

Improving Discoverability, Preventing Broken Links: Considerations for Land-Grant University Publishers

Mark Anderson-Wilk

Follow this and additional works at: <https://newprairiepress.org/jac>



This work is licensed under a [Creative Commons Attribution-Noncommercial-Share Alike 3.0 License](https://creativecommons.org/licenses/by-nc-sa/3.0/).

Recommended Citation

Anderson-Wilk, Mark (2011) "Improving Discoverability, Preventing Broken Links: Considerations for Land-Grant University Publishers," *Journal of Applied Communications*: Vol. 95: Iss. 1. <https://doi.org/10.4148/1051-0834.1175>

This Research is brought to you for free and open access by New Prairie Press. It has been accepted for inclusion in *Journal of Applied Communications* by an authorized administrator of New Prairie Press. For more information, please contact cads@k-state.edu.

Improving Discoverability, Preventing Broken Links: Considerations for Land-Grant University Publishers

Abstract

Publishers of online resources, including Extension publishing units at land-grant universities, are generally concerned about the discoverability of the resources they publish. Broken links create an obstacle between potential audiences and Extension resources, and thus reduce discoverability. This study examines the feasibility of Extension publishers using metadata (information about resources), digital object identifiers (instead of URLs), and/or institutional repositories (in combination with Extension catalogs) to improve the discoverability of their published resources. The use of rich, standardized metadata is recommended as a best practice in digital publishing. The Digital Object Identifier System provides publishers with tools to ensure persistent discoverability; however, the cost and time requirements may be impediments for some Extension publishers. Institutional repositories such as DSpace are underutilized and offer many benefits for Extension publishers to consider. In particular, using an institutional repository may be a low-investment option for Extension publishers to provide both access to and preservation of digital resources.

Keywords

Publishers, feasibility, Dspace, digital resources, metadata

Improving Discoverability, Preventing Broken Links: Considerations for Land-Grant University Publishers

Mark Anderson-Wilk

Abstract

Publishers of online resources, including Extension publishing units at land-grant universities, are generally concerned about the discoverability of the resources they publish. Broken links create an obstacle between potential audiences and Extension resources, and thus reduce discoverability. This study examines the feasibility of Extension publishers using metadata (information about resources), digital object identifiers (instead of URLs), and/or institutional repositories (in combination with Extension catalogs) to improve the discoverability of their published resources. The use of rich, standardized metadata is recommended as a best practice in digital publishing. The Digital Object Identifier System provides publishers with tools to ensure persistent discoverability; however, the cost and time requirements may be impediments for some Extension publishers. Institutional repositories such as DSpace are underutilized and offer many benefits for Extension publishers to consider. In particular, using an institutional repository may be a low-investment option for Extension publishers to provide both access to and preservation of digital resources.

Introduction

A common interest among publishers of online materials is the discoverability of the resources that they publish (Schnittman, 2008; Stranak, 2006; Lossau, 2004). Land-grant universities, as major providers of public education materials, need to improve the discoverability of the resources developed by their communities, including their Extension Services charged with the mission of public outreach (Heatley, 2007). This study evaluates current and potential systems used to facilitate the discoverability of materials published by the communications offices that publish public education materials on behalf of Extension at land-grant universities.

The primary problem addressed by the current study is the reduced discoverability of Extension resources over time due to broken links, also known as link rot. Broken links commonly occur when resources are removed from a website or when a website is reorganized (Stranack, 2006; Paskin, 2006). Users encounter this problem when they either click on a hyperlink or enter an address in their web browser's address bar and they receive a "404 Not Found" error.

Link rot has been identified as a critical problem on the web. Markwell and Brooks (2002) documented the problem of link rot for distance education courses. Over a period of 13 months, 16.5% of the URLs cited in online courses ceased to exist. They concluded, "The progressive disappearance of materials presents a major problem for courses developed to utilize these resources extensively" (Markwell and Brooks, 2002, p. 107). Another study found that "nearly 20% of Internet addresses

in a web-rich high school science curriculum became inactive” in the course of seven months (Delavalle et al., 2003, p. 787).

When URLs are used as launching points for discovery, finding a broken link rather than the expected resource is a loss for both the resource seeker and the resource publisher (and potentially for the third-party referrer as well). As Hellman (2003) puts it, “linking is an activity that enriches both the linker and the linkee” (p. 181). Broken links also create an ongoing web maintenance burden for anyone wishing to keep their website’s links up to date into the future (Kiernan, 2002).

A number of solutions have been developed to help publishers reduce the problem of broken links and increase the discoverability of published resources. This study evaluates the potential for discovery-rich metadata, digital object identifiers (DOIs), and/or institutional repositories to address concerns Extension publishers have regarding the discoverability of the resources they publish. Specific research objectives include (1) assessing whether Extension publishers are using DOIs and/or institutional repositories and (2) determining whether it would be beneficial for Extension publishers to increase their use of discovery-rich metadata, DOIs, and institutional repositories.

Literature Review

A literature review was conducted to provide a foundation for understanding the potential of discovery-rich metadata, DOIs, and institutional repositories to enhance the discoverability of on-line published resources. The state of knowledge on resource discoverability and link persistence has been developed through years of research within the publishing industry, the library community, and information technologists.

Discovery-rich Metadata

Metadata is data about data. In the field of publishing, metadata usually includes key identifying and descriptive information about a published resource such as title, creator, description, publisher, date, and format (Kasdorf, 2003; Brand et al., 2003).

Over the years, a variety of communities have established metadata standards for specific purposes. For example, the Library of Congress’s Machine-Readable Cataloging (MARC) standard was created in the 1960s to help libraries classify and catalog documents (Moura et al., 1998). More recently, the Dublin Core Metadata Element Set of 15 metadata elements was developed as a common metadata vocabulary to facilitate interoperability and information sharing on the Internet (Brand et al., 2003; Awre, 2004; Stranack, 2006; Heatley, 2007).

Metadata has been described as “information that makes data useful” (Moura et al., 1998, p. 221), because a published resource’s metadata affects whether and how the resource is used. Metadata can improve the discoverability of a published resource and can facilitate the exchange of a variety of information related to the resource – for example, in rights management, e-commerce, and library services (Kasdorf, 2003; Brand et al., 2003).

In 2003, Salokhe et al. (2003) argued that if resources were consistently published with well-structured metadata, neither locators such as URLs nor identifiers such as DOIs would be necessary any longer. The metadata itself would be sufficient to allow for accurate search and retrievals of information.

A few years later, however, Ward (2006) bemoaned that common search engines such as Google do not use keyword metadata in their relevance rankings. Even so, Ward argues that we “need metadata not only to record and trace information objects, but to reaffirm their currency of trust, cred-

ibility and authority. . . . Metadata provides the essential link between the information creator and the information user” (p. 2).

In the *Plan to Develop a Digital Information Infrastructure to Manage Land Grant Information*, Randall Heatley (2007) of Michigan State University recommends that land-grant universities develop metadata-rich institutional repositories. He recommends using the Dublin Core Metadata Element Set and metadata harvesting tools to facilitate sharing of resources among land-grant universities.

Digital Object Identifiers

The digital locators (addresses) and identifiers (names) associated with published resources on the web affect their ongoing discoverability and accessibility. Locators and identifiers are related, yet distinct in important ways (Paskin, 2005).

Locators are web addresses that are intended to point users to resources. The uniform resource locator (URL) is a widely recognized type of locator, often being used as a reference link or typed into a browser’s address bar. As noted earlier, URLs cannot always be relied upon to locate a resource because URLs commonly become broken links. A variety of solutions (such as OpenURL and the persistent uniform resource locator [PURL]) have been developed to create URLs that “resolve” through a service that redirects users to a resource’s current location (Rapple, 2004).

Identifiers, in contrast, are unique, permanent names given to resources (Paskin, 2005). The publishing industry has a long history of using identifiers for published works (e.g., the International Standard Book Number [ISBN]). In recent years, digital identification systems (such as the DOI System) have been developed to provide mechanisms for users to find the current web location of a resource using its identifier (its DOI).

Identifiers can be used for any digital media (publications, images, learning objects, etc.) and at any granularity (an identifier could be used for a whole book, for each chapter, or for any sized unit) (Powell, 2003; Paskin, 2006). To be effective, identifiers should have the following qualities: unique, persistent, resolvable, usable in web browsers, usable in nondigital environments (can be printed, dictated over the phone), simple to assign, transportable between locations, and free at the point of use (Powell, 2003). The DOI System provides these qualities and is widely used in digital publishing (Paskin, 2005, 2006).

The DOI System is managed by the International DOI Foundation. Registration agencies handle the registration of DOI names and provide a variety of related services to publishers (Paskin, 2005). A publisher’s choice of registration agency is based on which agency’s services best meet the publisher’s needs. The most widely used registration agency in higher education is CrossRef. CrossRef specializes in reference linking of scholarly works and supports rich metadata for the resources registered (Kasdorf, 2003). R.R. Bowker, the long-time ISBN agency, is now also a DOI registration agency serving the general publishing market.

The publisher works with its registration agency to use DOIs for its published resources. Publishers obtain a DOI name prefix and assign DOI names to resources following a system that ensures uniqueness. The publisher then submits the DOI along with current web location and metadata to the registration agency, which deposits the information in the DOI directory. If the location of the resource changes, the publisher updates the URL in one place (through their registration agency) (Pentz, 2006), as opposed to the URL needing to be updated in every place a URL reference exists, which of course is impractical.

The DOI resolution service can be used in two ways. A DOI name can be entered into the “Resolve a DOI Name” box at the International DOI Foundation website (<http://www.doi.org>) or the DOI can be written in URL form by appending the DOI name to <http://dx.doi.org/> and entering this into the address bar of a web browser. In either case, the user is immediately redirected to the current URL (Paskin, 2005, 2006).

CrossRef also provides a number of resolution and linking services geared toward the needs of researchers – for example, automated cross-publisher reference linking and a function called “forward linking,” which updates an article’s metadata into the future as new articles cite the article (this sometimes appears as a “this article has been cited by ...” feature) (Pentz, 2006).

The DOI System is an implementation of the Handle System (Kahn and Wilensky, 2006; Paskin, 2006). Another application of the Handle System is the DSpace institutional repository software, which refers to digital object identifiers as “handles.”

Institutional Repositories

Institutional repositories are digital asset management systems used to archive and provide access to digital materials produced by the members of a university community (Chan, 2003; Rogers, 2003). These systems are typically open, searchable, metadata-rich, interoperable (able to share information with other systems), and perpetual (Crow, 2002; Ware, 2004).

Institutional repositories have been heralded as revolutionary and transformative – “a new strategy that allows universities to apply serious, systematic leverage to accelerate changes taking place in scholarship and scholarly communication” (Lynch, 2003, p. 1).

Blythe and Chachra (2005) note the great variety of digital media that can be housed in and delivered through institutional repositories – not just text and PDFs, but also photos, audio, video, animations, data sets, session-captured recordings, and learning objects.

According to Clifford A. Lynch (2003), director of the Coalition for Networked Information, “an institutional repository is a recognition that the intellectual life and scholarship of our universities will increasingly be represented, documented, and shared in digital form, and that a primary responsibility of our universities is to exercise stewardship over these riches: both to make them available and to preserve them” (p. 2).

Lynch (2003) also framed the institutional repository as a mechanism for public universities to address their responsibilities to the public: “It is a new channel for structuring the university’s contribution to the broader world” (p. 3). This view of the institutional repository fits with the public service missions of land-grant institutions and the Extension Service in particular.

The value and utility of institutional repositories have been well documented. Foster and Gibbons (2005) interviewed faculty at their institution (University of Rochester) to identify the faculty needs and interests that might be served by an institutional repository. They concluded that the institutional repository function of providing users access to faculty work is the greatest value of repositories from the individual faculty perspective. Though information about the usability of institutional repositories from the information seeker perspective is limited (McKay, 2007), repository collections have been shown to be well discovered through Google searches (Organ, 2006).

Universities customize institutional repository software for their identity and needs. Smith (2002) explains that communities within an institution define what collections they would like to create, and each community defines the policies and procedures it will use to acquire material into its collections. The university library typically assumes asset management responsibilities (Tansley

et al., 2005). Most institutional repositories adhere to standards such as the Dublin Core Metadata Element Set and support the Open Archives Initiative protocol for metadata harvesting. Users can discover materials in institutional repositories through a variety of search mechanisms (Ware, 2004).

The most common institutional repository systems include DSpace developed at the Massachusetts Institute of Technology in 2002, EPrints developed at the University of Southampton in 2001, Digital Commons distributed by Berkeley Electronic Press since 2007, Fedora initially developed at Cornell University in 1997, and Greenstone developed at the University of Waikato in 2000.

The current study focuses primarily on DSpace, which is widely used in the United States (Kim, 2005) and which uses the Handle System to establish persistent identifiers (handles) for published objects (Chan, 2004). Many land-grant universities have DSpace institutional repositories.

Methods

Data were collected using two methods: (1) a survey of individuals working in Extension publishing, and (2) a review of current Extension publishing practices identifiable on the web.

Survey of Extension Publishers

A survey was used to gather information regarding the current practices and decision factors of Extension publishers. An online questionnaire with nine multiple choice questions and one short answer question was developed. The questions were as follows:

- 1) Confirm you have a professional connection to Extension publishing. (yes required to continue)
- 2) Indicate which of the following best represents your position. (unit head/management, editor/production, IT/web developer, or other)
- 3) What types of objects do you publish? (print publications, online publications, videos, images, audio files, including podcasts, websites, e-learning modules, and/or other)
- 4) What audiences do you serve? (public, practitioners, policy makers, researchers, and/or other)
- 5) Indicate your familiarity with DOIs and/or the Handle System. (very familiar, somewhat familiar, or unfamiliar)
- 6) Indicate which of the following best represents your organization's practices related to DOIs/Handle System. (use DOIs/Handle System, have considered but do not use DOIs/Handle System, have not considered using DOIs/Handle System, or don't know)
- 7) Describe your experience with DOIs or other permanent locator systems. (short answer)
- 8) Indicate which potential benefits of DOIs/Handle System are relevant concerns of your organization. (reduce broken links, provide an alternative to long URLs, content appears more prominently and accurately in indexes and search engines, persistent metadata, better management of digital assets and intellectual property, and/or don't know)
- 9) If the annual cost for a publisher to use DOIs were approximately \$500, would this cost affect your decision to implement the DOI system? (yes—would not implement DOIs because of the cost, perhaps—the cost would be a factor in our decision, no—the cost would not be a factor in our decision, or don't know)
- 10) Would the time and effort involved in implementing DOIs (one-time updating of links, entering of metadata into a database when items are published) affect your decision to imple-

ment the DOI system? (yes—would not implement DOIs because of the time and effort, perhaps—would need more information/consideration, no—the time and effort required would not be a factor in our decision, or don't know)

An email invitation to complete the questionnaire was distributed to members of the Association for Communication Excellence publishing and information technology special interest groups through their listservs. Targeted number of recipients was 76. The message stated, "This survey is intended to be completed by one person from units that publish Extension materials; please forward this request to the most appropriate person." The survey was completed by 21 individuals for a 27.6% response rate.

Review of Current Practices

The website of a publisher can be reviewed to determine whether the publisher uses digital object identifiers or an institutional repository to publish its online resources. In order to review the current practices of Extension Service publishers, a list of land-grant institutions was obtained from the Association of Public and Land-grant Universities. A search of the websites of these land-grant institutions found 75 sites with an Extension presence. Of these sites, 61 published digital materials online (through an online catalog database or web page listings) and the remaining 14 did not appear to publish materials online.

The 61 online publishing sites were reviewed to identify which types of online locators/identifiers were being used to direct users from a web page or list of search results to a resource when they click on the link to it. Specifically, the Extension publishers' use of the following was recorded: standard URLs, DOIs, DSpace institutional repository with handles, and/or other institutional repository software with persistent locators.

Findings

The survey asked respondents to confirm their position was in Extension communications/publishing. The responses showed that within Extension publishing the respondents represented unit head/management (38.1%, $n = 8$), editor/production (28.6%, $n = 6$), IT/web development (19.0%, $n = 4$), and other (14.3%, $n = 3$).

The respondents indicated that they publish the following types of objects: printed publications (85.7%, $n = 18$), online publications (81.0%, $n = 17$), websites (71.4%, $n = 15$), images/photographs (52.4%, $n = 11$), audio files (52.4%, $n = 11$), videos (47.6%, $n = 10$), and e-learning modules (28.6%, $n = 6$). The audiences served by the respondents' publishing efforts are presented in Table 1.

Table 1

Audiences served by Extension publishers responding to survey.

Audience	Number of respondents	Percent of respondents
Public	20	95.2%
Practitioners	17	81.0%
Policy Makers	16	76.2%
Researchers	14	66.7%

Note. The survey results included 21 responses from confirmed land-grant Extension publishers.

The survey and the review of current practices together provided data related to understanding of the problem, familiarity with and use of technologies to increase the persistence and discoverability of online resources, and decision factors related to cost, time, access, and preservation.

Understanding of the Problem

Extension publishers responding to the survey overwhelmingly identified the discoverability of the resources they publish as an issue of concern. Specifically, 90.5% (n = 19) identified broken links as a problem, 81.0% (n = 17) indicated a desire for their published resources to appear more prominently and accurately in indexes and search engines, and 71.4% (n = 15) acknowledged the need for information (metadata) about their resources to be more consistently available into the future.

The DOI System and institutional repositories are two solutions that can be used to minimize the problem of link rot and increase the discoverability of resources. Survey respondents were provided a list of specific problems that are addressed by the DOI System and/or institutional repositories and were asked to indicate which of the issues are relevant concerns of their organizations.

The responses confirm that the solutions offered by the DOI System and/or institutional repositories address multiple issues of common concern for Extension publishers (see Table 2).

Table 2

Issues of concern for Extension publishers at land-grant universities.

Issue of concern	Number of respondents	Percent of respondents
Reduce broken links	19	90.5%
Content to appear more permanently and accurately in indexes and search engines	17	81.0%
Persistent metadata	15	71.4%
Better management of digital assets and intellectual property	15	71.4%
Provide an alternative to long URLs	13	61.9%

Familiarity with DOIs and Institutional Repositories

Both the survey and review of current practices involved data gathering related to Extension publishers' familiarity with and use of DOIs and institutional repositories to increase the persistence and discoverability of online digital resources.

When asked to indicate familiarity with the DOI/Handle System, only 14.3% (n = 3) of survey respondents indicated they were very familiar. The remaining indicated they were either only somewhat familiar (42.9%, n = 9) or unfamiliar (42.9%, n = 9). One respondent stated, "We have not considered them before now because we were not familiar with them. However, thanks to your survey, I have read some background material, and based on this, am of the opinion that DOIs could be useful and solve a number of problems."

When asked whether they have considered using the DOI/Handle System, 14.3% (n = 3) of respondents indicated that they currently use the DOI/Handle System. Another 23.8% (n = 5) have considered but do not use the DOI/Handle System. In addition, 33.3% (n = 7) have not considered

using the DOI/Handle System, and 28.6% (n = 6) didn't know whether such a system had been considered.

The review of current practices provided data on how links to online published resources are currently handled by Extension publishers on their websites (see Table 3). As a general practice, Extension publishing websites provide links to published resources using standard URLs from a catalog database or web page listings. No sites were found to use DOIs.

Table 3

How links to online published resources are currently handled by Extension publishers.

Extension publisher practice	Number	Percent
Published objects linked using standard URLs from catalog database of webpage listings	61	100%
Published objects linked using DOI System	0	0.0%
Published objects linked using handles from DSpace institutional repository	9	14.8%
Published objects linked using persistent locators from other institutional repository software	3	4.9%

Note. The total sample of 61 websites includes those land-grant Extension websites that were found to publish digital resources.

While all sites used URLs through a traditional online catalog or webpage listings, some land-grant universities also provide some of their Extension materials through the university's institutional repository. The use of institutional repositories for Extension materials to date appears to focus on specific types of resources (e.g., historical publications and special collections). No Extension publisher exclusively uses an institutional repository for its entire online catalog.

Extension materials that are published using DSpace and its Handle System were identified in the following land-grant institutional repositories: Cornell University's eCommons, Ohio State University's Knowledge Bank, Oregon State University's ScholarsArchive, Texas A&M's Repository, the University of Hawaii's ScholarSpace, the University of Illinois's IDEALS, the University of Maryland's DRUM, the University of Minnesota's Digital Conservancy, and the University of Missouri's MOspace.

Extension materials were also found in other repository systems (e.g., using Berkeley Electronic Press's Digital Commons software) at other land-grant universities. Although these repositories do not use the Handle System, they do offer relatively stable URLs and rich, discoverable metadata.

Decision Factors Related to Cost and Time

Several survey questions asked respondents to assess whether the costs and time associated with implementing DOIs and institutional repositories would affect their decision to invest in the these technologies of discoverability.

The survey provided an estimate of the annual cost for an Extension publisher to use the DOI System (\$500). This estimate was a conservative projection of annual membership to a DOI registra-

tion agency and going-forward (not including past content) DOI deposit fees for an average-sized Extension publisher. When asked if this cost would affect their decision to implement the DOI system, 14.3% (n = 3) responded that they would not implement DOIs because of the cost, 57.1% (n = 12) responded that the cost would be a factor in their decision, 9.5% (n = 2) responded that the cost would not be a factor in their decision, and 19.0% (n = 4) responded that they didn't know whether the cost would affect the decision.

Survey respondents provided a number of specific comments regarding costs: "Due to recent budget cuts we are struggling to keep up with current demands but hope to move this up in our list soon." "Fees and work required looked like more than we wanted to deal with. We're trying to save money, this year especially." "I would say our Extension IT group is updating its e-publishing after about 10 years of almost standing still. Our group is also very conservative about perceived value of expenditures."

The survey described the time and effort required to implement the DOI/Handle System as follows: A one-time updating of links used on web pages and in publications catalogs and ongoing entering of metadata (title, description, location, etc.) into a system as items are published. (Note that within a university, installation and management of institutional repository software are typically handled centrally through its library and not the responsibility of participating university units/communities.) When respondents were asked if the time and effort required of them would affect their decision to implement the DOI/Handle System, 9.5% (n = 2) said they would not implement DOIs because of the time and effort, 66.7% (n = 14) said they would need more information to determine whether the time and effort would be worth it, 14.3% (n = 3) said the time and effort required would not be a factor in their decision, and 9.5% (n = 2) didn't know whether this would be a factor. One respondent indicated that DOI/Handle System implementation did not have high priority because it did not have a deadline.

Decision Factors Related to Preserving Resources versus Controlling Access to Current Information

Extension publishers face the challenge of at once providing a wide variety of audiences access to current research-based information and at the same time ensuring that the knowledge legacy of past investments are preserved. Survey respondents confirmed this dual demand by indicating that both discoverability (access to current information) and persistence (preservation of digital assets) are issues of concern.

One respondent explained his or her unit's experience as follows: "Our library has in the past assigned persistent URLs to born-digital extension publications for linking from the library catalog. Extension publications, however, have a review/revision cycle that takes many offline after their period of relevance has passed. We're still working out the kinks in a workflow to transition the publications from active distribution on the Extension website to digital archive and access from the library catalog. We use 'friendly URLs' on our website."

Discussion Cost and Time

The perceived cost and time required to implement the DOI System appear to be barriers to adoption. In particular, the benefits of the DOI System are generally not perceived to be great enough to justify the relatively modest investment necessary to benefit from the system.

While the cost of implementing the DOI System appears to be an impeding factor to implementation, using an institutional repository might be a more affordable option with the same benefits for Extension publishers. Investment in an institutional repository is typically made through the university library. The institutional repository approach involves a university community (in this case, Extension publishing units) partnering with the institution's library, with the library assuming some of the digital asset management responsibilities (Crow, 2002; Foster & Gibbons, 2005).

Preserving Resources and Controlling Access to Current Information

Inconsistent use of institutional repositories represents a missed opportunity to utilize the benefits of repositories to their fullest. For example, when institutional repositories are used only for historical resources, not current resources, the problem of link rot is perpetuated. That is, new resources are fated to initially be located using a URL in the Extension catalog/website, and then later be transferred to a new location with a handle identifier or persistent URL. All existing web links to the resource would thus become broken links unless they are manually updated.

Making resources permanently available could raise concerns within the Extension Service, especially regarding sensitive information that regularly changes as research findings and/or regulations change (e.g., pesticide application recommendations, food preservation practices). Although printed Extension materials have been preserved for perennial access in libraries for nearly a century, the openness and ease of access of institutional repositories could be cause for hesitation.

However, systems could be developed to minimize these potential problems. For example, Extension publishers could develop a model where Extension resources are initially published in an institutional repository and where the Extension catalog/website is used as a gate for users to access only the resources with current content, using the institutional repository's handle identifier (or persistent URL). Out-of-date resources in the repository could be tagged with a date and disclaimer for users who locate the resources through the repository interface rather than the Extension catalog/website.

Other Considerations

Extension publishers have many considerations to weigh in the decision-making process of whether to implement a discoverability technology, which one, or which combination.

For example, while most publishers are moving toward online digital content, some Extension publishers still have a number of resources whose full content is available only in physical print form. These publishers may be interested in publishing systems that can provide information about a product and how to purchase it without actually serving the product itself digitally through the web. The DOI System technically can be used to point to metadata for resources that are print-only as well as those that exist digitally. Institutional repositories, on the other hand, are designed specifically to hold digital resources; repositories generally are not used as indexes containing metadata without their associated resources.

Another consideration relates to charging access fees. The DOI System can be used with systems that charge users to access resources. Most institutional repositories, on the other hand, are designed to be open access and so are not well suited for resources that have access fees.

Non-exclusivity of Approaches

Discovery-rich metadata, the DOI System, and institutional repositories are not mutually exclusive. The use of standardized, harvestable metadata can be implemented using any approach,

including the DOI System, institutional repositories, and homegrown publishing systems. The DOI System itself is not a full publishing platform. It is a system to be used in conjunction with whatever mechanism is being used to publish resources. An institutional repository can serve as a publishing platform itself or in combination with a publisher's site. Users can access a repository from the repository's interface or through a separate search engine, index, or catalog system.

Conclusions

Extension programs at land-grant institutions have a strong tradition of publishing a great number of publications and other educational materials on the web. These resources are digital assets that should be managed both for discoverability and preservation. Scholarly publishers and university libraries have widely adopted the DOI System and institutional repositories, respectively, to preserve digital information and make it available on an ongoing basis. This study examines the applicability and feasibility of Extension publishers implementing harvestable metadata standards, digital object identifiers, and/or institutional repository collections to address the common concerns of discoverability and persistence.

The following conclusions and recommendations are drawn:

- Extension publishers recognize the importance of discoverability and the problem of broken links.
- Digital object identifiers and institutional repositories both provide solutions that can help Extension publishers address their concerns related to discoverability and persistence.
- Best metadata practices can be used with either the DOI System or an institutional repository.
- Most Extension publishing offices have not had great enough familiarity with the DOI/Handle System or institutional repositories to make informed decisions regarding their adoption.
- Extension publishers generally do not perceive the benefits of the DOI System to be greater than the financial investment required to implement it.
- Institutional repositories are being inconsistently used and underutilized.
- Extension publishers at land-grant universities with institutional repositories may find use of the repository to be a cost-efficient strategy for maintaining stewardship of digital resources.

Because "Extension publishing" is heterogeneous, a single discoverability strategy cannot fittingly apply to all Extension publishing units. Each publisher needs to make their decisions on how their digital resources are published based on their particular needs and conditions.

About the Author

Mark Anderson-Wilk is the publishing leader in Extension and Experiment Station Communications at Oregon State University. He has worked in a variety of positions in publishing over the last 20 years.

Keywords

access, digital asset management, digital object identifier, digital publishing, discoverability, institutional repository, metadata

References

- Awre, C. (2004). The JISC's FAIR Programme: Disclosing and sharing institutional assets. *Learned Publishing*, 17(2), 151-156. Retrieved May 4, 2009, from <http://dx.doi.org/10.1087/095315104322958535>
- Blythe, E., & Chachra, V. (2005). The value proposition in institutional repositories. *EDUCAUSE Review* 40(5), 76-77. Retrieved May 4, 2009, from <http://www.educause.edu/ir/library/pdf/ERM0559.pdf>
- Brand, A., Daly, F., & Meyers, B. (2003). *Metadata demystified: A guide for publishers*. Hanover, PA: Sheridan Press, Hanover. Retrieved May 4, 2009, from http://www.niso.org/standards/resources/Metadata_Demystified.pdf
- Chan, L. (2004). Supporting and enhancing scholarship in the digital age: The role of open-access institutional repositories. *Canadian Journal of Communication*, 29(3). Retrieved May 4, 2009, from <http://www.cjc-online.ca/index.php/journal/article/viewArticle/1455/1579>
- Crow, R. (2002). *The case for institutional repositories: A SPARC position paper*. ARL Bimonthly Report 223. Washington, DC: Association of Research Libraries. Retrieved May 4, 2009, from <http://hdl.handle.net/123456789/315>
- Dellavalle, R. P., Hester, E. J., Heilig, L. F., Drake, A. L., Kuntzman, J. W., Graber, M., & Schilling, L. M. (2003). Going, going, gone: Lost internet references. *Science*, 302(5646), 787-788. Retrieved May 4, 2009, from <http://dx.doi.org/10.1126/science.1088234>
- Foster, N. F., & Gibbons, S. (2005). Understanding faculty to improve content recruitment for institutional repositories. *D-Lib Magazine*, 11(1). Retrieved May 4, 2009, from <http://dx.doi.org/10.1045/january2005-foster>
- Heatley, R. (2007). *Plan to develop a digital information infrastructure to manage land grant information*. Lincoln, NE: American Distance Education Consortium. Retrieved May 4, 2009, from <http://www.adec.edu/adec-agnic-digital-inf.pdf>
- Hellman, E. (2003). OpenURL: Making the link to libraries. *Learned Publishing*, 16(3), 177-181. Retrieved May 4, 2009, from <http://dx.doi.org/10.1087/095315103322110950>
- Kahn, R., & Wilensky, R. (2006). A framework for distributed digital object services. *International Journal on Digital Libraries*, 6(2), 115-123. Retrieved May 4, 2009, from <http://dx.doi.org/10.1007/s00799-005-0128-x>
- Kasdorf, W. E. (2003). *The Columbia guide to digital publishing*. New York, NY: Columbia University Press.
- Kiernan, V. (2002). Nebraska research measure the extent of 'link rot' in distance education. *Chronicle of Higher Education*, April 10, 2002. Retrieved May 4, 2009, from <http://chronicle.com/free/2002/04/2002041001u.htm>
- Kim, J. (2005). Finding documents in a digital repository: DSpace and Eprints. *Proceedings of the American Society for Information Science*. Retrieved May 4, 2009, from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.103.5619&rep=rep1&type=pdf>

- Lossau, N. (2004). Search engine technology and digital libraries: Libraries need to discover the academic internet. *D-Lib Magazine*, 10(6). Retrieved May 4, 2009, from <http://dx.doi.org/10.1045/june2004-lossau>
- Lynch, C. A. (2003). Institutional repositories: Essential infrastructure for scholarship in the digital age. *ARL Bimonthly Report* 226 (February 2003), 1-7. Washington, DC: Association of Research Libraries. Retrieved May 4, 2009, from <http://www.arl.org/resources/pubs/br/br226/br226ir.shtml>
- Markwell, J., & Brooks, D. W. 2002. Broken links: The ephemeral nature of educational WWW hyperlinks. *Journal of Science Education and Technology*, 11(2), 105-108. Retrieved May 4, 2009, from <http://dx.doi.org/10.1023/A:1014627511641>
- McKay, D. (2007). Institutional repositories and their 'other' users: Usability beyond authors. *Ariadne*, 52(July 2007). Retrieved May 4, 2009, from <http://www.ariadne.ac.uk/issue52/mckay/>
- Moura, A. M. C., Campos, M. L. M., & Barreto, C. M. (1998). A survey on metadata for describing and retrieving Internet resources. *World Wide Web*, 1(4), 221-240. Retrieved May 4, 2009, from <http://dx.doi.org/10.1023/A:1019228117935>
- Organ, M. (2006). Download statistics—what do they tell us? *D-Lib Magazine*, 12(11). Retrieved May 4, 2009, from <http://www.dlib.org/dlib/november06/organ/11organ.html>
- Paskin, N. (2005). Digital object identifiers for scientific data. *Data Science Journal*, 4, 12-20. Retrieved May 4, 2009, from <http://dx.doi.org/10.2481/dsj.4.12>
- Paskin, N. (2006). Naming and meaning of digital objects. Presented at the 2nd International Conference on Automated Production of Cross Media Content for Multi-channel Distribution, Leeds, UK, December 2006. Retrieved May 4, 2009, from <http://www.doi.org/topics/060927AXMEDIS2006DOI.pdf>
- Pentz, E. (2006). CrossRef at the crossroads. *Learned Publishing*, 19, 250-258. Retrieved May 4, 2009, from <http://dx.doi.org/10.1087/095315106778690760>
- Powell, A. (2003). Identifiers for learning objects—a discussion paper. Bath, UK: UKOLN, University of Bath. Retrieved May 4, 2009, from <http://www.ukoln.ac.uk/distributed-systems/lo-identifiers/>
- Rapple, C. (2004). After the goldrush—the golden age of reference linking. *Learned Publishing*, 17(4), 299-304. Retrieved May 4, 2009, from <http://dx.doi.org/10.1087/0953151042321653>
- Rogers, S. A. (2003). Developing an institutional knowledge bank at Ohio State University: From concept to action plan. *Libraries and the Academy*, 3.1, 125-136. Retrieved May 4, 2009, from http://muse.jhu.edu/journals/portal_libraries_and_the_academy/v003/3.1rogers.html
- Salokhe, G., Weinheimer, J., Bovo, M. G., Agrimi, M., Keizer, J., & Katz, S. (2003). Structured metadata for direct resource location: A case study. Proceedings of the International Conference on Dublin Core and Metadata Applications, Seattle, Washington, September 28, 2003. Retrieved May 4, 2009, from <http://dcpapers.dublincore.org/ojs/pubs/article/view/743/739>
- Schnittman, E. (2008). Discoverability and access in book publishing. *Publishing Research Quarterly*, 24(2), 139-142. Retrieved May 4, 2009, from <http://dx.doi.org/10.1007/s12109-008-9058-y>
- Smith, M. (2002). DSpace: An institutional repository from the MIT Libraries and Hewlett Packard Laboratories. *Lecture Notes on Computer Science*, 2458, 543-549. Retrieved May 4, 2009, from <http://hdl.handle.net/1721.1/26706>

- Stranack, K. (2006). Getting found, staying found, increasing impact: Enhancing readership and preserving content for OJS journals. Vancouver, BC: Public Knowledge Project. Retrieved May 4, 2009, from <http://pkp.sfu.ca/files/GettingFoundStayingFound.pdf>
- Tansley, R., Smith, M., & Walker, J. H. (2005). The DSpace open source digital asset management system: Challenges and opportunities. Research and Advanced Technology for Digital Libraries. Heidelberg: Springer. Retrieved May 4, 2009, from http://dx.doi.org/10.1007/11551362_22
- Ward, F. L. H. (2006). Search engines and metadata. Rome, Italy: Food and Agriculture Organization, United Nations. Retrieved May 4, 2009, from ftp://ftp.fao.org/gi/gil/gilws/aims/publications/papers/metata_and_search_engines.pdf
- Ware, M. Institutional repositories and scholarly publishing. *Learned Publishing*, 17(2), 115-124. Retrieved May 4, 2009, from <http://dx.doi.org/10.1087/095315104322958490>