

Sustainable Data Management for the Solar Decathlon

Bridging the Spectrum Symposium
Catholic University of America
February 8, 2019



Patricia Kosco Cossard
David Durden
Michael Molyneaux-Francis



Data-Driven Information Services Session

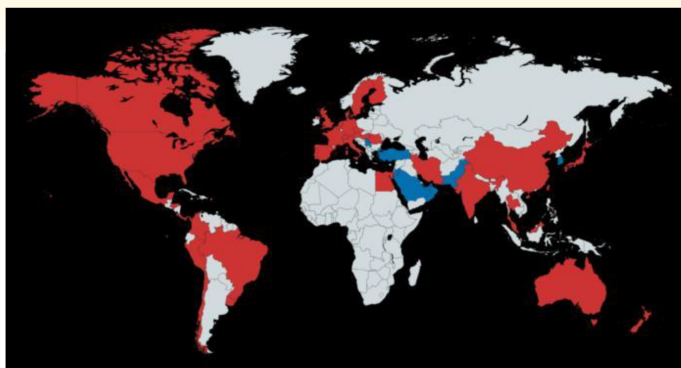
Great Room B, Pryzbyla Center
11AM-12:15PM

Attendees: 65

Patricia Cossard is the Art Librarian at the University of Maryland Libraries. Patti earned her Master of Library Science from Rutgers University in 1987 and holds a Master of Arts from the University of Toronto (1984). Patti's applied librarian practice spans 35 years at four academic/research institutions: Princeton University (NJ), St. Joseph's University (PA), The Historical Society of Pennsylvania (PA) and the University of Maryland (MD). An early scholar of Digital Humanities, her practice includes building bibliographical databases, cataloging, metadata migration and structures, curation and special collections, and subject specialist. Cossard believes that the future of subject specialization must include data curation. In the area of Information and data management she has been embedded into the UMD faculty leadership of three US Department of Energy Solar Decathlon international collegiate competitions: 2007 (top US team), 2011 (top team Worldwide), 2017 (top US Team. She is currently, embedded in the Solar Decathlon Europe team competing this Summer in Budapest, Hungary.

Context

In 2002 the U.S. Department of Energy inaugurated the Solar Decathlon. The program has expanded to events held on other continents: Africa, China, Europe, Latin America and Caribbean, and the Middle East.



Bridging the Spectrum Symposium - Catholic University of America - February 8, 2019

The Solar Decathlon is a series of international competitions that challenge collegiate teams to design, build and operate solar-powered houses that are cost-effective and attractive. It is based upon an initiative of the U.S. Department of Energy (DOE), the Solar Decathlon was inaugurated in 2002.

It consists of ten contests, juried or metered, with prizes for the overall three highest cumulative scores. Teams comply with competition rules, building codes, and graphic standards. Teams enter a Memorandum of Understanding with the Organizers that requires a scheduled set of deliverables. Teams may be either single institutions or a multi-institutional effort.

The Solar Decathlon has expanded internationally to include five additional worldwide competitions: Solar Decathlon Africa, Solar Decathlon China, Solar Decathlon Europe, Solar Decathlon Latin America and Caribbean, and Solar Decathlon Middle East. It is the largest sustainable building research and innovation competition in the world. Technology demonstrated at the decathlon's have a 5-10 year lead to broad market availability and adoption.

Tens of thousands of students, educators, and industry mentors at over 500 educational institutions from nearly 60 countries on five continents have raised public awareness of the importance of renewable energy, sustainability, and energy efficiency. It has emerged as a core 20th century pedagogy for educating future engineers, architects, and ecological scientists, operating in multidisciplinary teams, to advance sustainability technology. The very scale and complexity of competing in the Solar Decathlon are essential elements in mastering sustainability competence and leadership skills development.

Team Maryland competed as a single institution in 2002, 2005, 2007, 2011, 2017. Maryland took first place overall in 2011, and first place US institutions (second place overall) in 2007 and 2017. In 2002 UMD placed 4th overall. In 2005 Maryland received the People's Choice Award. Recognized for its record of excellence, Team Maryland has been invited to collaborate on multi-institutional teams for Solar Decathlon Europe and Solar Decathlon Africa, both held in 2019.

While more than 500 books, theses, reports and articles have been written about individual competitions in its nearly two decades fifteen years of existence, to date there has been no systematic archiving of the research, scholarly, and creative work created by these competitions.

Process



- At the Beginning: Be a team player
- During the Research Project: Be a record keeper
- At the end of the project/beginning of archiving: Be an investigator

Bridging the Spectrum Symposium - Catholic University of America - February 8, 2019

Patricia Cossard and David Durden (DSS) are developing a digital asset management standard and best practices that will be disseminated internationally to all teams and agencies, past, present and future. The hope is that by recommending deposit in an Institutional Repository (IR), e.g., D-Space, the records of this early 20th century engineering, architecture and environmental science pedagogy will be discoverable, sustained, and measured.

Guiding principles for their infrastructure development include: clear intellectual property attribution via DOI to home educational institution; sustained authentic, verified, and accurate “master copy” (limited editorial and management); perpetual institutional commitment to steward and curate the collection; generous use for educational purposes, freely available and licensed via Creative Commons; measure impact of student and faculty output and performance of University assets; and demonstrate the continued relevance of the Academic Libraries. Our institutional repository (IR) was the vehicle to drive these principles.

As a Subject Librarian, if offer this model for collecting primary digital data and scholarly communications. Be a team player. Embed in projects early, e.g., Co-Principal Investigator (PI) with responsibilities for information, data, and communication. Engage Data Management Librarian as early as possible to develop a strategic overview of libraries added value to project. Develop your own guiding principles for sustainable data products with PIs and other stakeholders for maximum buyin. Introduce *Metaliteracy* theory and practices as your instructional contribution.

Once you’ve established your role, and while active research is progressing, I recommend that you record taxonomy as it develops. Keep Author lists and track creative/intellectual attributions. In order to maintain a “master version,” develop protocols to establish documents of record and then limit editorial changes.

If you have managed to follow these guidelines, your tasks will be much easier once you begin archiving the research team’s output. When you are ready to create your collection, re-engage with your library’s Data Librarian or IR Manager. Together review data plan, current state, size, and virtual location of records. As questions will surely come up after-the-fact, it behooves you to keep contact with authors and share the collection’s milestones. Populate your “domain” team appropriately, then meet regularly for dialog and meta-reflections.

The Vastness of it All...

The nature of the Solar Decathlon as a pedagogy is based upon a participatory online environment where production and sharing of original and repurposed information is the basis of knowledge making.

Without including literacy in the theory and methods of metacognition and metamemory as learning outcomes, data management grew unwieldy.



Bridging the Spectrum Symposium - Catholic University of America - February 8, 2019

Metaliteracy provides a unified and comprehensive approach to learning, which encourages the production and sharing of original and repurposed information in participatory environments. It includes skills of metacognition, i.e., knowledge about when and how to use particular strategies for learning or problem-solving. Metamemory, defined as knowing about memory and mnemonic strategies, is an especially important form of metacognition.

By teaching and applying metaliteracy to the research practice of teams producing and managing digital assets, may help team members to engage in the added work of building a real-time archive. This type of digital asset management is opposed to the traditional archiving model in that documents are living, and in their pre-final versioning have a limited life-cycle, but the group decides when the asset is final (and therefore an asset of added value) and follows their pre-determined protocols for deposited in the institutional repository for long-term stewardship. This is different from the traditional archive model, where research teams collect their working files and hand them off to a curator/archivist the decisions and definitions of which are valuable legacy assets and therefore deserving of long-term stewardship.

The metaliterary processes are a part of the student-centered learning process as well as enhance the depth of learning. On creative and research projects such as the Solar Decathlon, reflection and categorization are proven ways to form a shared and articulated body of knowledge that is being produced by a large team that for the most part work in sub-teams.

Lessons Learned

Through the Lens of Digital Asset Management this project demonstrated that the Research Cycle is much more expansive than traditional models. The agency of library by organizing these digital assets is essential to the continued research funding potential.



Bridging the Spectrum Symposium - Catholic University of America - February 8, 2019

Dave

Bridging the Spectrum Symposium - Catholic University of America - February 8, 2019

Context

Data Management
or Records Management
or Digital Curation

Why not all three?

Bridging the Spectrum Symposium - Catholic University of America - February 8, 2019

As the data services librarian at the University Libraries, I often consult and assist with data curation and management projects. When I was approached almost a year ago by Patti to consult on developing data and records management workflows for the Maryland Solar Decathlon team, I originally thought I was going to be working on a single data management project. While there are plenty of data assets to be managed within Team Maryland, I soon realized that what I was really dealing with was a massive information creation process that spanned multiple years and involved a couple hundred participants.

While Patti and I are working through the specifics of data and records management policies for future Solar Decathlon projects at Maryland, we seized the opportunity to capture the output of the most recent competition by establishing a community managed collection in the Digital Repository at the University of Maryland (DRUM), which is a DSpace instance. DRUM hosts 22 other community managed collections, so using DRUM to publish and make accessible Team Maryland's output made perfect sense.

Process

- Identify and select materials for inclusion
- Determine appropriate access formats/original file formats
- Apply naming conventions for each deliverable type
- Arrange materials by competition
- Create item level metadata
- Prepare materials for batch-loading

Bridging the Spectrum Symposium - Catholic University of America - February 8, 2019

Patti and I met several times to discuss what types of materials would be best suited for inclusion in the collection. Because DRUM is a repository, we decided that only final documents, or competition deliverables, would be selected. Patti and Michael selected the individual files, and Patti and I established appropriate naming conventions for each deliverable type and identified the appropriate access formats as well as original file formats for architectural and design files. We also thought it best that each competition be given its own collection within the community so that we would have a sustainable and reproducible structure for adding future competitions as well as retroactively adding previous years.

The process of creating a community managed collection in DRUM was straight forward -- I have administrator privileges for the application so I created a test collection on our staging server. I assigned both Patti and Michael's accounts editor privileges so that they could also familiarize themselves within DRUM. I was also able to quickly build a metadata creation spreadsheet that would match the existing fields in DRUM, and Patti and Michael created the controlled vocabulary for item types and keywords, and author/contributor names.

Records in DRUM could be created manually, but I decided that once the metadata was created and the files were selected, I would use our batch loading process to add records once the collection goes live.

A Peek at What's to Come

- A Solar Decathlon: Team Maryland community has been created
- Competitions are organized by year/team name
- Other research output from UMD previously submitted by faculty can be mapped to display in the Solar Decathlon community

BROWSE BY

By Issue Date Authors Titles Subjects

Search within this community and its collections:

Go

The Solar Decathlon (SD) is a series of international Design/Build competitions for collegiate teams to design, build and operate solar-powered houses. Team Maryland has competed as a sole institution in the SD (US in 2002, 2005, 2007, 2011, 2017) and has expanded in 2019 to collaborate with International teams for Solar Decathlon Europe and Solar Decathlon Africa.

Collections in this community

Solar Decathlon 2017: Team reACT

Recent Submissions

90% Complete Design Drawings

Team reACT; Johnson-Williams, Malik; Oh Boun, Sandra; Elmahadi, Alla; Rockcastle, Garth C; Binder, Michael P (2017-11-17)

90% complete construction documentation: General, including solar envelope compliance, accessible tour route, finished square footage, water delivery and removal compliance information, and constructed footprint calculations. ...

Team Overview To Date

Team reACT; Campbell, Jamie; Cossard, Patricia Kosco; Rockcastle, Garth C (2016-09-15)

Updated 100-word description of team and its goals; Digital representation (renderings, graphic floor plan, photography of scale model, animation, etc.) of competition prototype design; Team photograph; Summary of unique ...

Schematic Design Summary

Team reACT; Rockcastle, Garth C; Binder, Michael P; Adomaitis, Raymond A (2016-04-29)

Mission Statement; Detailed strategy for winning the competition including a contest-by-contest breakdown; Design drawings and/or written description of the following systems and components, with identification of any ...

Project Management Plan

Team reACT; Rockcastle, Garth C; Adomaitis, Raymond A; Binder, Michael P (2016-03-31)

100 Word Description

Team reACT; Binder, Michael P; Rockcastle, Garth C; Adomaitis, Raymond A (2016-03-31)

100 word description of team and goals for placement on Solar Decathlon website. This can change later, but should serve as a current introduction.

[View more](#)

Bridging the Spectrum Symposium - Catholic University of America - February 8, 2019

The Unexpected

- Came in expecting to assist with a data management plan; came out consulting on a digital curation project
- Learned a lot about the Solar Decathlon
- The scope of the project changed several times over the course of a year

Bridging the Spectrum Symposium - Catholic University of America - February 8, 2019

This project evolved and went through several iterations before we settled on our current plan: establish the workflows and the repository collection to disseminate and make available the project deliverables for Team Maryland. Now that we have an idea of what the end of each competition's life cycle looks like, we are better prepared to develop records and data management procedures for the forthcoming competition teams.

Lessons Learned

- Never underestimate the ability of teams to produce a lot of records

Bridging the Spectrum Symposium - Catholic University of America - February 8, 2019

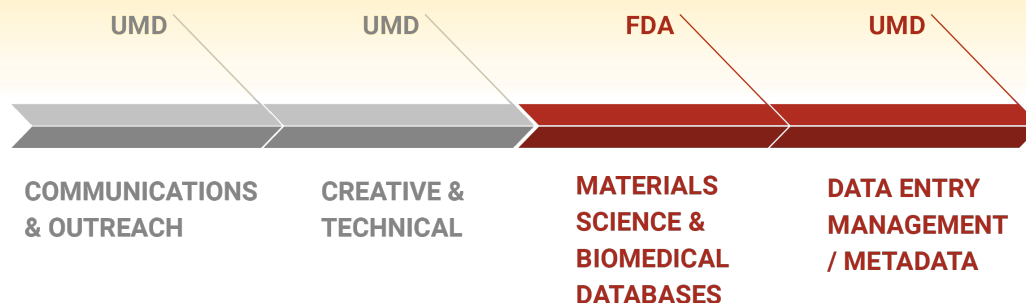
I only have one real 'lesson learned', which is to never underestimate the ability of groups to produce lots of information. I approached this project thinking that there would be some data, some research documents, notes, etc. I did not expect the amount of working, living documents that would be covered in the project. I am also grateful that Patti has been working so closely with the Solar Decathlon throughout the project life cycles -- without her immense knowledge of the project, it would have extremely difficult (and time consuming) for me to determine what was valuable.



Michael

Bridging the Spectrum Symposium - Catholic University of America - February 8, 2019

Context



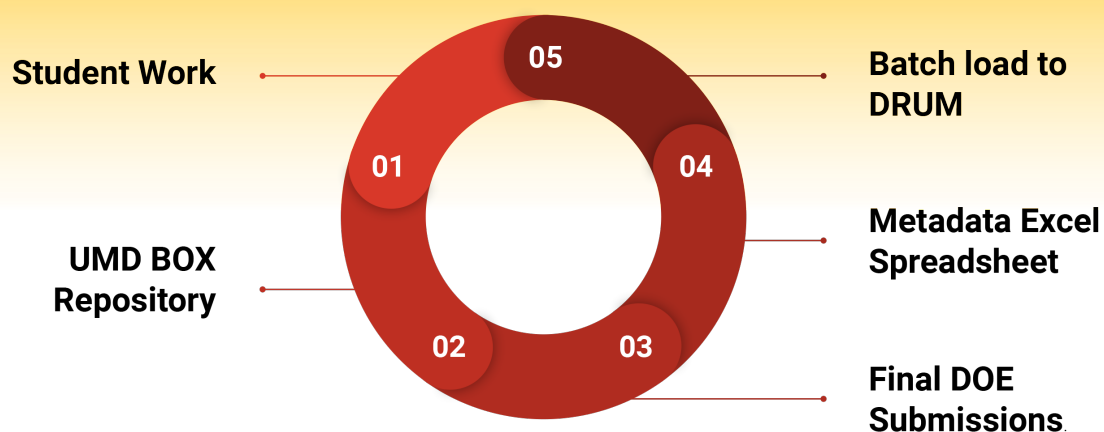
Bridging the Spectrum Symposium - Catholic University of America - February 8, 2019

My role: Data Entry Management

My path towards data entry management has been quite fluid. I first began the Solar Decathlon 2017 project as a student communications & outreach coordinator. In this role, I reached out to different indigenous groups (e.g. Chippewa tribes of Minnesota) as well as collaborated with my peers on sustainable living systems and practices germane to indigenous groups.

While collaborating on the Solar Decathlon project, I concurrently interned at the U.S. Food and Drug Administration (FDA) where I ran several materials and biomedical device databases (e.g. Granta Design and Medical Device Reporting (MDR) in response to spinal biomedical implant devices. Working with scientific and technical databases at the FDA has allowed me to transgress into library science metadata for the solar decathlon project. Currently my role in the project is data entry management (which takes into account the metadata attributed by University of Maryland's part in the Solar Decathlon 2017 global project). My data entry process is briefly discussed on the next slide.

Process



Bridging the Spectrum Symposium - Catholic University of America - February 8, 2019

My particular process in the matter is quite simple.

All of our data is housed in our UMD BOX database, where only those members a part of the sd2017 project can have access to. The files in BOX are broken up into their deliverable and assignment sections. From the BOX database, I collect the most pertinent files (i.e. the final documents that were submitted to the Department of Energy (DOE)) and transfer them into our Google Drive team page for conciseness and easy access. From there, I tabulate entries into our metadata excel spreadsheet pertaining to the final documents. Once all of the constituents of each entry are fully accounted for, the excel data is then moved to DRUM (either manually or via batch upload).

The Unexpected

- Developing global infrastructure
- Ongoing expansion of research
- Creating scholarly presence for students

Bridging the Spectrum Symposium - Catholic University of America - February 8, 2019

When I first heard about this project, I assumed it was just a more efficient way to organize and catalogue the achievements of our UMD's 2017 solar decathlon team. As I began working on the project, I realized that the extent of the project goes far beyond the scope of UMD. Creating a data repository on the UMD sd2017 achievements allows us to have a global reach. Furthermore, the ongoing expansion of research and the creation of a scholarly presence for students grew more apparent.

Lessons Learned

- Be consistent
 - Format in DRUM (e.g. file identifiers, sentence structure, labels, etc.)
 - Preserve authorship(s), sponsors, etc.
- No need for redundancy
 - Search engine optimization is more powerful than you think
- Take it one step at a time

Bridging the Spectrum Symposium - Catholic University of America - February 8, 2019

One important takeaway from this, is that it is extremely important to be consistent when creating information in DRUM. File identifiers, sentence structure, labels attributed to constituent entries must all follow the same format. Furthermore, preserving authorships, sponsors, and crediting those for each file upload is mandatory.

Another important takeaway from creating an online repository (e.g. DRUM) is that there is no need for redundancy. Search engine optimization is more powerful than you think. DRUM's indexing software creates tokenized words. Therefore, when you create a phrase (or even a word) it becomes a part of the infrastructure. For example, it will be unnecessary to create a file identifier with the name "solar," when "solar" is already in the title of the document.

Finally, as you are working through such a large database, it is important to take it one step at a time. There are cohorts of files that are in varying orders and are not always clear to comprehend or reference. Taking the time to inspect each file will create less errors when uploading final metadata into DRUM.

Thanks!

Patricia Kosco Cossard
Art & Sociology Librarian
pcossard@umd.edu

David Durden
Data Services Librarian
durden@umd.edu

Michael Molyneaux-Francis
Student Library Assistant
mmolynea@terpmail.umd.edu

Bridging the Spectrum Symposium - Catholic University of America - February 8, 2019