CASE 19

Dangerous Liaisons: Working Collaboratively with a Donor on a Mass-Digitization Project at UNC Charlotte

AUTHORS: RITA JOHNSTON
JOSEPH NICHOLSON
Special Collections and University Archives
J. Murrey Atkins Library
University of North Carolina at Charlotte

CASE STUDY DATE: June 2018

ISSUE: This case study focuses on a project involving the digitization of a large collection of photographs in Special Collections and University Archives at Atkins Library at the University of North Carolina at Charlotte in which the donor has been involved in every stage of the work, from providing input on the digitization of the images to creating detailed metadata for them. The study discusses the benefits and drawbacks of this intensive collaboration between archive and donor, the tools and workflows used for collecting metadata from the donor and digitizing the images, and the special human challenges of working closely with a donor on a complex project.

KEYWORDS: Commercial photography
Digitization
Digital asset management systems
Donor relations
Metadata
Photographic archives
Repositories
Background

Established in 1949 as Charlotte College, the University of North Carolina at Charlotte is one of the youngest institutions in the University of North Carolina’s sixteen-campus system and—with 28,000 students in 2018—the fastest growing. The university’s nine colleges offer twenty-three doctoral, sixty-four master’s, and 140 bachelor’s programs. The J. Murrey Atkins Library, a ten-story building at the center of campus, is a large research institution with two million volumes and ninety-six faculty and staff.

Until recently, Atkins Library’s special collections department devoted itself to husbanding a modest collection of rare books and local history materials. In the past few years, the recently renamed unit, Special Collections and University Archives (hereafter referred to as SCUA), has redirected its energies to strengthening its archival programs and launching mass-digitization initiatives. It has parlayed retirements into new archivist positions, purchased new digitization equipment, and acquired an open-source repository built on the Islandora platform. Despite recent faculty hires, the unit lacks the layers of specialized staffing and the robust budgets of special collections departments at larger institutions, as well as their technical expertise.

The most notable of SCUA’s recent projects is the digitization of the Bryant McMurray motorsports photographs, a vast collection of more than 450,000 images of motorsports events and personalities donated by a noted local motorsports photographer. Capturing the speed and danger of NASCAR races, celebrated drivers at play and in the heat of action, press conferences, parties, and motorsports fans, the collection charts the history of the sport, which has close ties to the Charlotte region, from the late sixties to the present day. The collection contains black-and-white and color photographs taken with both film and digital cameras in a variety of formats.

The project was conceived on ambitious lines. Rather than pare the projected digitized collection down to a more manageable size, SCUA committed itself to begin digitizing all the donated images in more or less the order in which the donor had arranged them, without setting forth a plan for selection. And instead of opting for a collection-level approach to metadata, it proposed to describe all of the images at the item level, assigning access points with controlled vocabulary to every person identified and every race, racetrack, and track location depicted in the images. The intent was to create a vast searchable database that would enable users to find images of drivers and other motorsports personalities and events with pinpoint accuracy through faceting and hyperlinked controlled value terms.

Working closely with the donor himself, Bryant McMurray, promised to be as unique a challenge as digitizing the images and creating metadata for the collection. Talkative and opinionated, a skilled raconteur, McMurray is deeply interested in every aspect of the project that is centered on his life’s work, from the digitization process to copyright statements to the metadata. Though he had committed to deliver all of the photographs to SCUA, the donor planned to use the image collection as his working archive and as a commercial showcase for the images at the same time that SCUA was digitizing them.
Most significant of all, the library had contracted with McMurray to provide metadata for the images, principally in the form of identifications of people depicted therein but possibly including other information as well, such as the make and model of racecars and captions for noteworthy images.

The difficulties of the project seemed considerable as staff began work. Early on, staff confronted the problem of determining what metadata elements the donor could reasonably be expected to provide and with what tools. Training a donor unfamiliar with library standards for description to create metadata was an obvious hurdle. There was the challenge of devising an efficient method of creating a huge volume of item-level metadata for the images in MODS (a standard that SCUA had not used before), mixing in the donor’s metadata, and ingesting the records into an unfamiliar repository. In addition, there were misunderstandings about the purpose and speed of the project that were almost certain to arise between a commercial photographer principally interested in rough and ready online access to his photographic archive and SCUA staff concerned with ensuring that the project met rigorous archival standards and specifications.

Case Methodology

Creator, Purpose, and Content of Collection

The Bryant McMurray Photographs and Personal Papers collection primarily comprises photographic materials created from 1969 to 2014 by Bryant McMurray, a 50-year veteran of photojournalism, marketing, and public relations in motorsports. After conversations beginning in 2008 between McMurray and departmental leadership in SCUA, UNC Charlotte entered an agreement with McMurray to obtain the collection and began acquiring McMurray’s materials in 2014 in multiple accessions.

During the initial talks with McMurray, he expressed a strong interest in having his analog photographic materials digitized, and assurances that the library would digitize a large portion of his film collection led to his decision to enter a donor agreement with UNCC. McMurray hoped that this partnership would lead to the creation of a vast repository of his digital content and had high expectations for the marketing and promotion of this collection, which would in turn increase demand for licensing requests for his photography.

SCUA leadership changed over the years leading up to the donor agreement, but the unit’s motivation for acquiring Bryant McMurray’s motorsports photography remained the same: the economic, cultural and historical ties the Charlotte region has to motorsports. As of 2011, the motorsports industry brings about six billion to North Carolina’s economy annually and creates over 20,000 direct and indirect jobs.1 The Charlotte region is the home of the Charlotte Motor Speedway, the NASCAR Hall of Fame, prominent NASCAR racing teams including Richard Petty Motorsports and

---

Hendrick Motorsports, and multiple NASCAR offices—UNC Charlotte offers a Motorsports Engineering concentration through the William States Lee College of Engineering. Charlotte also serves as a significant “memory community,” or space for fans to engage with a shared narrative of the history of racing.²

McMurray emerged from talks with the library with a strong sense that SCUA would be able to digitize more of his collection and devote more resources to its promotion than it realistically could with current resources. Though such commitments were not specified in the donor agreement itself, the library may have unwittingly played a part in encouraging high expectations in the donor that were almost certainly to be disappointed as the project moved forward and optimistic prognostications of swift progress collided with the unpleasant realities of recalcitrant equipment and budgeting and staffing constraints. These expectations led to tensions in the donor-library relationship that have never been fully resolved.

The Nature of the Records

The bulk of materials in Bryant McMurray’s motorsports collection are 35mm negatives cut in strips of six or fewer, but also include medium-format negatives, mounted color slides, and about 200 GB of born-digital images on a hard drive. McMurray held on to most of his material during the first several years of the project, transferring materials to the library periodically in small accessions, but conveyed the remainder of his motorsports negatives and a sampling of his slides to SCUA in November 2017. He continues to commercially license his photographs, so was reluctant to part with his entire collection before SCUA had established that it could provide access to a complete inventory of his work and fulfill requests on demand for images.

McMurray organized his negatives through a homegrown paper-based indexing system beginning early in his career, and SCUA deviated very little from the donor’s arrangement. He assigned sets of negatives taken during a motorsports racing event a unique number and kept the sleeves associated with the event in a labeled envelope. He maintained a typed alphabetical index of events and event numbers on notecards, and organized film strips in envelopes with the event name, year, and number. McMurray created contact sheets from negatives for all of the events which he used to locate and select images for a particular purpose.

Custodianship/Restrictions on Use

The donor agreement signed by Bryant McMurray and UNC Charlotte stated that McMurray would retain copyright to all materials in his motorsports photography collection until his death, upon which time full copyright would transfer to UNC Charlotte. It also stipulated that McMurray would transfer his materials in batches, and

that UNC Charlotte would digitize part of his analog collection. To prevent unauthorized use of McMurray’s collection and preserve his ability to financially profit from his work, all content made digitally available would need to have reasonable protections in place to block downloads.

The agreement additionally specified that for the first two years of the project the library would pay McMurray to supply metadata for his images, placing the library in the ethically awkward position of remunerating him for work that would increase the exposure and hence marketability of his own collection. The agreement did not indicate the type or extent of metadata he would provide or the tools he would use. Filling in these lacunae and ensuring that the project had a sound academic basis no matter what material benefits it might provide to the donor were to become some of the most challenging aspects of the project as it moved into its production phase.

**Long-Term Preservation of Digital Objects and Access**

SCUA had been publishing digital collections in CONTENTdm since the 2000s, but the department did not have an integrated preservation and access system. Also, the CONTENTdm repository was not a viable option for publishing certain types of content, including SCUA’s oral history interviews and electronic theses and dissertations.

After exploring options, Atkins Library decided to build a new digital asset management system utilizing the Islandora open source framework, which offers a balance between out-of-the-box and boutique digital asset management systems. The new repository was intended to be a home for content including digitized and born-digital materials held by SCUA, born-digital local government documents, and electronic theses and dissertations. At the same time, Atkins Library and campus ITS were moving away from permanently storing digital objects on local servers. The new repository would be integrated with Amazon Web Services (AWS), so that upon ingest, digital objects would be copied to AWS Glacier for long-term storage.

Atkins Library’s Technology and Digital Strategies team installed and configured the library’s Islandora repository, branded Goldmine, by late 2014. While they continue to work on optimizing and improving all aspects of Goldmine, the new repository was operational by the start of the McMurray motorsports digitization project.

**Pilot Project**

As promises to swiftly and exhaustively digitize Bryant McMurray’s photographic motorsports collection were a motivating factor in his donation of the collection to Atkins Library, SCUA embarked on a pilot digitization project from August 2014 to September 2015. The unit acquired a representative sample of materials from McMurray, primarily 35mm negative strips but also some medium format negatives and 35mm mounted slides.

From the digitization pilot study, the Digital Production Librarian and the Processing and Technical Assistant learned that existing equipment in the department was inadequate for
digitizing McMurray’s collection of motorsports photography. Pilot project staff used a flatbed scanner with plastic transparency holders, but the holders couldn’t flatten out the film enough to get a clear scan of the images. Additionally, using a flatbed to scan at the highest resolution and the post-scanning workflow was incredibly time-consuming. Because of these issues, SCUA was able to successfully make the case for purchasing a Hasselblad Flextight X5 scanner, which was acquired by the department in summer 2015.

One of the primary objectives of the pilot project was to test out methods of collecting metadata from the donor. McMurray’s previous experience with spreadsheets and SCUA’s heavy reliance on that tool in previous digital projects suggested that Excel was the most appropriate medium for gathering his data. The initial plan was to send him metadata spreadsheets for the images in each sub-collection within the larger collection of photographs via a Dropbox account. Working from small images loaded directly into the left-most column of the spreadsheets, the donor would enter identification of people, makes and models of cars, and free text comments in columns reserved for his data. Dropdown menus preloaded with driver names would assist him in his annotations. Unfortunately, the donor found it difficult to read the small script in the dropdowns and click the name terms into place, and he was unable to identify the makes and models of cars in the small images. Moreover, the cramped Excel format discouraged him from creating the detailed captions for which project staff were hoping.

The pilot project made it clear that metadata staff had inflated expectations about the extent of the metadata the donor could supply and the tools they had devised to collect it. Scaling them back, they decided to ask the donor only for identification of persons and occasional free-text notes that metadata staff would then expand into full captions. They also began to think about the possibility of finding tools other than Excel to gather his data. Yet it would be more than a year before a suitable method was developed.

**Current Metadata Creation Process**

The photograph collection is arranged in smaller sub-collections that document particular racing events, such as the 1977 Daytona 500 or the 1975 National 500. Each sub-collection contains somewhere between 50 and 300 images on average. For each racing sub-collection, project staff create an Excel metadata spreadsheet containing columns representing some 20 elements that correspond to MODS elements such as title, extent, subject, abstract, etc. They rely heavily on the MODS User Guidelines for version 3 developed by the Digital Library Federation and the Library of Congress but use MODS version 3.6 in metadata records created for the project. An Excel extension program, Kutools, is used to load JPEGs of the digitized images into the far left column of each spreadsheet. The images serve as useful reference points as staff assign subject and names terms.

Staff use several approaches to accelerate metadata creation. In lieu of devising an individual title for each image, they use a formulaic title based on the name of the race

---

(e.g., “1972 Coca-Cola 600 events”) to which is appended the incrementing phrase “image 1,” “image 2,” “image 3,” etc. for the remainder of the photographs in the sub-collection spreadsheet. Instead of Library of Congress Subject Headings, project staff prefer topical terms from the FAST (Faceted Application of Subject Terminology) vocabulary developed by OCLC, which student workers can apply more swiftly. These topical terms, along with controlled vocabulary for personal name subjects derived from Library of Congress authorities and local authority files, are placed in columns of the second tab of the spreadsheet and clicked into place in the appropriate columns in tab one using the Excel data validation functionality. Some tweaks to the Excel coding enable staff to insert more than one FAST topical and name term per cell.

Figure 1. Controlled vocabulary terms and uniform resource identifiers (when available) are inserted into Excel metadata spreadsheets with dropdown menus.

Once the metadata has been created and proofread, the spreadsheet is saved as a text file and then uploaded to OpenRefine. Prior to this stage, a template record is created in MODS in Notepad that contains every element and attribute used in the project. After applying the template record to the metadata in OpenRefine, staff export a collection-level record containing individual XML MODS records for every row in the spreadsheet.

At this stage, two simple XSLTs strip out detritus from the resulting file, split the large collection-level file into individual MODS files for each image, and rename the resulting files after the filenames of the images. The image files and metadata files are then united in an ingest directory on the NAS. UNC Charlotte developers have created a backend pipeline from the NAS to Islandora that allows project staff to ingest collections into the repository in a fraction of the time that ingest through the web interface would require.

**Current Donor Metadata Process**

Once project staff have uploaded the images and metadata files to the repository, they notify the donor that a new sub-collection is ready for his annotations. The donor then logs into the Islandora repository and completes his metadata for the sub-collection using an annotation tool that digital initiatives staff at Atkins developed for his use in 2016. The sub-collection remains under embargo until staff proofread the donor’s metadata.
Based on Drupal form technology, the tool enables the donor to fill out a simple form for each photograph in a specified collection. The form includes boxes for names of people and a larger text box for free-text notes. The zoom capabilities of the Islandora image viewer allow him to see the images with a clarity and detail that were impossible in Excel thumbnails, encouraging more detailed and accurate metadata.

When the donor completes each form, he clicks on a “submit” button and toggles to the next image in the collection. After he finishes his annotations, project staff ingest them into an XML metadata form they have created in Drupal and edit them to conform to project standards. Name forms entered by the donor are converted into authorized project forms. Free-text notes are edited into complete sentences or deleted as appropriate. When project staff finish editing and proofreading the donor’s metadata for a particular collection, the embargo is lifted and the sub-collection is made public.
Access Mechanisms

The digital Bryant McMurray motorsports collection is available through Goldmine, the public view of the Islandora repository. The Digital Public Library of America harvests the collection several times a year through the North Carolina Digital Heritage Center, the North Carolina service hub, and makes it available through its interface. Other potential access mechanisms include the library’s Ex Libris Primo discovery system, which could host MARC records derived from the Islandora MODS records or Dublin Core records harvested from the repository using OAI.

Partners in Forming a Digital Solution

Though SCUA staff have performed all imaging and metadata work for the project thus far, developers in the library’s Technology and Digital Strategies unit have been crucially important to the project’s success. They built the Islandora repository from the ground up and assisted project staff with their first experimental ingests of motorsports images and metadata. Chief among their contributions are the NAS ingest tool for quickly loading collections into the repository and the annotation tool for the donor.

Resources Required

Aware from the beginning that staffing and budget constraints would sharply limit the number of project participants, project staff attempted to design workflows that would enable a small number of staff to produce a large volume of images and metadata. They were also conscious that in the absence of numerous full-time staff in the department, student workers would need to perform highly specialized work ordinarily assigned to FTE. Carefully vetting student hires and extensively documenting workflows has allowed the department to make the most of a depleted workforce.
Staff involved in building the motorsports digital collection include one digital production librarian, a staff member responsible for managing day-to-day imaging workflows, and a metadata librarian. Seven to eight undergraduate student assistants perform tasks ranging from digitizing analog photographs and slides to performing quality control on images to creating and transforming metadata. During the summer of 2017, the department hired a full-time intern to extract born-digital images from the photographer’s hard drive, process the images, and devise born-digital metadata workflows based on already-established processes for the description of digitized images. Previously, project staff had concentrated exclusively on providing access to digitized photographic materials.

**Analysis**

**General Reflections**

Some of the difficulties that arose during the project can be traced to the heightened expectations that the library may have unintentionally encouraged in the donor early in the project. In the heady early days when a project is first taking shape, it is easy for staff and donors to allow rosy visions of phenomenal progress to obscure the gritty realities of budget and staffing shortfalls that are sure to complicate any undertaking as large as the McMurray motorsports project. In retrospect, it would probably have been prudent for project staff to confine their commitments to the donor in the beginning stages of the project to what they knew they could actually deliver. Doing so would almost certainly have eased frictions that developed during the project and made for a more harmonious donor-library relationship.

**Digitization**

Learning to work with the Hasselblad Flextight X5, training students on digital production processes, and ensuring quality of digital images remained challenges throughout the project. The donor had high expectations for speed, volume, and the creation of images with a finished look which Atkins, relying on a largely student workforce, couldn’t always meet.

When UNC Charlotte purchased the Flextight scanner and batch tray feeder, the digital production librarian had hoped that staff would be able to load multiple trays of 35mm negative strips and walk away while the machine scanned. Project staff learned that monitoring and occasional human intervention was needed during the scanning process to prevent various issues that would affect the quality of images. Also, more rescanning of images than we had anticipated was needed due to film shifting in the holders; holders failing to grip the film tightly during the entire scan, which resulted in blurry images; or too much dust, despite efforts to brush off debris prior to scanning.
Negative digitization projects are further complicated by the necessity of additional steps post-scan to produce an image that reasonably represents a print from the original object. Digitization staff are very conservative with image adjustments and avoid clipping detail in the light and dark regions. This would allow someone to use a copy of the preservation master digital file to emphasize various parts of the image, but depending on the photograph, may result in a flat, uninteresting digital image. Additionally, color negatives in the collection, particularly the medium-format negatives, are degrading unevenly across the surface of the film. While we want to digitize and preserve a copy of these materials as they are now before they break down further, it’s not possible to create digital images from the color negatives that would look the same as a print created from the film around the time of capture in the 1970s.
Metadata

![Table of metadata](image)

**Figure 6: Sample metadata record for motorsports image in Goldmine repository.**

Staff made a number of errors during the project. Many had less to do with project design than with a lack of sensitivity to the donor’s needs. A defensiveness about archival and metadata standards and an inflexible adherence to established workflows may have led project staff to view his difficulties with Excel as a peripheral problem, for instance. They were almost certainly too slow in devising a more simple and convenient tool for recording his metadata.

Reducing the difficulty of metadata creation, the new metadata annotation tool has accelerated the pace of the donor’s work and eased some of his concerns about the project’s slow rate of progress. It has not, however, cleared away the problems with communication that would be likely to arise in any project involving multiple participants creating complex metadata. In some cases, staff must contact the donor by email to get clarification about ambiguous metadata. Staff also contact the donor each time a new collection is ready to annotate. Sometimes the donor believes that he has been sent work he has already performed, an experience that increases his frustration with how rapidly the project is progressing. Though more than 8,000 images have been ingested since the
project began, the donor believes that SCUA should be digitizing and creating metadata for his photographs at a far faster rate. The rapidity with which he is able to complete his part of the metadata with the new metadata tool has increased his sense that SCUA’s own pace is laggardly.

Finally, the extremely granular, item-level metadata that the project designers committed themselves to record for each photograph has resulted inevitably in a slow rate of production for the project. It is not difficult to imagine approaches to creating metadata that would have been swifter and still made use of donor metadata: collection-level descriptions for each sub-collection that included general information about persons appearing in the images, for instance, but no metadata at the item level. Adopting these or similar measures might have increased output and assuaged some of the donor’s concerns about the project’s speed.

**Project Results**

Despite occasional misunderstandings, SCUA and the donor have together created an immense, still-growing digital photographs collection with rich, intensely detailed metadata about motorsports culture and personalities that project staff could never have provided on their own. The project has succeeded in capturing much of the donor’s encyclopedic knowledge of motorsports and passing it on to users in the form of free-text notes and controlled vocabulary that allow them to search his vast photographic archives with great precision and speed. Though mediated by staff members, the metadata he has contributed allows him to speak to users about his life’s work with a clarity and detail that is denied to most donors of archival collections, who must rely on archival intercessors to interpret their collections for the public.

Project staff have derived a number of benefits from the project, not least the confidence that comes from launching a large project with extremely limited means. They have applied many of the same techniques for creating and ingesting bulk metadata in this project to other work in the unit, and they are thinking about other more ambitious uses for the tools and workflows for collecting donor metadata—crowd-sourced metadata, for instance. Perhaps most importantly, they have gained a clear understanding that cooperative metadata creation is a deeply messy process attended by all of the errors and friction associated with any human effort.

**Next Steps**

Initially, SCUA focused its energies on digitizing negatives from the early 1970s phase of McMurray’s career, most of which are black-and-white images. More recently, it has turned its attention to processing and creating metadata for the large number of born-digital color images in his collection, as well as beginning to digitize color film from the 1980s and 1990s. The intent is to create a more balanced digital collection that represents all phases of McMurray’s photographic output.
At the same time, project staff are devoting more effort to promotional and graphic design aspects of the project that have been neglected until now. The current web interface through which users access the collection employs an out-of-the-box design scheme of drab greens and unprepossessing fonts that ill befits the kinetic events depicted in the images. In the coming months, a newly hired graphic designer will create a more visually attractive and engaging interface for the project.

**Future Plans**

It is unlikely that project staff will succeed in digitizing the entire corpus of McMurray’s photographs, at least not with current metadata workflows and staffing levels. The project could conceivably continue with a less rigorous metadata model that emphasizes collection-level rather than item-level records. But a more likely fate is a gradual winding down of the work after staff determine that they have compiled a collection that showcases most phases of the photographer’s work and offers a generous survey of motorsports culture from the 1960s to the present day.

Future SCUA initiatives, however, may draw on the large storehouse of images and metadata accumulated during the course of the Bryant McMurray motorsports project for entirely new purposes. Online exhibits, timelines, data visualizations, and digital humanities projects that explore the rich history and culture of motorsports are only a few of the possible ventures that might make use of the project’s raw materials. And the techniques and tools that unit staff developed for this project will almost certainly find new life in future projects involving donors and crowd-sourced metadata. Far cannier about the benefits and drawbacks of donor-supplied metadata than they were in the past, and more knowledgeable about how it can be collected and used, unit staff will be able to collaborate productively with future donors on projects that draw on their knowledge and expertise.