techniques, and who could capture spectral data with portable instruments, but could not always necessarily easily interpret that data. The need for efficient services led to a review of the workflow processes, and marked a transition away from access to, and reliance on, highly experienced scientists who could rapidly interpret spectra based on visual inspection and prior knowledge (e.g., identifying a material by characteristic features or recognizing the absence of diagnostic peaks).

Three areas of need were identified: training of staff, transfer of knowledge and expertise, and maintaining research services. Each of these required interactions between staff, but how to do this effectively, and safely, and achieve internal Library deadlines for research and testing services? The workflow process below will focus on research of collections, rather than the materials testing program.

Background

At the onset of the pandemic, PRTD convened virtual meetings via Zoom to bring together their team of scientists and collaborators to assess research priorities and determine paths forward. While such discussions were already part of standard practice, the absence of full on-site staff necessitated an adaptation of established workflows. In particular, greater emphasis was placed on developing structured and reliable modes of communication to replace the informal, in-person exchanges that had previously occurred somewhat organically. It was more effective to move the research process from a project being assigned to a singular staff working on a project, to a more structured team approach. This would allow multiple staff to work on the same research project but could come into the labs to use instruments and lab space on different days.

The phased approach enabled PRTD and the Library to utilize personnel in different ways – those who were proficient with instrumentation could capture the data (mostly non-invasive spectroscopies for collection items), and the subject matter expert (SME) or prime user for that instrument, could then complete the interpretation, even offsite if necessary. Accordingly, an (ongoing) issue that arose was getting access to proprietary instrument software on the Library network. Fortunately, we had laptops in the laboratory and many manufacturers were sympathetic to the challenges and allowed us to load multiple versions for a specific period of time.

Given the involvement of multiple staff members across individual projects, the research lab adopted a more proactive and structured approach to project coordination, ensuring clear communication of procedural steps, timelines, and workflows. Each project was initiated using a standardized template of a written outline of planned actions, proposed instrumentation, and concise documentation featuring images of the object with clearly defined regions of interest (ROIs) that had been established in discussion with other partners and collaborators.

Our standard process for characterizing and understanding a collection item was to start with multispectral imaging (MSI) to “map” condition and chemical composition across the entirety of its surface. Once the main image(s) of the collection item has been processed with principal component analysis (PCA), the resulting false-color, or pseudocolor, image allowed staff to quickly assess how many different chemical compositions/ pigments/ inks /material conditions etc. existed. These were used to determine the regions of interest for subsequent point source spectroscopies. These techniques are not limited to, but usually include most of the

following: microscopy, X-ray fluorescence spectroscopy (XRF), fiber optic reflectance spectroscopy (FORS), Fourier-transform infrared spectroscopy (FTIR) and Raman spectroscopy. Since many collection items have mixtures and layers of pigments, inks and other compounds, these complementary techniques allow us to capture both organic and inorganic information, especially important when we are seeing different interactions between the substrate (parchment – animal skin, or paper – cellulose) and the media – inks and pigments [1] [2].

Standardized documentation was the crux of efficiency and consistency across projects with multiple users. It began with standardized data capture parameters via standard operating procedures (SOPs) along with quality control checklists to ensure data integrity. Clear (and predetermined) file-naming conventions made it easy to identify the region of interest associated with each dataset, particularly when integrating complementary data from multiple instruments. A shared research report template, accessible to all staff via a centralized drive, allowed team members to input data in a uniform format. improved communication, reduced ambiguity, and supported more efficient, well-organized project completion.

Workflow Process Development

1300 Magna Carta and Carta de Forest

Our first critical project was a high-profile document with a very tight timeline. We were to analyze every component we could of the Sandwich copies of the 1300 Magna Carta and Carta de Forest with colleagues from the United Kingdom (UK). We had one day of viewing the document before it came into the lab and only 1-2 days of access before it returned to the UK.

In determining the workflow process, staff began by having a collaborative discussion in advance with curators and conservators, questioning the document and determining the critical locations and research areas to be explored:

- Characterizing the materials - why is it in this condition?
- Can we understand how to stabilize and preserve it?
- Can we image through stains and find the original text?
- What can each technique add to build upon our knowledge?

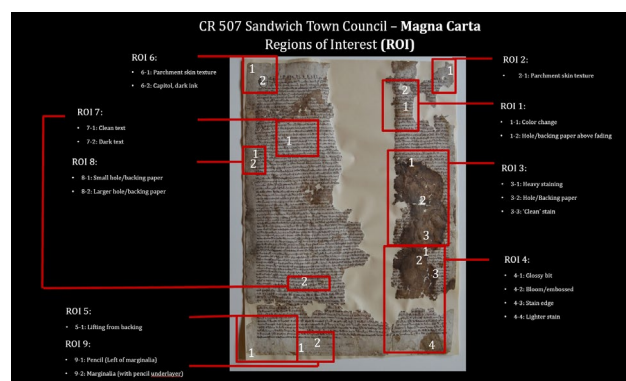


Figure 1. Regions of Interest for Examining the Magna Carta

The day before the document came into the lab, we used a small portable microscope with ultraviolet (UV), visible (VIS), and infrared (IR) capabilities to confirm the regions of the document that would best answer those research questions. We created a matrix of instrumentation to determine which were needed for each ROI and in what order – we'd begin like usual with (MSI) to map differences in chemical response across the entirety of the documents. However,

we didn't have the same amount of time to conduct and review PCAs before moving onto other spectroscopies, so two team members tackled image processing, prioritizing ROIs that would specifically benefit the next instruments to be used.

We then incorporated other non-invasive spectroscopies to add specific knowledge that could build upon one another to comprehensively answer the research questions. XRF was used to find trace metals and specific components of the inks as well as elements on the parchment; FTIR to look for degradation peaks in the parchment and use paper modelling to confirm the composition of the backing paper; and FORS for substrate, media, binding, and potential pigmentation aspects.



Figure 2. Colleagues and PRTD Heritage Scientists working on Magna Carta Documents with Fourier Transform Infrared (FTIR) Spectroscopy.

PRTD staff worked efficiently and effectively between the two fragile documents, working to answer all questions in the ROIs and utilizing all the space in the lab to assure safety of staff. Apart from answering the characterization and condition questions, two unexpected aspects of the documents were revealed to us in the process of data capture. We were able to determine the source of staining to be rodent urine, though we were not prepared to assess what type of rodent, to the dismay of the inquisitive conservator. Though more interestingly, the utilization of reference samples uncovered that the backing paper wasn't the type of paper it was originally believed to be.

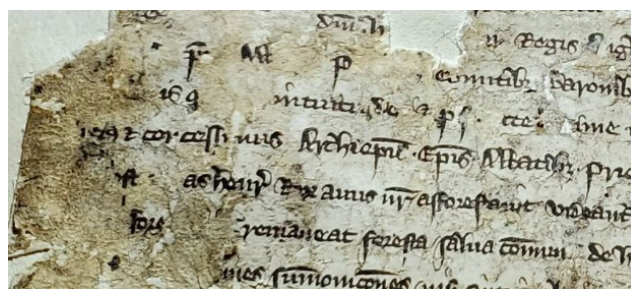


Figure 3. Section of Carta de Forest Parchment.

The backing paper on the two documents had been attached to protect the fragile parchment from degrading further but flaking and possibly previous treatments had caused changes to it. PRTD scientists were concerned that the backing paper might be off gassing and exacerbating degradation issues, so impromptu gas chromatography mass spectrometry (GC-MS) thermal desorption was undertaken to assess the presence of any volatiles.

Working within a limited access window underscored that success hinged on proactive planning, clear communication, strict project scope management, and a collaborative team dynamic that valued each area of expertise to achieve a complete dataset and holistic analysis.

Scientific Reference Samples

A critical aspect in the development of our workflow was the need for reference samples of heritage materials. Our analyses of collection objects rely solely on non-invasive techniques. This means no sampling and no chemical testing to identify complex materials. Having known reference samples greatly increases efficiency of data interpretation, especially when analyzing multiple layers and mixtures of media, all of which show up in a single spectrum and need to be separately discerned.

The Center for Heritage Analytical Reference Materials (CHARM) [3] is a *collection* of scientific samples of heritage materials. These materials are not part of the Library's collection and thus are not bound by the limits of non-invasive analysis. While their chemical testing does provide crucial degradation and preservation information of these mediums, the more pertinent aspect is that they are also characterized by the same suite of non-invasive instruments used with collection items. This helps streamline the identification of unknown materials in collection items by comparing their spectra to known inks, pigments, substrates, additives, and a plethora of other media.

Accumulated over more than 16 years (and continuously expanding), the breadth and diversity of our reference collection make it highly likely that corresponding matches to unknown materials already exist in CHARM. In addition, if time permits, we can age samples in a slow controlled manner to better replicate historic collection materials.

In the case of the Magna Carta analysis, a prior treatment report identified the backing paper type and a corresponding sample of that type of paper existed within CHARM. However, analytical results showed that the backing paper did not match this reference, indicating that a different paper had been used on the document. Through chemometric paper modeling [4], it was nevertheless still possible to predict the composition of the backing paper, and thankfully, the analysis indicated that it was not contributing further damage to the historic documents.

1560 Portolan Chart

Our workflow was further refined during the analysis of a 1560 Portolan Chart which presented us with a project requiring the identification of lots of different media in varying states of degradation. Portolan Charts are early nautical charts, usually on parchment, and often lavishly decorated. This Portolan was a map of the southeast coast of North America, created in Portugal c. 1560 by Bartolomeu Velho. It had been cut down from a larger map and likely bound into a book at some point in its history.

As can be seen in Figure 4 below, there were many ROIs required to confirm both the material compositions as well as the consistency across the object. Normally, MSI and PCA would attest to the consistency of the materials rather quickly and easily, but this was not possible across the entirety of the map due to its large size. There was extensive fading, loss and flaking or abrasion of the inks and pigments. In some cases, only stains remained as evidence of prior inks or pigments, thus it would be critical to detect the presence of trace elements that may no longer have strong indicators for identification. The chart was hinged to a backing board which made the verso inaccessible as well as adding some complexity in testing since the signal of the backing would need to be subtracted from the spectra.

Conversations with curators and conservators identified questions that would need to be addressed, including the identity and date-appropriateness of the inks and colorants, any evidence of later retouching, the nature of the white accretions across the top,

fluorescing regions and the condition of the darker regions of parchment. MSI, XRF, FORS, FTIR and multimodal microscopy were required to answer these questions.

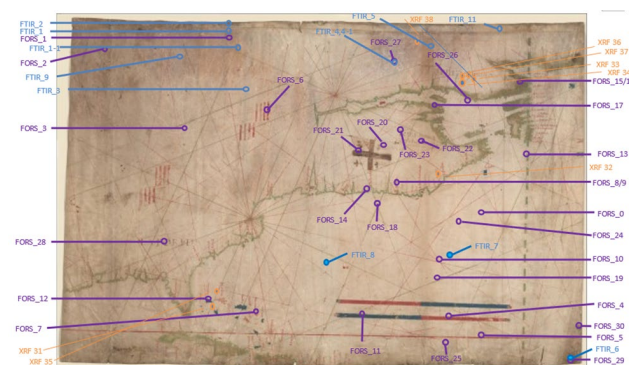


Figure 4. Regions of Interest for the 1560 Portolan Chart

The analysis was complex, not just in the sheer number of data points, but unexpectedly by the nature of the materials. The yellow pigment, seen in the middle of the map on the flags and some text, contained iron, sulfur and zinc, a strange combination of elements for a yellow pigment. With the aid of multispectral imaging which showed the material's transparency in the infrared and a red response in false-color infrared image processing, it was determined that the yellow pigment was in fact residual staining from iron gall ink that had flaked or abraded away.

Multispectral imaging revealed the use of three different greens across the map (Figure 5). The first green is seen in the coastline and vertical dotted line, the second as a wash blended out from the coastline, and the third as the green rhumb lines and outlines of the blue islands. Each green was copper-based, but with different signal strengths and trace elements.



Figure 5. Color (left) and false-color processed image (right) showing unique responses between areas of green pigment (coastline/dotted line vs. wash vs. rhumb line), detail.

“Green 1” had distinct copper peaks and no other characteristic elements in the XRF spectrum, most likely verdigris (copper acetate) based on the location of the absorption max seen in the FORS reflectance spectra. “Green 2” had a significantly yellower/browner hue and exhibited a less pronounced FORS absorption max. There was indication it could be verdigris in an early stage of deterioration or verdigris mixed with a yellow organic colorant. Examination under magnification showed that this wash had a mottled appearance, more suggestive of verdigris degradation than of a mixture. “Green 3” had an even less pronounced absorption in the FORS spectra and a weak XRF copper peak. But closer examination indicated a peak present for tin and small peaks for lead, silica, and iron. This copper-based green, still likely verdigris,

appeared to be mixed with a small amount of lead-tin yellow type II, $\text{Pb}(\text{Sn}, \text{Si})\text{O}_3$, a pigment popular from the 13th – 17th centuries.

This example highlights the need for expert users to carefully review and interpret the data and not make assumptions based on one instrumental technique, or even from the presence of specific trace elements. The greens required a complex interpretation, since the small additional metal features might have been overlooked by less experienced users.

Block Books, c. 1470

An exemplary project that clearly demonstrated the effectiveness of our developed workflow was a long-term, in-depth analysis of the materiality of block books. Block books are books where both text and illustrations were carved entirely onto single wooden blocks, printed, and then hand-colored. Their popularity was due to providing an inexpensive, visually driven resource used as religious education for the general illiterate public. However, very few remain, and even fewer as full volumes.



Figure 6. Example page from each of the five primary block books analyzed in this project, showing consistency of pigment palette but variety of application.

The project was grounded in well-defined research questions developed by curators with deep historical expertise, providing a strong intellectual framework from the outset. This guidance was especially valuable in selecting which books to analyze and in recognizing the unique nuances of each volume and its contents. As the work progressed, the project evolved through an iterative process shaped by ongoing curatorial input, which helped refine the scope, manage an increasingly large body of data, and direct attention toward lines of inquiry that meaningfully advanced scholarly understanding of the collection.

Analysis began with multispectral imaging in multiple modalities which helped establish how these books were made, their original provenance, and to begin characterizing some of the materials before moving to spectroscopy-based instruments. Low-angle raking illumination enhanced the impressions the woodblock left in the paper, which allowed assessment of the objects' construction techniques. Transmissive MSI and PCA were critical for capturing clear, detailed images of watermarks within pages of heavy content. UV reflected illumination aided characterization from the fluorescence of some pigments and IR reflected illumination characterized the ink.

Spectroscopy techniques worked in concert to build a holistic understanding of the media used to create these books. Elemental analysis with XRF, for example, indicated the presence of copper in a green pigment, while FORS allowed for more specific identification as verdigris rather than malachite, and FTIR would then contribute binder and sizing information. Although the palette was remarkably consistent, there were a number of unexpected results that needed extensive assessment and interpretation, such as an interesting use of verdigris and unconventional choices of binders.

Microscopy further refined these interpretations by revealing particle characteristics, helping to confirm in several cases whether pigments were used alone or as mixtures. For example, when *Passio*

Christi was examined under the microscope, both fluorescent and non-fluorescent particles were seen in the red robe. This suggested the possibility of a mixture, such as brazilwood (indicated by its absorption properties) and lac (indicated by fluorescence). Examination of a series of historically re-created mixtures of brazilwood and lac, in varying ratios, enabled confirmation that the red in the robe was simply a brazilwood lake. The presence of both fluorescent and non-fluorescent particles was due to a mixture of free brazilwood dye and laked brazilwood, which was common during production as it cannot be expected that every free dye molecule in a solution will attach to a substrate during reaction.

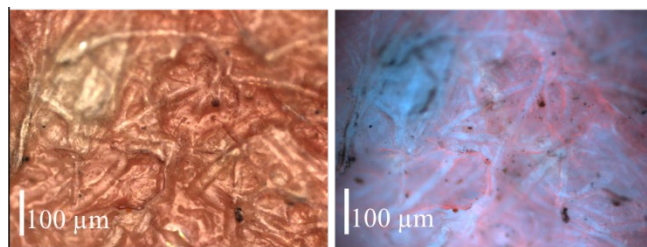


Figure 7. *Passio Christi* red robe microscope images, brightfield (left) and fluorescence (right).

The use of CHARM was critical for confirming the identification of many of the pigments used in these books. Pigment reference samples were created from historic recipes in a range of organic and inorganic materials specific to the palettes in the block books. This included recreating various extraction and application methods of organic media, mixtures, and more.

In this extended project, instrument portability supported more effective time management and use of staff resources by facilitating analysis in the curatorial reading room, eliminating the need to routinely transport materials to the laboratory.

This project demonstrated how an integrated, well-coordinated workflow can successfully support complex material analysis, even across an extended timeframe. The combination of curatorial expertise, iterative research design, and complementary analytical techniques enabled both efficient data collection and nuanced interpretation. Together, these approaches produced a comprehensive understanding of the block books' materiality while advancing broader scholarly knowledge of these rare objects.

Conclusions

The discussion above has illustrated how a focused approach and a thoughtful re-assessment of workflows can greatly assist reduced resources, while assuring preservation remains the prime focus. This work highlights that an effective analytical workflow depends on deliberate planning, clear communication, and strategic collaboration. Establishing research questions and regions of interest in advance, along with standardized documentation practices, supports efficient time management and focused data collection. Ongoing dialogue with curators provides essential historical context and ensures that analytical efforts remain aligned with meaningful scholarly questions. At the same time, balancing team roles by leveraging staff for efficient data capture while engaging subject matter experts for complex interpretation maximizes both productivity and analytical depth. The integration of complementary techniques, including imaging, spectroscopy, and microscopy, enables a more holistic understanding of materials than any single method alone. Finally, the use of well-developed reference collections allows for rapid and confident identification of

media, streamlining interpretation and strengthening results. Together, these elements demonstrate a workflow that is both efficient and rigorous, capable of supporting high-quality research under a range of project conditions.

The authors would like to acknowledge and thank their colleagues in the Preservation Research and Testing Division who have contributed their knowledge, time and expertise to the continued refinement of divisional workflows.

References

- [1] Biolcati, V., Wilson, M., Fiddymment, S., Unitt, R., Connelly Ryan, C., Hoffmann, A., Gillis, J., France, F., Ó Macháin, P., and Iacopino, D., “The Book of Uí Mhaine: An Interdisciplinary Analysis of the Materiality of the Gaelic Manuscript Tradition”, *Heritage* (2023) 6, 5393–5409, <https://doi.org/10.3390/heritage6070284>.
- [2] France, F.G., “Heritage Science – Perspective, Provenance, Preservation”, *Archiving IS&T*, Washington D.C., April 2024.
- [3] HSAI Webpage, ICCROM, Center for Heritage Analytical Reference Materials (CHARM), Preservation Research and Testing Division, Library of Congress, USA
<https://samplearchives.iccrom.org/oggetti/41712-center-for-heritage-analytical-reference-materials-charm?i=0>
- [4] France, F.G., and Davis, A., “Preservation Equity: Decision Making and Data”, *Archiving IS&T*, (2024) <https://doi.org/10.2352/issn.2168-3204.2023.20.1.33>

Author Biography

Dr. France, Chief of the Preservation Research and Testing Division at the Library of Congress, researches non-invasive techniques and integration and access between scientific and scholarly data. An international specialist on environmental deterioration to cultural objects, her focus is connecting mechanical, chemical and optical properties from the impact of environment and treatments. She maintains collaborations with colleagues from academic, cultural, forensic and federal institutions through her service on a number of international bodies.

Meghan Hill is a Preservation Science Specialist in the Preservation Research and Testing Division at the Library of Congress. She specializes in spectral imaging and collaborates with cultural heritage institutions to integrate multispectral workflows into digitization and preservation programs. Her work applies noninvasive optical technologies to support collection care and advance the understanding of historically significant materials.