11-2015

Collaborate to Innovate: Expanding Access to Faculty Patents through the Institutional Repository and the Library Catalog

Andrew Wesolek

Clemson University, andrew.j.wesolek@vanderbilt.edu

Jan Comfort

Clemson University, comforj@clermson.edu

Lisa Bodenheimer

Clemson University, bodenhl@clermson.edu

Follow this and additional works at: https://tigerprints.clemson.edu/lib_pubs

Part of the [Library and Information Science Commons](#)

Recommended Citation

Wesolek, Andrew; Comfort, Jan; and Bodenheimer, Lisa, "Collaborate to Innovate: Expanding Access to Faculty Patents through the Institutional Repository and the Library Catalog" (2015). *Publications*. 101.
https://tigerprints.clemson.edu/lib_pubs/101

This Article is brought to you for free and open access by the University Libraries at TigerPrints. It has been accepted for inclusion in Publications by an authorized administrator of TigerPrints. For more information, please contact kokeefe@clermson.edu.

TECHCAST

Collaborate to Innovate: Expanding Access to Faculty Patents through the Institutional Repository and the Library Catalog

ANDREW WESOLEK, JAN COMFORT, and LISA BODENHEIMER

Clemson University, Clemson, South Carolina

Institutional repositories and library catalogs have long been used to promote and provide access to scholarly articles, monographs, and other “traditional” forms of scholarship. However, faculty at many institutions participate in “non-traditional” forms of research, such as patents. This work explores the use of Clemson University’s institutional repository and library catalog in promoting and providing access to the patents granted to its inventors. More specifically, the authors outline the project from generating metadata for patents, and cross-walking it into a form that can be batch uploaded to a bepress Digital Commons Repository. They then explore the process of harvesting the records from the repository and using MarcEdit to crosswalk it into the library catalog. The process could easily be adapted to other non-traditional forms of scholarship.

Keywords: Patent Research, Institutional Repository Content, Metadata Crosswalking

Clemson University is a medium sized, public land grant institution located in the foothills of South Carolina, midway between Charlotte, NC and Atlanta, GA. Student enrollment includes more than 17,000 undergraduates and nearly 5,000 graduate students, taught and supported by about 5,000 faculty and staff. There are more than 80 undergraduate and 110 graduate degree programs with an emphasis on the STEM disciplines of Science, Technology, Engineering, and Mathematics. In 2014, Clemson University was ranked 20th in the nation among public universities by U.S. News and World Report.

The Libraries house over 1.5 million print volumes (including federal and South Carolina documents), and host 163,000 eBooks, 45,000 ejournal subscriptions, and nearly 500 databases. In addition, the R.M. Cooper Library is a Patent and Trademark Resource Center (PTRC), participating in a program administered by the US Patent and Trademark Office (USPTO). As a PTRC, The Library receives a variety of patent and trademark searching tools and librarians receive extensive training in using these resources. In the spirit of the University's land grant mission, the Library is open to the public, and services are free of charge. The Libraries are committed to "connecting the university community with resources vital to learning, teaching, research, and innovation..." Assisting Clemson students, faculty, and staff with their patent information needs is one way librarians support this mission (Clemson University 2015).

To this end, in 2014, the library embarked on a project to collect all patents produced by Clemson researchers and not only showcase them in the newly-launched institutional repository, but harvest them into the library catalog. The purpose of this project was to leverage the search engine optimization and discoverability of the institutional repository, with

the robust search functionality of the online catalog. This work explores the process of batch uploading patents to a bepress Digital Commons repository and using MarcEdit to harvest patents from the IR to the catalog, as well as our successes and next steps.

WHAT ARE PATENTS AND HOW DO THEY FIT INTO THE RESEARCH PROCESS?

A patent is a monopoly conferred by a government authority for a limited time to prevent others from making, using, selling, or offering to sell an invention. In general terms, the USPTO grants an inventor 20 years of patent protection, and in exchange the inventor must fully disclose the workings of their invention. In order to be patentable, an invention must be new, useful, and non-obvious. In addition, an invention must be reduced to practice, meaning it is not possible to get a patent for an idea (U.S. Patent and Trademark Office 2011).

The academic research process is well-documented and well-understood. Faculty conduct research, draft articles about the research, then submit them to publishers of academic journals. Once received by the publisher, articles undergo a peer review process and the approved articles will eventually be published (hopefully within several months). Faculty are evaluated based on the number of their publications, and the perceived value of the journals that publish their work.

But patents are not journal articles. The purpose of a journal article is to report on original research. It may involve a patented process or machine, but it is the results of the research that are most important. The purpose of a patent is to document a process or discovery. The path to obtaining a patent differs from publication of an article, and it is much less familiar to researchers. Patents undergo an examination process as opposed to peer

review. There can be significant changes during this period, which takes an average of 36 months.

Increasingly, faculty are being encouraged to pursue patent protection for commercial aspects of their research in addition to writing articles. As faculty patents become more common, the research they represent will be more highly valued, but generally speaking, patents are not given the same weight as journal publications in promotion and tenure decisions. Legal requirements for patent protection can also compete with the publication process, creating further challenges for researchers. These are some of the reasons that approximately 80% of the information found in a patent is not available anywhere else in comparable detail (U.S. Patent and Trademark Office 1977).

WHERE CAN YOU FIND AND ACCESS PATENT INFORMATION?

Patents are generally freely available on the internet. The USPTO provides access to all US patents, going back to 1790, in their patft database www.uspto.gov/patft. Patents from the European Patent Office (EPO) and other patent-granting authorities are available in the Espacenet database worldwide.espacenet.com/. The World Intellectual Property Organization (WIPO) also provides access to patents from many countries via their PATENTSCOPE database www.wipo.int/patentscope/en/. Even Google has a stake in the game. Their patents search site www.google.com/patents is an excellent resource and provides improved keyword access to patents from around the globe, with embedded links to Espacenet and several USPTO resources.

There are also a number of commercial patent databases; Lexis/Nexis and Derwent World Patents Index are the two best known. Other commercial scientific and technical databases include coverage of patent bibliographic data. However, their coverage is selective rather than comprehensive, based on the subject focus of the particular database. Because these databases are expensive – and license restrictions can limit their access – they are not available to all researchers (Reinman 2013).

As a PTRC Clemson librarians are particularly supportive of patent research support and remain alert to opportunities to provide access to patent information to our constituents. The PTRC research guide highlights services, offers strategies for effective patent searching, and provides links and instructions on how to use the various databases. In addition, there are links to commercial databases with selective subject coverage for patents which are available to Clemson affiliates; for examples, see <http://libguides.clemson.edu/ptrc>. Another way that the library provides access to patent information is through the discovery layer of its catalog. The discovery layer is an instance of Summon from ProQuest, which includes patents. Users can facet their results to content type “patent” to access patent information from the EPO as well as the USPTO.

EARLY DEVELOPMENT OF CLEMSON UNIVERSITY’S INSTITUTIONAL REPOSITORY

In addition to its role as a PTRC, Clemson Libraries hosts a thriving institutional repository. TigerPrints, our instance of the bepress Digital Commons institutional repository platform was launched in October 2013 during Open Access week, with the goal of capturing the intellectual output of Clemson University and disseminating it broadly, rapidly, and openly while making it

easily discoverable around the world. The initiative found early success, and as of July 2015, the platform hosts more than 8,000 works of scholarship that have been downloaded over 500,000 times.

Critical to sustaining and expanding on this early success is the Libraries' ability to cultivate faculty buy-in and recruit content. From the Libraries' perspective, of course, there are numerous reasons to make scholarship open access through an institutional repository (Crow 2004). However, conveying these to busy research faculty on any campus requires a deep understanding of the particular institutional environment. With such an understanding one can craft winning arguments that speak to the motivations of one's faculty.

Clemson University is currently refocusing its strategic plan to include a greater focus on research output and impact. Literature continues to suggest that making one's work openly available online, whether through open access publication or self-archiving in an institutional or disciplinary repository increases both the speed and frequency at which it is cited (Hitchcock 2004) Therefore, the library decided to focus its outreach efforts on the repository's potential to increase the impact of one's research.

WHY ADD PATENTS TO A REPOSITORY?

Content that is uploaded to Digital Commons repositories is indexed in Google and other major search engines. However, the all-important ranking of that content can be impacted by quality of metadata, linking of the repository and content to known trusted websites, and overall traffic to the repository in general. While there are steps repository managers can take to improve metadata in their repositories, and link content to known and trusted websites, these

are outside the scope of this work. Here, the focus is on early additions of content to the repository to generate traffic to the repository generally, thus helping to boost the visibility of scholarship recruited from our faculty.

Many repositories begin by hosting electronic theses and dissertations (ETDs) completed by the students at their parent institutions. ETDs are an attractive target for managers cultivating a recently launched repository for three reasons:

1. Typically, students grant to their university a nonexclusive license to host and distribute their thesis or dissertation, so these are free from most copyright concerns.
2. ETDs are, by definition, already digital, allowing them to be uploaded without digitization.
3. ETDs are heavily used by current students, future students, and others, so they play a significant role in driving traffic to repositories.

Like many other institutions, Clemson University focused on populating our newly-launched IR with ETDs. In our case this involved migrating nearly 3,000 ETDs dating back to 2006 from an aging homegrown system to our new repository. As this project neared completion, a casual conversation between Comfort and Wesolek, two authors of this paper, resulted in an epiphany. As Wesolek described the process of harvesting ETDs and the reasons why they were so desirable for TigerPrints, Comfort wondered aloud, “what about patents, could we add those to the repository?” And then, “ETDs have bibliographic data and links to TigerPrints in the Library catalog. Could we add bibliographic records to the catalog for Clemson patents, too?”

Indeed, patents have many of the same benefits as ETDs to an early repository. Copyright is a major impediment to posting content to institutional repositories, but since patents are produced by the federal government, and are thus in the public domain, they are free from copyright restrictions. Similarly, patents are also available electronically, which eliminates the need for digitization. The faculty at many institutions likely will have filed a number of patents over time, making them a good source of bulk content for the repository, and as discussed above, they are of scholarly significance and used by many researchers.

Additionally, the repository can be used to highlight the inventions created by faculty at Clemson University. Scholarly articles published by faculty at an institution are diffused through many different journals and publishers. For this reason, IRs are often touted as showcasing the scholarly output of an institution by providing a “one stop shop” for the scholarly articles published there. While the USPTO produces all patents, they are similarly diffuse, because they are not organized by institutional affiliation. Organizing them in this way through the repository benefits the institution.

Moreover, about two thirds of the traffic to TigerPrints comes from Google, Google Scholar, or other search engines. This enhanced visibility coupled with the analytics provided to authors and inventors who contribute to the repository also benefits them directly. Finally, Clemson faculty inventors are notified when their patents are published in TigerPrints, creating another opportunity to promote the repository.

HOW CLEMSON DID IT

Metadata for all patents associated with researchers at Clemson were compiled primarily through Free Patents Online (<http://freepatentsonline.com>) (FPO) and the US Patent and Trademark Office (USPTO). A spreadsheet was created consisting of standard metadata fields used predominantly among patent information specialists. Fields include patent number, publication date, abstract, application number, assignees, filing date, primary class, and other class (see Figure 1).

Example of spreadsheet containing metadata of patents from Clemson Faculty

Document Number	Publication Date	Title	Abstract	Inventor Name	Application Number	Primary Class	Other Class
2564150	1951-08-14	Tensiometer		Brown, Hugh M.	US8127349A	73/862.471	
2590374	1952-03-25	Yarn twister for spinning frames		Brown, Hugh M.	US12014849A	57/336	
2629254	1953-02-24	Tensile strength tester		Brown, Hugh M.	US17066650A	73/806	73/834
2638727	1953-05-19	Sweet potato digger		Park, Joseph K.; P	US24520351A	171/5	56/192, 171/69,
2650883	1953-09-01	Color-stable red pepper composition and process of producing		Blaricom, Lester O	US22097451A	426/270	
2659433	1953-11-17	Porosity tester		Brown, Hugh M.	US14212150A	368/2	73/38, 368/101
2713209	1955-07-19	Lap thickness meters		Brown, Hugh M.	US37607653A	33/792	33/501.02
2718111	1955-09-20	Yarn twister for spinning frames		Brown, Hugh M.	US23333651A	57/337	
2781655	1957-02-19	Machines for continuous dyeing of webs under pressure		Brown, Hugh M.	US46083254A	68/22B	68/43, 118/405
2795102	1957-06-11	Twister spindles		Brown, Hugh M.	US52099855A	57/58.72	
2796040	1957-06-18	Ends down indicator		Brown, Hugh M.	US51624655A	116/303	
2797468	1957-07-02	Warp feed equalizer		Brown, Hugh M.	US44031254A	28/194	66/146, 139/97,
2804771	1957-09-03	Flow meters		Brown, Hugh M.	US39398253A	73/861.72	
2812779	1957-11-12	Tension control device		Brown, Hugh M.	US33631153A	139/100	
2906109	1959-09-29	Knitting machine		Efland, Thomas D.	US70395457A	66/1R	
2908133	1959-10-13	False twister		Brown, Hugh M.	US60297056A	57/336	

Digital Commons allows for great flexibility in describing repository content, as well as batch ingest features. In collaboration with bepress consulting services, Clemson librarians created a new series in TigerPrints to house the patents represented on this spreadsheet. A metadata schema was constructed based on the standard digital commons template, with the addition of the fields indicated above. The standard “keyword” field was modified to display “patent number” and the Digital Commons-specific document type “patent” was also added (see Figure 2).

Example of metadata prepared for batch upload to DigitalCommons

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
1	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract
1	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract
2	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract
3	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract
4	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract
5	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract
6	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract
 
24	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract
25	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract
26	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract
27	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract
28	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract
29	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract
30	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract	Abstract

With this new ingestible spreadsheet created, the metadata was crosswalked from the original spreadsheet to the ingestible spreadsheet created for TigerPrints. This allowed us to batch upload the metadata to the repository, but unfortunately, not the files for the patents themselves.

By adding a publicly accessible link that prompts a download to a batch upload spreadsheet, one can use the Digital Commons platform to batch ingest files in addition to metadata. However, the terms of use posted on FPO's site precluded this option. While patents and the metadata that describes them are in the public domain, users are still bound to the terms of use of the FPO site. Unfortunately, FPO does not permit users to access the site in an automated manner, so batch ingestion of patents from FPO to TigerPrints is not possible.

Instead, librarians downloaded all 420 patents from the USPTO and stored them in a Dropbox folder and subsequently generated a batch revise sheet from TigerPrints, containing the metadata for all patents. Then the public URL from Dropbox was added for each patent to the fulltext_url field on the batch revise spreadsheet, and resubmitted it to TigerPrints. This allowed for the batch ingestion of the PDF for each patent from Dropbox to TigerPrints.

RESULTS AND NEXT STEPS

The batch upload of Clemson Patents was completed in March of 2014. At the time of this writing, these patents have been downloaded 1,966 times, with 301 of them being downloaded at least once (see Figure 3).

Downloads and page-views of Clemson faculty Patents from TigerPrints

As hoped, the patents were driving traffic to the repository, but this project had an additional benefit. Current researchers receive download reports on the patents that have been uploaded to the repository, thus giving them concrete evidence of TigerPrint's effectiveness in making scholarship highly visible through Google and other search engines. As a result, several faculty

who were exposed to TigerPrints through their patents have submitted their CVs in hopes of adding additional content. Moreover, each new patent that is granted offers the opportunity to add content to the repository and to consult directly with the inventor about other forms of research he or she may want to contribute to the repository.

THE TECHNICAL SERVICES SIDE: REPURPOSING IR METADATA FOR THE LIBRARY CATALOG

Records in TigerPrints are harvested by and discoverable through our discovery layer, Summon by ProQuest. However, harvesting metadata from TigerPrints for inclusion in the catalog brings with it additional benefits. Because patent information is often difficult for library patrons to access, it is at risk of being overlooked. The library catalog enhances access by virtue of being one of any library's most authoritative and widely available resources. In addition, it is more familiar to many researchers than the institutional repository. The library catalog is also used by researchers worldwide, either directly or through WorldCat, and when the content it has catalogued is made openly available, such as through an institutional repository, those researchers may access it. Finally, harvesting and cross-walking institutional repository metadata in qualified Dublin Core to MARC also extends the function of the catalog to include non-traditional library materials.

In addition to the benefits of including patents in the library catalog, this project offered a great opportunity for collaboration between three library units: Information and Research Services, Technical Services, and Library Technology. It was also a testbed in Technical Services for other automated batch creation processes involving IR metadata, like electronic theses and dissertations.

In her article, “Making IR Content Discoverable,” Fang Wang describes a two-part workflow that uses an in-house program to harvest the metadata and translate it into MARC (2012). In contrast, technical services librarians at Clemson established a workflow using MarcEdit to harvest and translate the metadata. Our implementation process consisted of five major steps, explained briefly below. This process produced brief MARC records with author, title, patent number, date of submission, abstract, and a genre heading for Patent.

1. Harvesting Metadata from the IR with MarcEdit.

MarcEdit is a well-known application used in many libraries for manipulating MARC bibliographic record files, particularly of electronic resources and other materials for which MARC records are supplied in batches. It also has a harvesting function that allows the user to query an OAI-compliant repository and download metadata records. It translates the harvested metadata into MARC-compliant records using XSLT stylesheets integrated into the product. See Appendix 1 for additional information.

2. Examining Harvested Metadata; Examining ILS Validity and Indexing Tables; Making Decisions about what MARC Tags to Use for New types of Metadata.

In the initial test, once the harvested metadata could be looked over, the MARC validity tables and indexing rules in our integrated library system ILS (Innovative Interface’s Millennium) were examined. While most of the metadata was common to other library materials like author, title, and URL to the IR, a new field we wanted to provide access to was the patent number. The

decision was made to use the MARC 024 field (Standard Number) for the patent number largely because this field was already indexed in our ILS.

Example: 024 bb Patent number <xxxxxxx>

A load profile, which specifies how data coming into our catalog will be treated, was created in the ILS, allowing the option to change incoming bibliographic data, standardize system data, and create additional bibliographic data for these records. For example, through the load profile, an author heading for Clemson University was added, and the patent number was moved from a 590 field to the 024. And, a record template was used to create a specific location for electronic patents so that these records could easily be pulled together later.

3. Manipulating the Harvested Metadata with MarcEdit Before Loading.

MarcEdit provides a variety of ways to modify data in a file. It can be as easy as a find and replace function, which was used to delete unwanted HTML formatting in the metadata output, or as complicated as creating an 008 field for every bibliographic record in the file and inserting information into the field. See Appendix 2 for additional information.

4. Creating Documentation and Training Materials.

Since this was a new process for Clemson, documenting the processes was important. Clemson library's intranet has a designated place for unit-produced documentation. Three pieces of

documentation covering the harvest, the modification of metadata, and the load into the catalog were produced. This was used in training the cataloger who did the work.

5. Testing of Metadata Going into the Catalog.

A big part of any process development is testing. It allows one to see what works and what does not work, and how metadata displays in the catalog and how it can be searched. Using a system test environment, we tested the harvesting, manipulation, and loading of bibliographic records before loading them into the production database. In addition, we tested searching in the public test catalog. Readers may contact the authors for additional information, or visit [Clemson Libraries StaffWeb](#) (Clemson Libraries 2015).

These five steps yielded brief, minimal-level records with the author, the title, the full publication date (year/month/day), a summary of the content, a link to the patent document in our IR, an added entry for Clemson University, the patent number, and a genre heading for Patent. The author and title are searchable using the author and title indexes. The summary is keyword searchable. The patent number is in the standard number index and is searchable when prefaced by the words “patent number.”

Public view catalog record

Author [Sun, Ya-Ping author](#)
 Title **Fluorescent Carbon Nanoparticles**
 Publication info. 2015-01-13

Links
Connect to this object online.

Summary Disclosed are photoluminescent particles. The particles include a core nano-sized particle of carbon and a passivation agent bound to the surface of the nanoparticle. The passivation agent can be, for instance, a polymeric material. The passivation agent can also be derivatized for particular applications. For example, the photoluminescent carbon nanoparticles can be derivatized to recognize and bind to a target material, for instance a biologically active material, a pollutant, or a surface receptor on a tissue or cell surface, such as in a tagging or staining protocol.

Genre [Patent](#)
 Added author [Clemson University.](#)
 Standard no. Patent Number 8932877

Staff side view of catalog record

b30472325 Last Updated: 02-12-2015 Created: 02-12-2015 Revisions: 1			
Language	eng English	Cat. Date	02-12-2015
Skip	0	Bib Level	m MONOGRAPH
Location	ip Electronic Patents	Material Type	m COMPUTER FILE
Bib Code	z NON-CONTRIBUTING	Country	scu South Carolina
<p>MARC Leader ##### am a #####3u</p> <p>Misc. 008 150212s9999 scu o d eng d</p> <p>Standard No. 024 8 Patent Number 8932877</p> <p>Misc. 042 dc</p> <p>Author 100 1 Sun, Ya-Ping^eauthor</p> <p>Title 245 1 0 Fluorescent Carbon Nanoparticles</p> <p>Publication Info. 260 c2015-01-13</p> <p>Note 520 Disclosed are photoluminescent particles. The particles include a core nano-sized particle of carbon and a passivation agent bound to the surface of the nanoparticle. The passivation agent can be, for instance, a polymeric material. The passivation agent can also be derivatized for particular applications. For example, the photoluminescent carbon nanoparticles can be derivatized to recognize and bind to a target material, for instance a biologically active material, a pollutant, or a surface receptor on a tissue or cell surface, such as in a tagging or staining protocol.</p> <p>Subject 655 7 Patent 2local</p> <p>Added Author 710 2 Clemson University.</p> <p>Misc. 856 4 1 uhttp://tigerprints.clemson.edu/clemson_patents/451 zConnect to this object online.]</p>			

Considerations for Technical Services

The authors did not experience resistance from our technical services team when introducing this project, but there are some requirements which must be taken into account for a project like this to become a reality. First, having a cataloger participate in the project is the most essential way to ensure that expert knowledge of MARC is available. Furthermore, working with MarcEdit was a very important part of this project, so knowledge of its capabilities, or the time and energy to learn them, is of critical importance. MarcEdit supplemented our ILS's functionality, and it was indeed essential for harvesting and manipulating metadata. Even so, it

is simple to use, and there is much documentation and help available to users via the web and a listserv.

In addition, knowledge of XML and XSLT transformations in MarcEdit is helpful so that metadata can be effectively transformed into MARC. The stylesheets that are part of MarcEdit are sufficient for the minimal-level records shown above; however, knowledge of how the stylesheets are built would allow more customization and better exposure of the IR metadata. Knowledge of integrated library system structure and parameters is also essential, as it allows one to index new types of metadata. Our system support manager was an excellent resource, as was documentation from our system vendor. Knowledge of local cataloging policies, particularly permitting the use of minimal-level records in the library catalog is another key. Clemson Libraries has a long history of creating local minimal-level cataloging records for various types of materials. This history allowed us to go forward with this project. A library that is not accustomed to these kinds of minimal-level records must consider the impact of their use on the catalog and the patrons who use it.

Also, this project has an impact on databases outside the library such as the library's discovery tool, Summon, and the consortial union catalog. In the library's workflow, bibliographic records for newly cataloged materials are pulled together and uploaded to Summon. There is a similar workflow for metadata records in TigerPrints. The TigerPrints records for electronic patents are uploaded to Summon because the metadata in it is more extensive than in the catalog. Because all Tigerprints records are harvested into Summon, the catalog records for the patents had to be eliminated from our normal catalog uploads to Summon to prevent duplication.

If the library is a member of a consortium with a union catalog, its cataloging policies will need to be considered. After consulting with our consortium technical staff, it was clear that these minimal-level records should not go into the union catalog. To this end, technical services coded a system-specific fixed field to keep them from loading in the union catalog. Additionally, there are opportunities to build relationships with outside databases, such as OCLC Worldcat. The existence of the OCLC Digital Collection Gateway, for example, allows for the possibility of adding resource records for Clemson Patents to WorldCat in the future.

CONCLUSION

It is important to realize that repurposing metadata for the catalog is a work in progress, as is all cataloging work. It depends on the ILS used, the expertise of the people doing the work, and changes in standards and technology over time. It is very important to realize that the metadata and the resulting catalog records are not the final word. With system technology that allows searching for specific data in a bibliographic record, any changes necessary can be made quickly and effectively. They can change and grow as expertise and technology change and grow. This is just one way that the capabilities of the library's ILS and catalog can be used to provide access to little-known materials and create an efficient internal workflow for providing access to them.

All known Clemson patents have been added to TigerPrints, with additional access points added to the Libraries catalog. A process is in place to systematically add new patents as they are issued. The project is now focused on discovering additional Clemson patents. Patents

dating from 1976 were relatively easy to identify based on keyword searching, and assignment data maintained by the USPTO; however, patents issued prior to 1976 are image files and not fully searchable. They can be searched by patent number, issue date, and class/subclass. To identify all Clemson patents, it was necessary to consult resources housed in the Clemson University Archives. An initial search uncovered fifteen patents that were not listed on the original spreadsheet. Further searching may yield others, which, along with the supporting archival material, could form part of an interesting digital collection. Consultation with colleagues in Special Collections is essential to the success of this aspect of the project.

Feedback from PTRC colleagues has been very positive. The authors sent an email message in April 2014 to see if any other PTRCs were doing anything similar, but our project appeared to be a unique effort. Since that time, the authors have worked with five academic libraries which are in various stages of implementing their own process to harvest patents from their institutions. As of this writing, Rice University has completed the process of adding patents to their IR.

The benefits of adding patents to an institutional repository and library catalog extend beyond participants in the PTRC Program. Any institution that has an active research program could include patents in their own institutional repository. It would be especially easy for libraries that already work with bepress. Libraries in institutions that are not actively involved in patenting can also benefit from this project. Are there other unique collections that could be added? Perhaps a collection of departmental technical reports? The process developed at Clemson Libraries could easily be adapted to other collections.

REFERENCES

Clemson Libraries. 2015. "Loading a File of Bib Records into Millennium Using Data Exchange."

Technical Services and Collection Management.

<http://library.clemson.edu/depts/tscm/policies-procedures/loading-a-file-of-bib-records-into-millennium-using-data-exchange/>

Crow, R., SPARC (Organization), and Association of Research Libraries. 2002. *The Case for*

Institutional Repositories: A SPARC Position Paper. Washington, D.C: SPARC

http://works.bepress.com/ir_research/7/

Hitchcock, Steve. 2004. "The Effect of Open Access and Downloads ('Hits') on Citation Impact: A Bibliography of Studies." *The OpCit Project*, last modified June 25, 2013,

<http://opcit.eprints.org/oacitation-biblio.html>

Reinman, Suzanne L. 2013. "Patent Research." In *Information Literacy*

Instruction That Works (2nd ed), edited by Patrick Ragains. Chicago: Neal-Schuman. 299-313.

U.S. Patent and Trademark Office. 1977. *Technology Assessment and Forecast: Eighth Report*.

U.S. Washington DC: National Government Publication.

U.S. Patent and Trademark Office. 2011. *General Information Concerning Patents*. U.S.

Washington DC: National Government Publication.

<http://www.uspto.gov/patents-getting-started/general-information-concerning-patents>

Clemson University Libraries. 2015. "Vision, Mission, Goals & Philosophy." *Clemson University*

Libraries. http://www.clemson.edu/library/lib_overview/missvissg.html

Wang, Fang. 2013. "Making IR Content Discoverable." *AALL Spectrum*, 17 (1): 32-33.

<http://www.aallnet.org/mm/Publications/spectrum/archives/vol-17/No-1/ir-discovery.pdf>

APPENDIX 1

PROCEDURES FOR HARVESTING OAI METADATA FROM TIGERPRINTS FOR THE CATALOG

1. Open Marc Edit and go to the Add-Ins menu. Pull down to OAI Harvester Tools, then to Harvest OAI Records. You will see a window that looks like this (see Figure 1, Appendix 1)

OAI Data Harvester

The first time a harvest is done, one needs to input settings into this window. These will be saved after the first time.

Server: <http://tigerprints.clemson.edu/do/oai/>

Set Name:

(for patents): publication:clemson_patents&metadataPrefix=dc

(for dissertations): publication:all_dissertations&metadataPrefix=dc

(for theses): publication:all_theses&metadataPrefix=dc

Metadata: Dublin Core

Crosswalk: C:\Program Files\MarcEdit 6\xslt\OAIDCtoMARCXML.xsl

Don't worry about Advanced Settings.

2. Click OK. You will see a window with various messages such as "Harvesting Records" and a count of how many records are being harvested, along with a green arrow showing the progress of the harvest in numbers of records.

3. At the end of the harvest, you will see a screen giving you information like "resumption token." This is not necessarily something to ignore but for now, you can.

4. A window with MARC records in .mrk format will open. This is the file that you can search and edit.

5. Do your edits via Find/Replace or other MarcEdit utilities.

APPENDIX 2

MODIFICATION TO TIGERPRINTS METADATA RECORDS USING MARCEDIT

All of these modifications should be done prior to loading TigerPrints metadata into Millennium. While all of these but number 4 could be done with Millennium's Global Update, it's a bit more efficient to do them with MarcEdit. The final edit, adding the 008, can be done with MarcEdit ONLY.

MarcEdit creates and manipulates files with three extensions: .mrc and .mrk. Characteristics of these are:

.mrc: this is the file of **raw MARC that is created by the harvester**. It cannot be read with the naked eye. If you "open" an .mrc file, you will open the MarcBreaker window in MarcEdit, and you will have to process the file so that you can read it in the MarcEditor.

NOTE: You MUST load an .mrc file into Millennium. Loading a file with any other extension will not work. (Trust me, I've tried.)

.mrk: this is an eye-readable file that you can save, read, and make modifications to later. You can also resave an .mrk file to be an .mrc file.

In your .mrk file:

1. Delete <p> and </p> tags from 520.

MarcEdit Find/Replace Window

Go to Edit → Find/Replace. You are replacing <p> and </p> with a space. Simply type what you want to find and then insert a space in the Replace box.

2. Remove timestamp from 260.

Go to Tools → Edit Field Data and use the **regular expression T.+** in Find: and a period (.) in Replace.

Edit Field Data Window with a Regular Expression

Be sure to click the checkbox for Use Regular Expressions.

Click Process.

3. Add 008 field to each record. This is a two step process.

A: Go to Edit-> Insert/Edit 008 and pull down to "Electronic Resource." The following window will open:

006/008

006/008 Options:
Type: Electronic Resource

Type: m GPub: \
Audn: \
Form: \
File: u

008 Options:

Srce: \
MRec: \
Lang: \\\
Ctry: \\\
DtSt: s
Dates: 9999 , \\\

OK
Close

MarcEdit Insert/Edit 008 Window

Edit the following fields:

Form: o

File: d

Srce: d

Lang: eng

Ctry: scu

(A backwards slash means blank.)

Click OK. This will create one 008 in ONE record. The =LDR field will append to the end of the 008. Put your cursor at the end of the 008, hit enter. Your 008 will look like this:

=008 141014s9999\\scu\\o\\d\\eng\\d

B: Go to Tools → Add/Delete Field Utility.

Pull down to the 008 field in the window. Copy and Paste the 008 from your .mrk file. Click Add Field. Close.

MarcEdit Add/Delete Field Utility

This will create an 008 for each record in your file.

When you have completed your edits, name and save your file on your computer as an .mrc file by pulling down to File → Compile file into MARC. This is the file you will load into Millennium.

Name it something intuitive such as Patents <date> or ETDs <date> so that you know when you did the harvest.