Introduction

Burton Callicott, David Scherer, and Andrew Wesolek

HISTORY OF REPOSITORIES: HOW WE GOT WHERE WE ARE

Institutional repository initiatives consist of a suite of services intended to support the preservation and organization of, and access to, the intellectual output of the institution in which they are housed. The institutional repository (IR) itself typically refers to the software infrastructure on which these initiatives depend. More than that, though, institutional repositories were developed to be a solution to some of the problematic aspects of scholarly communication in a digital age. Specifically, they were and continue to be seen as a way to introduce competition to a monopolistic traditional publishing system by offering the possibility of immediate publication, long-term preservation, and barrier-free global access to those publications.

The promise of repositories in general was immediately apparent with the launch of the disciplinary-specific repository arXiv in 1991. On its debut, electronic communication of scholarly literature via this preprint server was rapidly embraced by high-energy physicists, and this has since expanded to include related areas of physics and mathematics while hosting more than one million EPrints. The revolutionary potential of this new mode of communication was recognized and embraced soon after its launch, and as early as the mid-1990s, some began recognizing the broader potential of such repositories to revolutionize traditional scholarly communication systems (Ginsparg, 1997).

Beginning in the early 2000s, the potential of disciplinary repositories to disseminate scholarship immediately and openly began to be applied at the institutional level. The year 2002 marked a watershed, seeing the first public release of the open source institutional repository software DSpace, along with the publication of the Scholarly Publishing and Academic Resources Coalition (SPARC) position paper, *The Case for Institutional Repositories* (Crow, 2002). These two events provided broadly accessible software support for institutional repositories, as well as a compelling case which tied institutional repository initiatives to institutional visibility and prestige. As a result, institutional repository programs began to grow at an exponential rate, now numbering in the thousands.

GROUNDING THE VISION

Although institutional repositories had lofty goals and intentions, the actual practice of repositories, and the activities undertaken to populate them, did not match the zeal of the library community. In her canonical 2008 article, Dorothea Salo (2008) stated that academic libraries were enticed into the wind and that the whole project might have been a waste of time. Many institutional repositories encountered unforeseen problems and a surprising lack of impact. Clunky or cumbersome interfaces, lack of value and use by scholars, fear of copyright infringement, and the like tended to dampen excitement and adoption.

Even today, libraries that have repositories (or those considering whether or not to take the plunge) have been questioning:

- What are the best containers/platforms?
- Should we host or not host?
- What are the best ways to make content visible and discoverable?
- What is the role of IRs in providing "green" open access to work published elsewhere?
- What should go in (and what should be kept out)?
- What is the role of IRs in being publishing platforms for original and unique institutional publications?
- What measures of success matter? Which measurements matter to whom?
- How are access and use measured—downloads, altmetrics, and so on?
- What is the impact of an institutional repository?

While repository initiatives have had to fight an uphill battle, widespread adoption and use indicates that they are here to stay and will have an impact in the evolution of scholarly communication. Libraries and those within the libraries who manage repositories have learned through their experiences and have demonstrated that the initial problems that they encountered can be overcome and that successful institutional repository initiatives are possible and replicable.

STRATEGIES FOR SUCCESS

Making Institutional Repositories Work takes newcomers as well as seasoned practitioners through the practical and conceptual steps necessary to have a successful IR customized to the goals and culture of their home institutions. Over the course of the last 10-plus years, much digital ink has been spilled discussing and debating the more technical aspects of IRs including platform design, methods of integrating datasets, open access initiatives, copyright considerations, and so forth. The result is a lack of practical and straightforward literature available to those considering an IR initiative at their institutions and for current practitioners seeking to increase the success of their current repository initiatives in a holistic way.

Making Institutional Repositories Work intends to fill this void. We asked several established and highly regarded experts in the world of institutional repositories to take a step back from the theoretical and highly technical details surrounding repository initiatives and share their real-world experiences, observations, and premonitions about the practice and shape of repositories. This volume contains their experiences, case studies, and strategies for success, as well as their perceptions on the future of institutional repositories and their role within the scholarly communication landscape.

THE STRUCTURE OF THE VOLUME

This volume is arranged in four thematic parts intended to take the pulse of institutional repositories—to see how they have matured and what can be expected from them, as well as to introduce what may be their future role. To keep the content grounded and practical, the volume also contains a series of case studies in which librarians at institutions of different sizes, repository platforms, and research focuses describe how and why they

initially created IRs and how the role of the IR has evolved. The work concludes with a vision of the future of IR initiatives by detailing some of the challenges they face and strategies for sustained success.

PART 1: CHOOSING A PLATFORM

In the broadest sense, an institutional repository initiative seeks to capture the intellectual output of an institution and make it openly available in perpetuity. Launching such an initiative requires specialized software. Part 1 will focus on the many repository platform options available and the desired outcomes that influence software decisions. Does the institution wish to invest in the technical staff to develop its own repository or support an open source solution? Or would it be better suited to an "out of the box" proprietary option? In addition, Part 1 covers content and how these decisions will impact platform choice: what types of items will the repository hold? Articles, theses and dissertations, datasets, library-published materials? Finally, to what degree is discoverability important? If it is, what are some steps that the institution can take to enhance the discoverability of its repository's content?

Chapter 1, "Choosing a Repository Platform: Open Source vs. Hosted Solutions," by Hillary Corbett, Jimmy Ghaphery, Lauren Work, and Sam Byrd, lays out the major considerations that go into selecting an institutional repository platform. Those new to repositories will discover that what may appear to be a murky and even scary array of factors to consider can become quite clear with a simple assessment of key components. The chapter also offers insights and advice to readers who have an existing repository but are considering a platform change. The authors outline the major differences between open source and proprietary systems using DSpace/Fedora and Digital Commons to illustrate the relative advantages of each system. Drawing from the experience of Virginia Commonwealth and Northeastern University, separate sections detail the processes and considerations that go into switching from an open source to a proprietary system as well as the reverse.

These initial platform decisions will also have an impact on the types of data storage services that may be offered as part of a repository initiative. The use and sharing of research data is of increasing interest to funders and publishers, but repositories are often responsible for the long-term

storage of and access to this information. Chapter 2, "Repository Options for Research Data," by Katherine McNeill, discusses the relationship between research data and repositories, in light of the experiences of the Massachusetts Institute of Technology (MIT) library system. McNeill examines the various types of data repositories currently available to academic institutions, noting key differences on several important characteristics. She suggests that institutions need to consider several questions when either developing a stand-alone data repository service or accepting research data into an existing institutional repository. She points out that there is no single solution, and for the foreseeable future, institutions will choose between varying repository models that will best fit the needs of their local context, and enable the best models of data storage and sharing.

The most beautiful and intuitive repository interface is functionally useless if its content is not discoverable by researchers. In Chapter 3, "Ensuring Discoverability of IR Content," Kenning Arlitsch, Patrick OBrien, Jeffrey K. Mixter, Jason A. Clark, and Leila Sterman explore the key factors that will reliably enhance search engine optimization. They start with metadata and provide tips that can enhance efficiency as well as effectiveness. The authors then provide suggestions for structuring IR sites that will enable search engine crawlers to more readily index content. In addition to providing useful approaches to data maintenance and cleanup, the authors outline best practices that will minimize overhaul work as search engines evolve, and as repositories become more integrated into various databases and new modes of research strategy.

PART 2: SETTING POLICIES

After selecting a platform to support an institutional repository, one must consider which policies are to be put in place. Part 2 examines the theoretical aspects and practical applications of two important policy decisions: institutional open access policies and published theses and dissertations.

Due in large part to the advent of repositories, many colleges and universities have passed or are in the process of passing open access policies. In Chapter 4, "Open Access Policies: Basics and Impact on Content Recruitment," Andrew Wesolek and Paul Royster explore the different types of open access policies currently in place and discuss steps and methodologies that can lead to development and passage. Wesolek served as scholarly

communication librarian at an open access policy institution and as the chair of the Coalition of Open Access Policy Institutions (COAPI), while Royster manages the remarkably successful institutional repository at the University of Nebraska–Lincoln, an institution that made the conscious decision not to pursue passage of an open access policy. From these differing perspectives, Wesolek and Royster seek to answer the question, "Are open access policies necessary for successful repositories?"

In Chapter 5, "Responsibilities and Rights: Balancing the Institutional Imperative for Open Access with Authors' Self-Determination," Isaac Gilman makes a broader investigation of the ethical dimensions of an open access policy. Gilman makes the case that institutions have a clear and often explicitly stated goal of making locally created knowledge openly available to the world, while faculty and students, as rights holders, have an equally clear right to self-determination. He concludes that institutional repositories should play an essential role in fulfilling an institution's mission to share knowledge as broadly as possible while respecting faculty rights.

We then focus on the more concrete aspects of open access policy implementation with Chapter 6, "Campus Open Access Policy Implementation Models and Implications for IR Services," by Ellen Finnie Duranceau and Sue Kriegsman. In this chapter, the authors offer a snapshot of the institutional open access policy implementation landscape in an effort to build a roadmap for others moving forward in this "nuanced" environment. The authors report on a survey conducted by the COAPI that was designed to discover and chart the scope of the coalition membership's policies and their methods of implementation. Based on the data from this survey, Duranceau and Kriegsman provide a suite of strategies modeled on institutions with open access policies in place that have been employed to both meet faculty needs and successfully populate institutional repositories.

Gail McMillan then covers the most fundamental content of institutional repositories, electronic theses and dissertations (ETDs), in Chapter 7, "Electronic Theses and Dissertations: Preparing Graduate Students for Their Futures." Here, McMillan outlines some of the policy considerations associated with integrating an ETD program into an institutional repository. Institutional missions and ETD stakeholders as well as the impact these policy decisions may have on student-authors are discussed and contextualized.

Finally, Megan Banach Bergin and Charlotte Roh discuss key aspects of the ETD policy decisions made at the University of Massachusetts Amherst in Chapter 8, "Systematically Populating an IR with ETDs: Launching a Retrospective Digitization Project and Collecting Current ETDs." Ultimately, both chapters recommend empowering student-authors through educating them about their rights as authors in a landscape that is rapidly shifting toward open access.

PART 3: RECRUITING AND CREATING CONTENT

Once a platform is in place and policies have been adopted, institutional repository managers can begin to focus on content. From previously published materials, expanding forms of gray literature and other existing works, to the emerging field of repository-based publishing programs, the chapters in Part 3 cover the array of content that can potentially be added to an IR. This part also outlines challenges that institutions can face in terms of marketing IR services and soliciting scholarship while presenting strategies to meet and rise above real and perceived recruitment barriers.

In Chapter 9, "Faculty Self-Archiving," Stephanie Davis-Kahl identifies faculty resistance to self-archiving journal articles in institutional repositories. Davis-Kahl argues that while open access has become increasingly accepted, and IRs have contributed to that acceptance, there are still many points of confusion and concern regarding repository self-archiving practices including (but not limited to) repository awareness, copyright, time, perceptions of self-archived materials, and disciplinary culture and practices. She suggests that faculty use and perceptions of research may shift with the use of social media programs, such as ResearchGate and Academia.edu, to engage with faculty by enhancing the activities and practices faculty use to interact and communicate with colleagues and disciplinary counterparts.

As many early adopters have demonstrated, a repository cannot become successful by simply being built, regardless of the quality of the platform. In Chapter 10, "Incentivizing Them to Come: Strategies, Tools, and Opportunities for Marketing an Institutional Repository," David Scherer discusses that while repositories continue to emerge, they have not lived up to their expectations for growth and coverage. Based on his experience at Purdue, Scherer provides tried and true methods that can lead to

a diverse, active, and constantly evolving marketing plan that emphasizes benefits and incentives to stakeholders, repository offerings, and additional resources that increase participation and use.

As libraries begin to collaborate with university presses at an ever expanding rate, institutional repositories are staged to play an important and active role in these new and budding programs and partnerships. In Chapter 11, "Repository as Publishing Platform," Simone Sacchi and Mark Newton discuss why institutional repositories are in a position to provide opportunities for current and future researchers to better understand the scholarly communication/publication process, how the institutional repository can be utilized as a publishing platform, and what may be the future of repository-based publishing.

Not only can repositories serve as new venues for publishing models, they can also serve as new training grounds to inform and educate those involved in the publication process, ranging from students becoming acclimated to academic activities and dissemination, to academic journal editors interested in new publishing models. Chapter 12, "Publishing Pedagogy: The Institutional Repository as Training Ground for a New Breed of Academic Journal Editors," by Catherine Mitchell and Lisa Schiff, explores the role of the institutional repository as a pedagogical tool and resource for campus stakeholders on several publishing topics and activities, including copyright, licenses, types and quality of peer review, and journal sustainability and business models. Mitchell and Schiff also discuss how their interactions with campus stakeholders have informed the California Digital Library (CDL) development plans and policies for the University of California's institutional repository, eScholarship.

PART 4: MEASURING SUCCESS

This final thematic part attempts to encapsulate all the tools and data that can reliably measure IR success for managers, contributors, users, departmental and institutional administrators, and other stakeholders. It seeks to answer the question, "So, I have an IR; now, how do I know that it is effective?"

In Chapter 13, "Purposeful Metrics: Matching Institutional Repository Metrics to Purpose and Audience," Todd Bruns and Harrison W. Inefuku tackle IR metrics that can be generated through repository platforms as

well as third-party sources such as Google Analytics and Altmetrics. They provide methods of turning raw metric data into useful information parsed to the appropriate audience and purpose. The authors outline the ways that metric data captured and presented correctly can provide an avenue for establishing institutional repositories as an integral technology in the research enterprise of the institution.

Kim Holmberg, Stefanie Haustein, and Daniel Beucke introduce readers to the rapidly evolving and increasingly important realm of altmetrics. Chapter 14, "Social Media Metrics as Indicators of Repository Impact," inventories and assesses the various means of measuring impact through social media. They show how these measures can bring to light potentially more timely, granular, and nuanced measures of use and impact than what has been used previously. The chapter presents concrete examples from institutions that currently employ altmetrics as well as a likely future of this burgeoning approach to assessment.

Tacking away from raw numbers and metrics, "Peer Review and Institutional Repositories" (Chapter 15), by Burton Callicott, addresses the potential impact IRs may have on the peer-review system and the ways in which IRs may begin to play a significant role in credentialing and assessing scholarship. By exploring the ways in which gray literature has risen in prominence, availability, and legitimacy due to its inclusion in IRs, this chapter charts the ramifications this may have for "white" or more traditional scholarly publications—journal articles and monographs. Due to the radical increase in production of scholarship and the role of the repository in the process, this chapter also describes publishing trends and avenues of scholarly communication that are affected by repositories and the concomitant effect this will likely have on the peer-review system.

Marianne A. Buehler's "Defining Success and Impact for Scholars, Department Chairs, and Administrators: Is There a Sweet Spot?" (Chapter 16), the final chapter in Part 4, attempts to bring all the various assessment measures together such that they have value and resonance for all the major institutional constituents. Buehler outlines the ways that the primary interests of scholars and administrators may seem at odds on some levels but when viewed holistically can be seen to have shared goals that can be documented and graphed when success measures are implemented and reported in a way that reveals the "sweet spot" that has resonance and value for all involved.

PART 5: INSTITUTIONAL REPOSITORIES IN PRACTICE: CASE STUDIES

Part 5 presents four case studies from institutions of varying size and mission that describe the implementation and application of the concepts and activities described in the previous parts.

Princeton University

In "Creating the IR Culture" (Chapter 17) Anne Langley and Yuan Li present a case study that maps out the creation of the institutional repository culture at Princeton University. Langley and Li describe how their experience was unique due to the fact that, unlike at most schools, an open access policy predated their repository. They emphasize the creation of a strong base of support across campus by telling the story of open access, while also being careful how the message was created to fit the needs of their audience.

College of Charleston

James Tyler Mobley takes readers through the decision-making process that led to an open source (DSpace) repository at the College of Charleston in "On Implementing an Open Source Institutional Repository" (Chapter 18). The case study illustrates the realities that many mid-sized state schools face when they want to play a part of the IR movement. Based on his experience, Mobley outlines what is required and what can be expected when the choice is made to go with a "free," open sourced platform with a limited number of staff members who have various levels of expertise and coding skills. As anyone who has attempted to employ open source software knows, unlicensed applications invariably come with unexpected cost expenditures in terms of staff time and training. Mobley provides an invaluable case study that can greatly impact a major IR decision both in terms of creating an IR from scratch or switching from a proprietary to an open source platform.

Purdue University

David Scherer, Lisa Zilinski, and Kelley Kimm's case study, "Interlinking Institutional Repository Content and Enhancing User Experiences" (Chapter 19), focuses on the connection and linkage of published research findings available in Purdue's textual-institutional repository, Purdue e-Pubs, to published datasets available in the Purdue University Research

Repository (PURR). They discuss a partnership with the Joint Transportation Research Program (JTRP) to develop these two repositories to further enhance two intersecting publishing workflows to account for enhancements and presentation of content, and to further develop the user's experience with an overall goal of increasing access and visibility of published technical report publications and published datasets.

Utah State University

Betty Rozum and Becky Thoms describe a strategy of populating an institutional repository through relying on subject librarians and cultivating grassroots efforts in Chapter 20, "Populating Your Institutional Repository and Promoting Your Students: IRs and Undergraduate Research." In coordination with its subject librarian, the Physics Department at Utah State University recognized the opportunity of the IR, DigitalCommons@USU, to showcase the department by combining student and faculty research and organizing it by research area. As a result, many student and faculty works that might not ordinarily receive a great deal of attention, such as posters and conference proceedings, have been discovered and utilized by scholars inside and outside of the Utah State system. Utah State's story demonstrates the potential of IRs for all schools.

PART 6: CLOSING REFLECTIONS AND THE NEXT STEPS FOR INSTITUTIONAL REPOSITORIES

The main purpose of Part 6 is to provide a better understanding of the priorities and challenges institutional repositories will face in the coming years by highlighting the broader factors that will most likely affect the development of repositories, repository services, and the roles of those directly involved including scholarly communications librarians, repository managers, and the library administrators in charge of making resource decisions.

CONCLUSION

The number of institutional repositories established and the total amount of content they hold has exploded in recent years. While institutional repositories are entering their second decade with rapid growth, they are still in their infancy and have yet to reach their fullest potential. We hope that this volume offers a bird's-eye view of the scholarly communication landscape

and a clear picture of where IRs have been, where they are today, and where they will be in the future.

With this book, we hope that you will find one source that will allow you to gain a fuller grasp of the concept of institutional repositories as well as introduce you to strategies that have worked to make IRs relevant, useful, and vital at institutions nationally as well as internationally. We hope that those looking to launch a repository will find this volume helpful and that those of us who have been furiously working to cultivate thriving repositories will find new ideas and models for collaboration, innovation, and success within the following pages.

REFERENCES

Crow, R. (2002). The case for institutional repositories: A SPARC position paper. *ARL Bimonthly Report,* 223. Retrieved from http://www.sparc.arl.org/resources/papers-guides/the-case-for-institutional-repositories

Ginsparg, P. (1997). Winners and losers in the global research village. *Serials Librarian, 30,* 83–96.

Salo, D. (2008). Innkeeper at the roach motel. *Library Trends, 57*(2), 98–123. http://dx.doi.org/10.1353/lib.0.0031