Rethinking the Digital Media Library for RIT's The Wallace Center

April Younglove

Rochester Institute of Technology

Follow this and additional works at: http://scholarworks.rit.edu/article

Recommended Citation


This Article is brought to you for free and open access by RIT Scholar Works. It has been accepted for inclusion in Articles by an authorized administrator of RIT Scholar Works. For more information, please contact rit scholarworks@rit.edu.
Rethinking the Digital Media Library for RIT's The Wallace Center

Abstract

In 2012, the Digital Preservation Team at Rochester Institute of Technology's The Wallace Center undertook an assessment of its Institutional Repository, the Digital Media Library. The Digital Preservation Team looked at the repository's current performance and requirements for ensuring its future success, and then examined four different philosophical approaches to configuring an IR that are in use by research institutions today: All-in-one, Archival, Researcher-centric and Subject. The team also compared open source vs. in-house repository providers and services including DSpace, Digital Commons, IR+, and Discovery Garden. A decision was made to move from the open source software DSpace to the full-service hosted software bepress from Digital Commons. The new repository will have a narrower research focus and will be rebranded RIT Scholar Works. It will be launched in fall of 2013. This article describes the configuration approaches that were considered and how each approach would impact the Digital Media Library, leading to the final recommendation.

Introduction

Rochester Institute of Technology's The Wallace Center will soon bring major changes to its institutional repository, the RIT Digital Media Library Repository (DML). In late 2013, the DML will be rebranded as RIT Scholar Works. The DML's multi-disciplinary institution-wide acceptance policy will be replaced with a narrower research focus and will only house materials such as faculty written articles and RIT theses. Significantly, RIT's institutional repository will no longer use open source software, but will be built on the proprietary software Digital Commons from bepress. These changes are based on a 2012 review of the DML's performance by the Digital Preservation Team (DPT) at RIT's The Wallace Center (TWC). The team assessed the DML's current performance, looked at alternative repository models, and generated a market analysis of repository solutions before making its decision to head in a new direction.

1 Background and Initial Assessment

The DML was created with the open source software DSpace and was released publicly to the RIT campus in 2002.¹ The DML's webpage offers two search portals — RIT Scholars and RIT Digital Archive — and repository content includes submissions of research materials from the RIT researchers, as well as institutional records from the broader university community, and some special collections materials. RIT customized its DSpace installation's style with custom CSS. It also adjusted the overall structure of its installation with the toolkit Manakin, customized java pages, database tables and more. This customization was extensive and involved the use of student talent to achieve. When students moved...
on to other projects, much of the knowledge about those customization efforts was lost. Due to the level of customization and concerns about developing and maintaining it, TWC chose not to upgrade its version of DSpace and is currently two upgrades behind.

In June 2012, there were a total of almost 13,000 individual items in the DML. This submission total was achieved without the use of any departmental mandates. As of June 2012, according to the website, Ranking Web of Repositories, RIT's DML ranked 40th out of 100 on “Ranking Web’s” list of top USA and Canada repositories. Although the total amount of content in the DML seems admirable, closer analysis reveals that the bulk of the article submissions came in to the DML in the distant past. Recent submissions are primarily required Electronic Theses and Dissertations (ETDs) and a trickling of content from the Academic Affairs office. Most of the communities lie dormant. This pattern of acquiring many submissions early on and then later faltering seems to be typical of Institutional Repositories.2

The Digital Preservation Team (DPT) identified several major barriers to the DML’s success. Notably, the lack of a project manager in the past caused user support and IT support to suffer. Project managers stay in touch with content creators, evaluate services, manage collections and monitor metadata. They also assist with monitoring, debugging and upgrading software, and developing system enhancements.3 Universities typically have either the equivalent of a ½ time position devoted to ongoing support and management of its institutional repository4 or they designate various IR responsibilities, including that of overall manager, among existing staff members.5 In light of this, a project manager was formally identified and other supporting roles were clarified.

In addition, the DML was originally intended to support special collections materials and institutional content, but the DSpace software cannot adequately meet these needs. The Wallace Center stores some of its digitized special collections in the DML; however, both the rare books collection (Cary Collection) and the university’s Archives display most of their digital content in the library catalog, and in a series of uniquely designed online Drupal exhibits. Currently, the DML is not able to replace these catalog entries and piecemeal Drupal exhibits because it does not have suitable metadata categories or the ability to display high resolution image files. The Archives stores some university records in the DML; however, the DML is not capable of standing in for a true Electronic Document and Records Management System (EDRMS) because it lacks key privacy and legal auditing features.6 Specialized software replacements outside of the IR will have to be found for both of these purposes in the future.

DSpace also offers a clunky user interface. Multiple input screens, and long load times between steps of the submission process discourage would be submitters of scholarly content. Other software issues include:

- Staff time and skills required for support and development are high
- Customization doesn’t “stick” upon upgrading
- Initial learning curve for software’s back-end management is steep
- Many administration features can only be accessed via DOS interface

2 Comparing IR Philosophies
Before selecting a software alternative to DSpace, the Digital Preservation Team asked:

- Will the IR continue to have a both an RIT Scholars area and an RIT Digital Archive area or will these be two separate services?
- Who is the IR's intended audience?
- How will RIT's IR relate to its researchers' processes as a whole?
- What role will the IR play in the university's larger archives and preservation process?

Then, the Digital Preservation Team looked at four different philosophical approaches to configuring an IR that are already in use by research institutions today: all-in-one, archival, researcher-centric and subject.

**All-in-one**

The all-in-one digital access model conceives of the IR as a place to deposit many or all of an institution's documents. In this configuration, materials the IR houses could include: finished research papers, research work that is currently in progress, institutional records, special collections materials, promotional materials for the campus, and even visual or audio files for classroom use. The all-in-one strategy overlaps with other strategies mentioned here; however, the primary emphasis of an all-in-one repository is on providing a single place for the entire academic community to store and access its materials.

**Strengths:**

This model can potentially solve many of an institution's digital storage needs with a single solution. Using a single platform to serve many different academic communities instead of purchasing multiple software solutions that are customized for different purposes can also help justify the amount of staff time and money that a library invests in its IR. An all-in-one model might also help promote cross-disciplinary fertilization and give greater visibility to digital materials that would otherwise have been stored in smaller, less-frequented storage spaces. Finally, if the entire campus community is persuaded to use an all-in-one solution, then the university archives would have less work to do transferring files from one software system to another and seeking out obscure locations where valuable digital materials on campus might be located.

**Weaknesses:**

Instead of persuading one user group to submit materials into its IR, librarians promoting the all-in-one repository must get depositing buy-in from the entire campus. Different expectations and storage needs from different populations make outreach challenging. Additionally, when some user groups avoid depositing materials, their empty categories can create the impression for users that the repository is not a success. An all-in-one repository may be rejected by potential depositors if it is not tailored to meet to the depositor's specific needs. It also lacks focus, which may deter users if they are confused by the IR's purpose and either avoid visiting it or fail to realize that it houses desirable materials. Standardizing submission forms and monitoring metadata are also difficult due to format and content variety. Finally, there is the challenge of how to sort temporary use documents from those
with archival value.

*Examples of IRs following this model:*

- RIT Digital Media Library (Contains: theses, research papers, some special collections, some institutional records. Software: DSpace)
- IDEA: Drexel E-repository and Archives (Contains: theses, research papers, special collections, institutional records, library documents and publications. Software: DSpace)

**Archival**

The archival model conceives of the IR as an electronic outgrowth of the physical archives. In this approach, the repository may preserve a single user community's digital documents, or it may preserve digital materials from multiple groups. The focus of the repository is on gathering only materials that are deemed to be of long term value for the institution and preserving them as far into the future as possible. Typically, archival IRs use mandates, schedules of deposit, and personal contact with depositors to ensure that the desired materials are being collected. Often the role of the depositor in submitting materials is minimal and the archives department creates all necessary metadata. This is one of the most widely used IR models within academic institutions.

**Strengths:**

The strength of this approach lies in its clear mission and focus. It is easy to explain the purpose of the IR to possible contributors. In addition, because its mission is strongly tied to preserving historically important materials that administrators value, it can be easier to gain the authority necessary to seek out and collect electronic materials of interest from creators. An archival approach is not a faddish approach and does not risk sinking a lot of time and money into a system that relies on sustained academic interest on the part of contributors (which can wane over time). It also provides a clear path into the existing archives for staff who might otherwise have difficulty determining how to cull important materials from the IR for long term storage. With an archival approach, integration between and institution's past paper records and present electronic documents is clearer than with other models, and the IR experiences less risk of mission creep.

**Weaknesses:**

This method relies on having enough archival staff to support it. Staff is needed to act as gatherers, enforcers, and catalogers. If staff falls behind in their efforts, there is no strong motivation on the part of the contributors to keep up their depositing, and the overall system may deteriorate. If electronic records are not being actively submitted by contributors in a timely manner, they may be deleted rather than deposited. Researchers who wish to use the IR to show off their current and past research may dislike the archives approach as author information is typically stored and displayed in a minimal fashion — a practice that is rapidly losing favor with IRs.

*Examples of IRs following this model:*
Researcher-centric

The researcher-centric model is a newer model that has not yet been widely adopted by college and university repositories. Its roots lie in the open access publishing movement within libraries. This model envisions the IR as a platform for university researchers to collaborate with one another, create customized promotional author pages, store data sets, and even publish their research through the library.

Some researcher-centric IRs are also part of a Current Research Information System (CRIS). The CRIS was developed by the European Organization for International Research Information in order to promote a Common European Research Information Format (CERIF). According to euroCRIS,2 “A CRIS consists of a data model describing objects of interest to R&D and a tool or set of tools to manage the data.” The CRIS is gaining in popularity, but primarily outside of the United States. This is due, in part, to the fact that many European countries are beginning to institute nation-wide standards for research and have identified libraries as key partners in developing and providing platforms for these standardized services.

There is no one particular type of software platform (i.e., commercial or open source) associated with a CRIS. Instead, CRIS’ are typically composed of a variety of tools and websites (including existing institutional repositories) that in concert manage research information in the areas of staff research, publications, project proposals, impact measurement, ethical guidelines, post-graduate research, and key performance indicators.3 This suite of tools can be managed with services like Symplectic or PURE. Several universities are also experimenting with other ways to make their repositories become more like CRIS systems.2

Strengths:

This model may be very attractive to researchers and may boost deposit rates by staff and faculty significantly. It may also be attractive to libraries that are interested in open access publishing. Because it is focused entirely on the needs of researchers, this approach can specifically address problems that are unique to storing research — linking together related research and correctly indexing materials that are authored by research groups. This model also has the potential to store data sets, which is a very real and growing need for those in math and science disciplines. Finally, since it is designed to showcase the work of researchers, it also helps to better advertise the accomplishments of researchers to prospective students and potential funders.

Weaknesses:

There are several obstacles that researcher-centric models must overcome. It is not clear how IRs in
this category will deal with faculty pages after faculty retire or take positions at other colleges. It is also not clear how pages devoted to showcasing the work of research groups will be maintained and archived in the long term since graduate students frequently join and leave these types of groups. There is the risk that this model might be too dependent upon how much researchers like the results. That is, if buy-in is not achieved, the added features for researchers could become liabilities that depositors and users must work around instead of assets that help researchers to leverage their work. Because significant power to edit and manage materials is given to the researchers themselves, there is also a potential challenge with deep archiving materials that may be in a semi-constant state of change. Finally, a research-centric model does not have the capacity to grow into a system that will accommodate other types of materials such as special collections or institutional documents.

Examples of IRs following this model:

- Keele Research Repository (Symplectic)
- UR Research (IR+)
- University of California eScholarship (PKP suite)

Subject

A subject or discipline repository solicits research-related materials (generally open access copies of published articles and preprints) that are associated with a specific area of study rather than with a specific institution. In this sense, the subject repository is not a true institutional repository at all. However, even though a subject repository is heavily influenced by those within a disciplinary community, libraries frequently provide the repository platforms and help to arbitrate the submission process, thus providing an IR-like function.

Strengths:

Subject repositories are enormously popular places to submit and read research within the academic, health and science communities. Because they are community driven rather than institutionally driven, hosting institutions do not need to spend enormous amounts of time doing outreach to the scholarly communities that use the repositories. Depositors are motivated to contribute material to a subject repository because their work will gain visibility on a platform that other professionals in their discipline regularly visit. In addition, most repository users prefer to search for materials in a subject repository rather than in an institutional repository because searchers are typically not looking for information based on where it comes from but rather are looking for research that is associated with a specific topic of study.

Weaknesses:

Subject repositories do not assist an institution in gathering and preserving research generated by its own faculty and staff, so there is the possibility that libraries might be asking researchers to deposit their work multiple times in multiple places in order to properly archive it. Subject repositories only focus on a single area of research and do not assist libraries in preserving any type of non-research related materials. There are also significant long term costs associated with the support and
maintenance of a subject repository that the larger university libraries may not wish to fund since the service provided could be perceived as an optional public good rather than as a key library service. Many subject repositories rely almost entirely on outside funding. Finally, the initiative for creating a subject repository needs to come from within an academic community rather than from a library if it is to be successful. In order for this to occur, those within a discipline must approach an institution based on its existing reputation for digital services. In other words, subject repositories are not a good option for universities that are looking to improve existing services or break into digital preservation for the first time.

*Examples of IRs following this model:*

- **UK Pubmed Central** (Contains: articles and references for open access medical articles available in the UK. Software: PMC)
- **ArXiv** (Contains: electronic preprints of scientific papers in the fields of mathematics, physics, astronomy, computer science, quantitative biology, statistics, and quantitative finance. Software: ArXiv)

### 3 Choosing a Philosophical Model

When it was first created in 2003, the DML was proposed as a “multi-disciplinary database of RIT-generated media.” In other words, it was envisioned not only as a place to store finished research, but also as a [learning object repository](#), a place for instructors to store and share materials associated with teaching, and as a place to access promotional audio and video files for use by the entire campus. While it is possible to maintain a repository system that performs multiple functions for multiple audiences, the lack of clear focus has made marketing the system more difficult and the repository has not met any particular audience's specialized storage needs.

The DPT looked at institutional and departmental mission statements and decided to limit its future IR to content produced by scholarly activities on the RIT campus. Approaches that most closely align with the envisioned IR as a research repository are the archival and researcher-centric models. In the end, the team envisions creating an IR that will be primarily research-centric but that will also work in partnership with Archives in order to ensure that Archives can access and be involved with preserving digital research content that has been identified for long term storage. The Digital Archive portion of the DML will be eliminated, and the IR will instead focus on the RIT Scholars area exclusively.

The team recognized that the decision to discontinue its IR's all-in-one model will have a direct impact on those who use the DML for non-research related materials and that a temporary home for these types of materials will need to be identified before they are removed from the DML. It also acknowledged that software solutions more tailored to archival collections and institutional records will need to be investigated in the future. In addition, TWC is aware that it will need to budget time for the set-up, rebranding, and marketing that will need to take place in order to transition from one repository model to another.
4 Selecting a Research Repository Platform

Hosted vs. in-house IRs

RIT's DPT has likened the library's adoption of the open source platform DSpace as being more like a free puppy than a free beer. That is, it has required a significant amount of time to maintain and needs constant work, which costs money.\textsuperscript{11} It is important for planning purposes to realize that a successful open source repository, whether hosted or in-house, will, in the long run, incur significant monetary costs.\textsuperscript{12} \textit{D-Lib Magazine} recently published research that indicated that “institutions that use open source applications have lower implementation costs but comparable annual operating costs with institutions that use proprietary solutions”.\textsuperscript{13}

Comparing the Options

Because of The Wallace Center’s negative experience with locally installed open source software, and because of the comparable operating costs, the DPT chose to look only at fully or partially hosted software solutions. Several different options were considered for The Wallace Center: staying with DSpace, but purchasing support in order to better maintain customization and meet technical needs; migrating to the full service commercial vendor Digital Commons; migrating to the emerging platform IR+ with the understanding that it will require significant staff technical investment; or selecting Discovery Garden (a vendor that builds its product on the open source product \textit{Islandora}).

Staying with DSpace

Although the \textit{DSpace} software platform has not cost RIT any money in terms of licensing fees, there were costs associated with setting it up, customizing its look and submission functionality, and entering all of RIT’s metadata and files into the system. Migrating to a new system requires a significant prospective cost of time and money upfront to migrate RIT’s existing data, install a new software, adapt it to RIT’s needs and workflow and then work out the initial bugs before release, whereas improving the existing system by purchasing outside support and services would not incur this same initial cost penalty in order to progress.

In order to improve the DML enough to stay with DSpace, several issues would need to be resolved, which could be accomplished with some work. First, in order to reduce the possibility of errors during future upgrades, customization should only be performed on non-key functionalities, in an organized way, and under the proper supervision. Statistical tracking and quality-control procedures would also have to be improved. Quality control checks could be done by either batch editing metadata with spreadsheets in DSpace, by using a data analysis tool like \textit{Google refine}, or by purchasing a module from \texttt{@Mire}. The company \texttt{@Mire} creates commercially licensed modular add-ons for DSpace and sells them à la carte. Some of the existing add-ons available, such as the \textit{Metadata Quality module} and the \textit{Content and Usage Analysis module}, appear to fully resolve some of the existing issues that TWC has with its DSpace installation and claim to be upgrade compatible.
Alternately, full customization and upgrade support could be commercially purchased. Longsight and Open Repository are the leading service providers for DSpace users. Both companies provide server hosting, offer customer support and training, perform upgrades (often seamlessly), can perform software customization, and have data preservation options available. These services can be obtained for a renewable annual fee. Purchasing support service for an open source service has the advantage of not costing any money for the software license itself, and the customer's data is not locked into a proprietary format that is hard to migrate, should the customer decide to end the service contract.

Although solutions could be found that would enable RIT to stay with DSpace, the Digital Preservation Team concluded that they would prefer to find a software that more closely matches what the institution wants straight out of the box instead of sinking more energy into work-arounds. They also noted that the DML's DSpace driven deposit process would need to become more attractive to researchers or a deposit mandate would need to be imposed, and that the DML’s digital storage capabilities would need to be upgraded so that it meets the standards for a preservation system.

**Migrating to Digital Commons**

Digital Commons is a hosted repository software platform and full support service is provided by the Berkeley Electronic Press, aka bepress. According to their website, “Digital Commons offers the features of a traditional IR as well as professional-grade publishing tools and our SelectedWorks™ individual scholar pages … SelectedWorks is an optional add-on to the Digital Commons suite, which allows faculty to create and manage their own personal researcher pages.”

Statistics are available in administration that provide download and referral information for individual articles and departments, or for the whole institution. Authors are notified monthly with download statistics for their papers. Google Analytics is used to provide web traffic information.

Clients of bepress repeatedly praise their excellent customer service, and their finished sites are visually appealing and easy to navigate. Many sites use their IR pages to partner with their University Press. Editkit, bepress’ publishing and editorial management software, is wrapped into a Digital Commons Software as Service (SaaS) contract, and provides five free journals to get users started.

One of the more interesting ways that Digital Commons has been used is described in the April 2011 report "Strategies for Developing an Institutional Repository: A Case Study of ScholarWorks @ UMass Amherst". In the article, author Yuan Lee describes how the University of Massachusetts has worked with their instance of Digital Commons to harvest articles written by UMass faculty from known major subject repositories and then bulk ingest them into their database. Although harvesting articles and metadata from subject repositories is a very efficient way to populate institutional repositories, some argue that such harvesting runs contrary to the spirit of the institutional repository, which is to serve as a primary source of research content rather than as an aggregator of content from elsewhere. Researchers seem unconcerned with such distinctions though, and for the purposes of maintaining high citation numbers in their discipline’s subject repositories, often prefer to have university eprints servers point to their disciplinary repository rather than the other way around.

Currently, although the Digital Commons website does not describe its ability to export data in detail, it does claim that "we have a multi-tiered disaster recovery plan utilizing fail-over servers and regular
on-site and off-site backups. We make copies of repository content available to subscribers so they can create their own back-ups, and support LOCKSS, an OAIS-compliant preservation strategy that provides tools for libraries to back up each other’s content. Libraries that request quotes for services from Digital Commons should inquire about whether Digital Commons can export their particular structural and descriptive metadata.

Cost for a Digital Commons installation is based on Carnegie list FTE schedule total campus population. The basic Digital Commons subscription of five publishable journals can be supplemented with additional journals for a moderate one-time setup fee ($500). Customization such as branding and other options is included. Advanced customization can be obtained for an additional fee.

**Migrating to IR+**

IR+ is a relatively new institutional repository software from the University of Rochester. It’s most attractive features have to do with its researcher-centric nature. After doing an anthropological study on its campus stakeholders, it was determined that faculty and graduate students don't see any need for an institutional repository. They collaborate with others in their discipline rather than with other researchers on campus, and don't need to access work that they have done in the past. Instead, they need to continuously generate new work in order to achieve funding and tenure. The problem for libraries, however, is that they have been charged with gathering, preserving, and showcasing the university's scholarly output. Therefore, IR+ was designed to provide faculty and graduate students with attractive authoring and collaboration tools, network file sharing, and portfolios to showcase current as well as past author works in order to improve submission rates.

The Researcher Pages offered by IR+ are a significant improvement over the default researcher pages offered by DSpace. The most noticeable change is the addition of profile pictures and a section to describe the researcher’s work generally. Another significant improvement is the software's authority control feature for author names. This feature allows contributing authors to identify and connect alternate names in the database that they have published under, thus solving the library staff's headache of trying to link together multiple versions of an author's name manually on the back-end. Finally, researchers can store research in an unpublished state indefinitely and to share documents with networked colleagues while the documents are still unpublished. The spirit of this idea is pro-collaboration, but researchers at RIT are already able to do this type of sharing through email and Google docs and may not need this feature.

While adopting IR+ offers the excitement of getting in on the ground floor of a new and promising software product, with developers who are in close proximity to RIT and may be able to collaborate in person, there are still some drawbacks. The software is immature and a robust user community does not yet exist. At present, a single developer does all of the software support interactions on forums.

**Migrating to Discovery Garden (Fedora/Islandora)**

The [Fedora Repository Project](https://www.fedora.info) (sometimes referred to as just FEDORA, an acronym for Flexible Extensible Digital Object Repository Architecture) was developed by Cornell and is now under the management of non-profit group DuraSpace (which also manages DSpace). Not to be confused with the LINUX operating system of the same name, Fedora is a toolkit that provides an architecture for
developing custom repositories. There are several commercial Fedora SaaS providers. The highly-rated company Discovery Garden creates custom Islandora repositories. Islandora is an open source software mash-up between Fedora and Drupal. Since both types of repositories are customized to customer specifications, installations can differ significantly from one institution to another and can be research focused, collection focused or learning object focused, as required. In addition, Discovery Garden offers server hosting, server maintenance and upgrading, and consultation and preservation services for a fee.

The raw capabilities of Fedora are extensive and could be leveraged to create almost any type of configuration imaginable. In addition, it offers solid pathways to SWORD and iRODS, technologies that could help build a true preservation repository. However, given that TWC would like to spend less time on technical support rather than more, building a repository from scratch with a Fedora toolkit is out of the question. Although Islandora offers the advantage of using Drupal (the same CMS that RIT uses to build its websites) in addition to Fedora, and existing users have certainly used it to generate some repositories with a beautiful look and feel, installing and supporting Islandora on its own would generate about the same amount of work that installing and supporting DSpace current generates. Behind the scenes, management is simple enough, but because it operates on CMS principles, its management is quite different from any other IR software on the market. The administrator and contributor displays involve Solutions Packs (out-of-the-box packages for adding formats such as audio or PDFs) and has a folder and drag and drop menu system. This paradigm is different enough that prospective users coming from a DSpace background should take a careful look at the Islandora sandbox before making any final determinations.

The company Discovery Garden provides Islandora implementation and long term product support. Each installation and set-up of Islandora by Discovery Garden is tailored to the requesting institution. It is not clear how Discovery Garden handles customizations and future upgrades over the long term. Discovery Garden has partnered with third party vendors like Rackspace and SOAR (Truman Technologies) to provide cloud storage including scalable, secure, and almost dark preservation options.

**Selecting a Software**

When considering whether to stay with DSpace or migrate to another platform, consideration was given to the classic sunk cost dilemma. When choosing between staying with a project where considerable time, effort, and money have been invested in it, or abandoning a project and starting over, staying with the project will always prove to be more profitable in the short term. However, in the long term, staying with the project will not be more profitable if the project will eventually fail, albeit slowly. The team asked itself whether the DML is not currently succeeding as much as hoped because of a lack of consistent ongoing project management (which might be achievable with newly increased staffing or with purchased hosting support) or whether the software is responsible for dragging down the progress of the project. The answer appeared to be both, and so plans were made to better support the repository in the future and to look for a new software platform.

Although the idea of being involved with the development of a promising new research-centric software was initially appealing to the team, it was quickly determined that IR+ would not be a good fit
because of TWC's preference for an out of the box solution that will not involve extensive customization or maintenance.

Final consideration was given to both the hosted Fedora service Discovery Garden and the fully-hosted product Digital Commons. After considering bids from both companies and looking more closely at how the differing functionalities aligned with RIT's needs for a new research-based IR, Digital Commons was selected. The team recognizes that an increasing number of people are locating IR materials through general Internet searches rather than through the IR interfaces themselves, and so it also gave considerable weight to Google Scholar's recommendation, "If you're a university repository, we recommend that you use the latest version of Eprints (eprints.org), Digital Commons (digitalcommons.bepress.com), or DSpace (DSpace.org) software to host your papers."

5 Future Plans and Reflection

Although the team is excited about the Research-centric possibilities of creating custom author pages and implementing CRIS features, it was determined that these types of upgrades will not be part of the first release of RIT's new research IR. Once the new research IR is established, faculty and staff will need to be surveyed to ensure that the system is as user friendly as possible and they will be asked questions to determine whether RIT researchers would be interested in CRIS-style features.

Through the process of self-assessment, looking at alternative software platforms, and considering how other universities have chosen to be more or less selective about their missions and content, RIT's The Wallace Center came to the conclusion that some significant changes to its repository were in order. Its existing DML will be broken into several different smaller repositories with clearer missions and tools better suited to differing types of specialized content. Server and software maintenance of its research component will be hosted outside the university. New library staff time will be devoted to promoting the research IR to faculty and staff, managing its quality control, and collaborating with RIT's Scholarly Publishing Studio to a create a place for archiving and displaying research.

In the spring of 2013, The Wallace Center signed a contract with bepress and worked with the company to redesign the look and feel of the DML. Scholarly content (ETDs and articles written by faculty) will be transferred to a new website, RIT Scholar Works. Non-scholarly content will remain in the DML temporarily while the DPT investigates and implements improved storage solutions that are better suited to special collection and institutional document needs. Before the scholarly data can be transferred from the DML to RIT Scholar Works, the metadata will be enriched and edited to meet new higher standards put in place by the project manager. After the scholarly data is transferred, and workflow is formally established, the new site will be beta-tested. The Digital Preservation Team is looking forward to launching RIT Scholar Works to the campus and general public in mid-fall 2013.

References

1. For more details about implementation, see "The evolving impact of institutional repositories on reference librarians" by RIT staff members Marianne A. Buehler and Adwoa Boateng.
2. As Li and Billings (2011) report in their article, "Strategies for Developing an Institutional Repository: A Case Study of ScholarWorks@UMass Amherst", "Many IRs have experienced difficulties in content recruitment after their establishment . . . Since then [2006], developing content recruitment strategies has become a top priority for all IR practitioners".


5. Marilyn S. Billings, "Institutional Repositories Sabbatical report January 30 - July 9, 2005".


7. EroCris "CRIS Concept and CRIS Benefits".


9. For example, see the Harvard repository DASH or Indiana University's OJS based online journal hosting service and a research data publishing service that allows researchers to publish, preserve, and reference datasets for use on current projects.

10. See Leslie Carr's discussion of the topic.

11. For instance, arXiv, Cornell's physics repository also running on DSpace, requires an annual budget of $400,000 and estimates that each submission costs the institution $7 due to staff time and hardware requirements, which is hardly free.


14. Digital Commons, "Digital Commons FAQ".


17. Ben Webster, "What are the downsides of the arXiv?", posted in the arXiv, May 19, 2011.

18. Note that there is a trademark dispute at this time.