11-2023

Evaluating an institutional repository: A case study of Singapore Management University

Pin Pin Yeo

Evaluating an Institutional Repository
A Case Study of Singapore Management University

Yeo Pin Pin
Singapore Management University
ppyeo@smu.edu.sg

ABSTRACT

An institutional repository (IR) was launched in 2010 for a university in Singapore. The motivations for setting up an IR were to disseminate the research output through open access, to promote the university and researchers, and to help increase their research impact. Some quantitative and qualitative performance indicators were selected and used to measure the success of the IR. For the review and the performance indicators, the data used came from the IR at Singapore Management University (SMU), interviews with faculty members, and a study of a set of journal articles available in the IR to test a hypothesis. The results of the evaluation of the IR using various indicators were presented, and the factors that were successful in building the IR were having a workflow process between the research information system and the repository, having effective promotional activities, leveraging on institutional policies and research funder requirements to obtain full text, and proactively searching for full text. Some of the success factors and performance indicators might be of interest to other institutions to consider when implementing and evaluating an IR.

Keywords: Institutional repositories, academic libraries, evaluation, performance indicators, strategies, content recruitment

INTRODUCTION

An institutional repository (IR) was launched in 2010 at Singapore Management University (SMU). SMU was set up in 2000 focussing on Management, Economics, and Business studies. Computer and information systems were then added in 2002, and followed by social sciences and law that were added in 2007.

The objectives of the IR which is named Institutional Knowledge (InK) @ SMU are:
- To create visibility and raise the profile of SMU research and scholarly assets;
- To provide organisation of the records of SMU research and scholarly assets; and
- To provide access to the full text where possible of SMU research and scholarly assets.

The IR is hosted on the Digital Commons' platform which offers a "Software as a Solution" (SaaS) and storage of the records and full text on the cloud. The IR started with over 8,000 records and 600 (7.5%) full text in 2010. It has grown to 32,000 items and 16,000 full text at the end of 2022 with a cumulative total of 6.7 million downloads. Several indicators were selected to review the IR and the results shown.

LITERATURE REVIEW

Gibbons (2004) defined an institutional repository (IR) as having the following core features: digital, community-driven and focussed, institutionally supported, durable and permanent, and accessible. Lynch (2003) described an IR as "a set of services that a university

offers to the members of its community for the management and dissemination of digital materials created by the institution and its community members."

The motivations for SMU to establish an IR were similar to the findings of Baughman, Roebuck, & Arlitsch (2018) which were: collect, organise, and preserve digital versions of the institution's scholarship, provide increased visibility to institution's scholarship, and demonstrate commitment to open access principles.

Giesecke (2011) described the challenges faced by IRs and offered some models for an IR to offer a successful set of services. Mark and Shearer (2006) reviewed the content recruitment strategies of IRs. Armstrong (2014) discussed the mediated approach to populate the IR and the need to focus on supporting faculty in the dissemination of their scholarship. Marsh (2015) reviewed best practices in managing institutional repositories and identified some ways of addressing some of the perceived barriers to populating repositories and the benefits for stakeholders. Hwang, Elkins, Hanson, Shotwell, and Thompson (2020) found that most active promotion efforts were geared towards faculty and less marketing efforts were made for end-users. Bell, Fried Foster, and Gibbons (2005) explored why faculty members did not take full advantage of IRs and proposed strategies for outreach by liaison librarians.

Thomas (2007) offered a framework for performance indicators to evaluate IRs for researchers, institutional, and national perspectives. Westell (2006) proposed indicators to measure Canadian IRs with a focus on integration with research initiatives. Bruns and Inefuku (2015) highlighted some metrics and linked the metrics to purpose and audience. Serrano-Vicente, Melero, and Abadal (2018) used a set of indicators to study how Spanish repositories were run. Demetres and Delgado (2020) carried out a systematic review on the impact of IRs and found that IRs appear to have a positive impact on citation count, exposure or presence, and administrative burden. Bashir, Gul, Bashir, Nisa, and Ganaie (2022) found that repositories enhanced the preservation of institutional research output with increased viewership and prestige apart from achieving a potential research impact. Baro and Nwabueze-Echedom (2023) identified the challenges for repository development in African universities included collecting material for IRs, a lack of skilled information and communications technology personnel, copyright issues, and an absence of IR policies.

METHODOLOGY

The experiences and data from running the SMU IR were used to review the strategies used for populating the IR, increasing the full text in the IR, and promoting the IR. Some quantitative and qualitative performance indicators were selected and used to measure the success of the IR at SMU. Selected findings from interviews with 20 faculty members were also used. A study of a set of journal articles available in the IR to test the hypothesis that downloads in the IR were positively correlated to Field-Weighted Citation Impact (FWCI) was included.

FINDINGS

Westell (2006) proposed input indicators to measure the success of IRs that included mandate, integration into institutional planning, interoperability, and promotion. Thomas (2007) proposed performance indicators in a framework that would assist policy and decision-makers to evaluate IRs as required by the individual organisation. Serrano-Vicente, Melero, and Abadal (2018) grouped the indicators into five dimensions comprising technology, procedures, content, marketing, and personnel. Bruns and Inefuku (2016) related the indicators to the audience and purpose and found that metrics were a basic tool for demonstrating the

value of repositories to stakeholders. Some relevant indicators would be selected for evaluating the IR at SMU. The qualitive indicators considered were:

Mandate: Westell (2006) thought that a clearly defined mandate was key. Although there is no mandate at SMU for open access, an Open Access Policy was put in place in 2013 which set out the faculty/student role, the library's role, and what would be collected in the repository. The policy was brought before the Provost Council and the Research Council which included the Deans and Vice-Deans for Research of the schools and key administrative heads in SMU. The policy was approved and had the support of the key stakeholders in SMU.

Interoperability: Westell (2006) thought interoperability indicated the institution's openness to contribute to international scholarship. The IR at SMU is compliant with the Open Archives Initiative-Protocol for Metadata Harvesting (OAI-PMH). This compliance is important as it ensured that the IR can be harvested by other services like CORE and Unpaywall as well as Internet search engines. The IR is also registered with the ROAR (Registry of Open Access Repositories) and OpenDOAR.

Integration into institutional planning: Westell (2006) stressed the importance of the goals of the repository to be aligned with those of the institution. In *SMU Vision 2025*, one of the strategies is to address societal challenges through cutting edge research and to conduct leading research with direct societal impact. The goals of the IR at SMU are aligned to the strategy for research.

Availability of public-funded research: Thomas (2007) also listed the "extent to which public-funded research available and accessible" (p. 144) as an important performance indicator for the research office, the university and for national research and development. A field for indicating the research funder can be increasingly found in various repositories and even databases like Scopus and Web of Science, which enabled Kirkman and Haddow (2021) to study the level of compliance with an Australian research funder mandate.
In February 2016, the major research funders in Singapore added a clause in their requirements:

> *Institutions shall ensure that all publications arising from publicly funded Research are made openly available no later than twelve (12) months after the official date of publication.*

The SMU Office of Research (ORE) contacted the library to check if these public-funded publications were available in the IR. It was found that the IR had some of the full text, but it was not complete. The library took the opportunity to proactively upload the full text for those that were not in the IR. The full text were sourced from asking the faculty members and from searching the Internet. Further rounds of checks were done from 2019 to 2022 to check if the public-funded publications were in the IR. Hence, there was high confidence that the publications from public-funded research at SMU were available in the IR.
In March 2020, the National Research Foundation (NRF) required all principal investigators submitting annual progress reports to provide NRF with a web-link to the deposited publication or author-accepted manuscripts in a publicly accessible repository. This requirement from the funder further incentivised faculty members who had grants to upload

their papers to the RIS and the IR at SMU. Having requirements from the research funder for open access of publications made the job of asking for full text easier and in fact some faculty members contacted the library to request that their papers be uploaded to the IR so that they could have the web-link to add to their progress report for the research funder.

Figure 1: Sample of progress report for NRF for web-link to publication in the repository

Progress Report	: NRFPR-0000000XX
Award ID	: NRF-CRPXX-201X-0X
Project ID	: NRF-CRPXX-201X-0X
Title of Research Project	: XXXXXX
Reporting Period (FY / Mid-Term, Completion Review, Final, etc.)	: FY201X / Mid-Term / etc.

Journal Publications (Top 10%)				
Title of Publication	**Name of Journal**	**Published Date**	**Deposited Date**	**Web-link to Publication or Author-Accepted Manuscript in openly accessible repository** (please provide reasons if a web-link is not available)
XXXX	XXXX	dd/mm/yyyy	dd/mm/yyyy	XXXX

Source: SMU Intranet

Fit between IR and RIS: The fit between IR workflow processes with research management information system was a performance indicator that Thomas (2007) listed. At SMU, the research information system (RIS) is linked to the IR and there is a workflow to pull records from the RIS to the IR. This workflow has been a key factor in the success of the IR, as faculty members were motivated to update the RIS for the purpose of appraisal. Leveraging on the addition of records in the RIS was more effective than asking faculty members to self-submit to the IR. The library has a role to play in the RIS and the process between the RIS and the IR. The library is the business owner of the publication module in the RIS, the library staff edits the records to ensure high quality metadata and exports the records and full text to the IR from the RIS.

Reports: The qualitative performance indicators proposed for the library by Thomas (2007) included monthly and annual reports. For SMU, selected metrics for the IR are reported in the Library's annual report and in the Research Librarians' annual reports to the schools. Key milestones in downloads are celebrated via emailers to the community, with announcements in the State of the University address by the President and in the SMU annual report. The platform also generates monthly emails to the faculty with the number of downloads and a link to their personal dashboard. From the interviews done in 2019, the dashboard and download counts were useful to most of the faculty members.

Promotion: Westell (2006), Serrano-Vicente, Melero, and Abadal (2018) considered promotion a crucial factor. Westell (2016, p. 217) gave the example of Cornell University Library which "introduced an incentive plan … where faculty compete for grants which are used to develop and add content to the repository." At SMU, the Library carried out a range of promotion activities for the IR, ranging from events, information literacy sessions, emailers, using the data from the IR for "shout outs" for top downloaded authors and papers, web pages with information about the IR and its benefits, and most importantly, engagement with the

faculty members through the Research Librarians, with some combination of activities done throughout the year.

The quantitative indicators considered were:

Number of items: Thomas (2007), Bruns, and Inefuku (2016) found that a widely used method to measure the 'success' of the IR was to count the gross number of items in the IR. Some criticisms were levelled at looking at the total item counts, and Carr and Brody (2007) suggested looking at daily ingestion rates to attempt to measure engagement with the community. At SMU, the workflow is based on a weekly ingestion of records from the RIS to the IR. *Figure 2* showed the cumulative number of items in the IR with steady growth in the number of items. *Figure 3* showed the ingestion activity for a typical year, with the highest peak in January, as all faculty update their curriculum vitae (CV) and their publications in the RIS in preparation for their appraisals that are usually carried out in January. Serrano-Vicente, Melero, and Ernest (2018) suggested the volume "must be related to the institution's total volume of research and the policies adopted". Hence, one possible way to benchmark this indicator would be to take the ratio of the number of ingested records in the IR to the number of publications by the institution, with a ratio of one as the ideal case.

**Figure 2:** Number of items in the IR (cumulative) 2010-2022
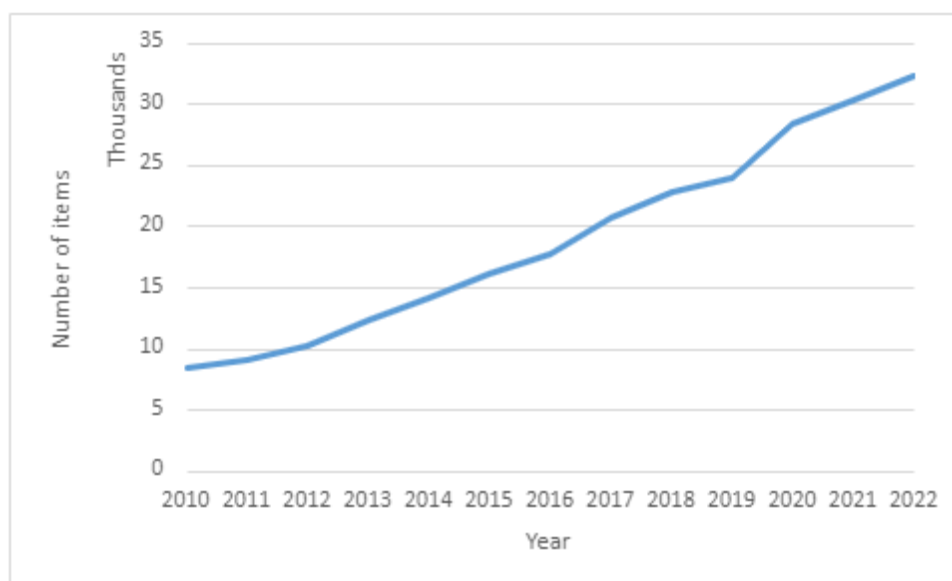
Figure 3: Number of items uploaded to the IR for the period January 2022 to December 2022
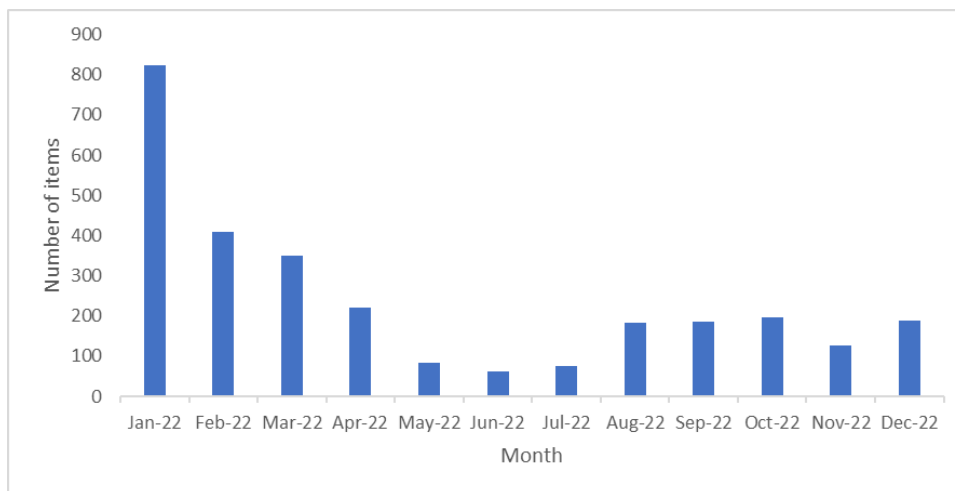
Percentage of full text: Bruns and Inefuku (2016) suggested that the percentage of items in the repository with full text would be useful in assessing the success of the library's engagement in scholarly communication and open access discussion on campus (p. 217-218). It would be difficult to benchmark with other repositories due to the difference in policies, for example a repository that only adds a record for an item that they have the full text, their percentage of full text would always be 100%.

At SMU, our policy is to add records of all faculty publications regardless of whether there is a full text version that can be uploaded to the IR. The overall percentage of full text in the IR had grown from 7% in 2010 to 51% in 2022 (see *Figure 4*) which showed that the library was effectively engaging faculty and to encourage them to add full text to the IR. The aim would be to further increase the percentage of full text, with the acknowledgement that there would be some full text that cannot be uploaded due to copyright and type of publication like books and book chapters.

Figure 4: *Growth in percentage of full text items in the IR for the period 2010 to 2022*
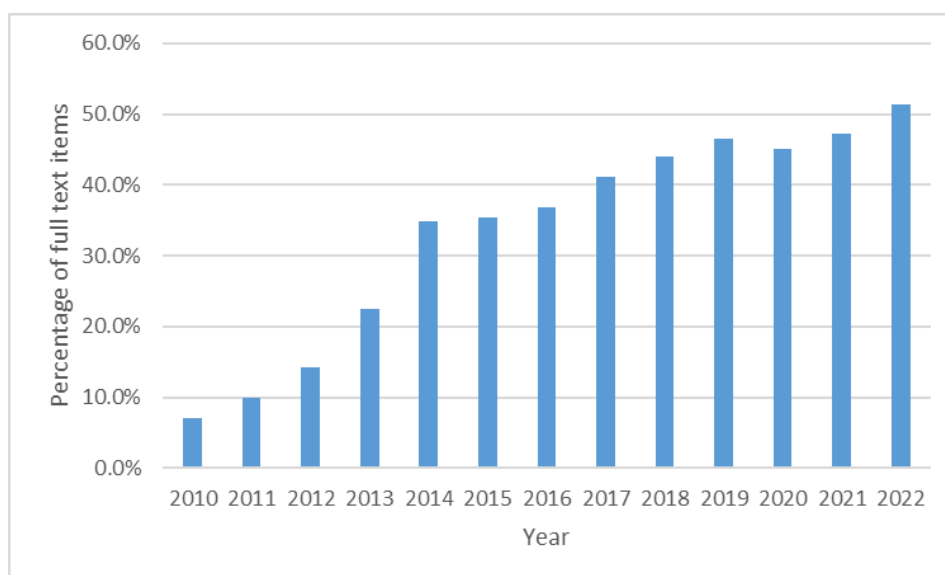
Downloads: Baughman, Roebuck, and Arlitsch (2018), Bruns, and Inefuku (2016), and Thomas (2007) found that the number of downloads was another often-used metric to assess the IRs. Thomas (2007) included hits and downloads as one of the input performance indicators. At SMU, it took six years after the launch of the IR to accumulate one million downloads in 2016, the next million downloads took two years and subsequently it took one year to accumulate the next million downloads as seen in *Figure 5*. As suggested by Bruns and Inefuku (2016), this metric may have contributed to faculty continuing to deposit new material.

Figure 5: *Growth in cumulative downloads in the IR for the period 2011-2022*
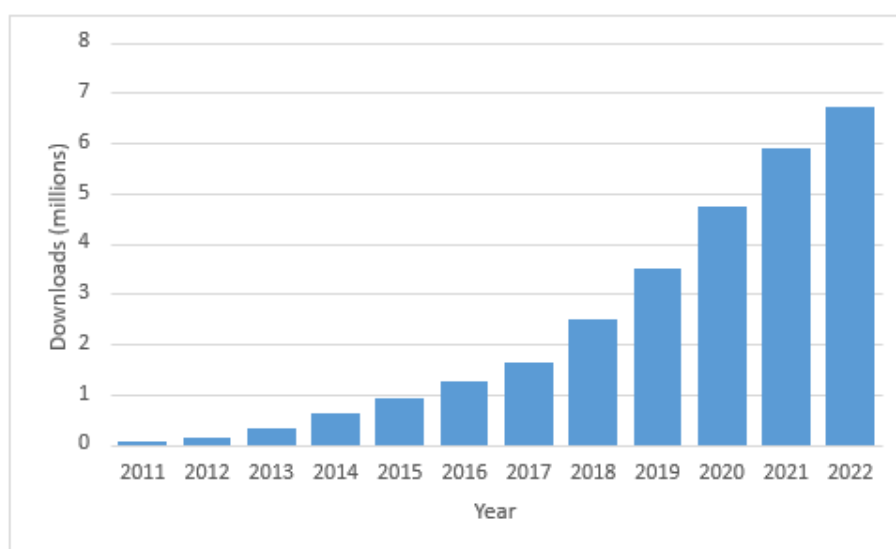
Visitor locations: Bruns and Inefuku (2016) proposed using visitor locations to demonstrate the visibility and reach of scholarship. At SMU, 62% of visitors who downloaded the full text were from the education sector and 31% were from the commercial sector and 2% from the government sector, across 225 countries, with the United States, the United Kingdom, India, China, Germany, Philippines, and Singapore among the top 10 countries. The data showed that SMU research was not just read by the education sector but also having an impact in the commercial sector locally in Singapore and the world.

Rate of deposit: Thomas (2007) also proposed using the rate of deposit which Bruns and Inefuku (2016) called item uploads as an indicator. At SMU, the rate of deposit is mediated by the Library and pulled from the research information system. The typical pattern for the rate of deposit can be seen in *Figure 2* with records added to the IR on a weekly basis.

Range of items: Thomas (2007) proposed using the range of items in the IR as an indicator to measure the exposure and the preservation of institutional research. At SMU, the range of items in the IR started with only the publications by the faculty members in the form of journal articles, conference papers, and book chapters, and has expanded to include working papers and reports published not just by the schools, but also those published by the centres and institutes in SMU. The IR also includes the records for teaching cases and the full text of magazines produced by SMU. The IR started with theses and dissertations by

postgraduate students and has grown to include other publications by postgraduate and undergraduate students like journal articles and conference papers.

Satisfaction levels: For the researcher as an end-user of the IR, Thomas (2007) proposed using satisfaction levels as an indicator. A study was conducted in 2019 with 20 faculty members representing all the six schools in SMU. They were interviewed about their perception of the IR. All the faculty interviewed appreciated that all their works was available in one place and made discoverable. They gave high importance to their papers being discoverable in Google and Google Scholar. They usually looked at the download numbers shown in the monthly email sent by the IR platform, but only 30% clicked on the link in the email to view the dashboard and they looked mostly at the world map showing the readership distribution. The faculty members were not asked to rate their satisfaction with the IR on a scale, the positive feedback collected from the interviews indicated high satisfaction.

Faculty A reported that having her publications in the IR led to requests for collaboration from the industry and the requestor specifically mentioned finding her papers in the IR. Faculty A said: "I am quite excited to know that my work is read by industry people and relevant to current industry practice."

Faculty B said "I really appreciate the service provided by the IR. I have seen my citations rise after my papers were available in the IR."

Faculty C reported that he received invitations to speak at universities about the topics in the papers that were made open access on the IR.

Citation rate: Bruns and Inefuku (2016) proposed using number of citations as to demonstrate scholarly impact for the author. As seen in *Figure 6*, the number of citations for SMU publications in Scopus had grown steadily since 2011, but was there a correlation to the growth in downloads in the IR? Faculty members had asked about the implications of high downloads for a paper in the IR, and whether it will lead to more citations.

Eysenbach (2006), Piwowar, et al (2018), and Langham-Putrow, Bakker, and Riegelman (2021) studied the open access citation advantage, and Langham-Putrom et al (2021) found that the quality and heterogeneity of the studies posed challenges for generalisation. To verify whether the downloads were correlated to citations for papers available in the IR at SMU, a study was conducted in 2019. A set of journal articles with SMU affiliation published between 2009 and 2018 was extracted from Scopus and SciVal including the citation count and the Field-Weighted Citation Impact (FWCI). The number of downloads for the full text in the IR were extracted for the same set of articles. The hypothesis that downloads in the IR were positively correlated to FWCI was accepted where N = 1,091 and $p < 0.001$ as shown in *Figure 7*. The study showed that the IR had a positive impact on the citation rate which was in line with the findings of Demetres and Delgado (2020).

Figure 6: Growth in number of citations in Scopus for SMU affiliated publications (2011-2022)
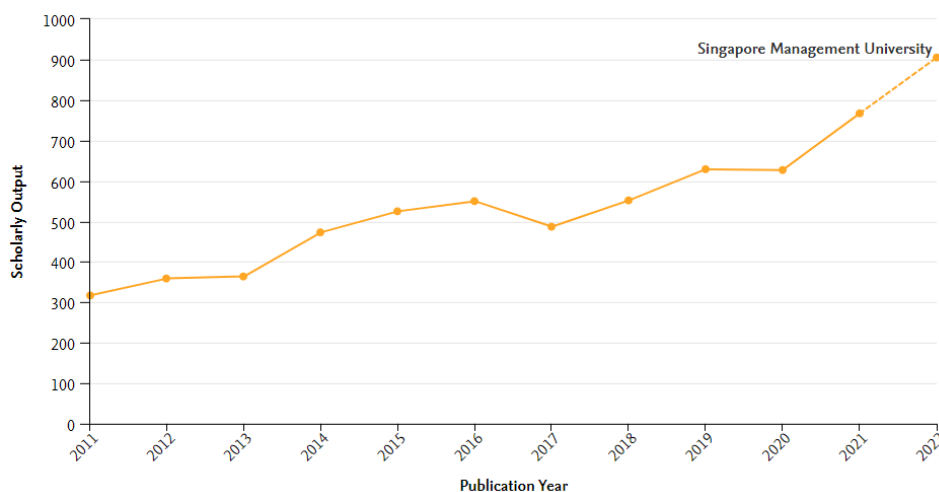
Figure 7: Correlation matrix for FWCI and Downloads in InK (*N* = 1091)

Correlation Table

			Spearman		Kendall	
			rho	p	tau B	p
Field-Weighted Citation Impact	-	Downloads in InK	0.386***	< .001	0.265***	< .001

* p < .05, ** p < .01, *** p < .001

CONCLUSION

The evaluation showed that the IR at SMU has met its objectives. For the objective: To create visibility and raise the profile of SMU research and scholarly assets, the records are discoverable in search engines, the downloads were high, and positively correlated to citation impact. For the objective: To provide access to the full text where possible of SMU research and scholarly assets, the number of full text in the IR has grown to over 16,000 with 51% of the items in the IR having full text. For the objective: To provide organisation of the records of SMU research and scholarly assets, the library is the business owner of the Research Publications module in IRIS. There is a workflow process where the library staff checks and edits the records and uploads them to the IR. This gives the library a critical role in ensuring good metadata for the RIS and the IR and to grow the IR in a sustainable way.

Bruns and Inefuku (2016) noted that an important criteria of performance indicators would be to connect the repository to the heart of the research community of the institution and to match the institution's strategic goals. One of the strategies in *SMU Vision 2025* is to address societal challenges through cutting edge research and to conduct leading research with direct societal impact. From the indicators relating to research, which were: integration into institutional planning**,** the fit between IR and RIS, availability of public funded research, and citation rate, the performance of the IR at SMU, it showed that the IR was aligned to the SMU strategy for research. This would be continually monitored and to assess if adjustments would be needed in the future to ensure alignment with the institutional strategic goals.

The first strategy that worked well for SMU to build and to sustain the IR was having a workflow process to harvest publication records and full text from the research information

system. This ensured a regular influx of records without depending on self-deposit by faculty. The next strategy that contributed to the success of the IR was the proactive engagement with faculty and the search for full text to add to the IR. The projects implemented by the library included actively contacting faculty members for full text, searching the Internet for a suitable version, and tracking publisher policies for changes like when the published version was allowed in IRs. These projects helped to increase the number of full text in the IR. Having an Open Access Policy provided opportunities to raise awareness of open access initiatives and to start conversations with the stakeholders and to encourage faculty to submit their full text. Changes to research funding requirements were leveraged on and the library played a role in supporting the faculty members and the institution to fulfil funder requirements. Having good search discoverability of the IR was also particularly important to the faculty, and for end-users who often started their search in their favourite search engines and discover the papers in the IR from their search results.

The lessons above were gathered from the evaluation done, and it is hoped they would be useful to other libraries starting a repository or planning to evaluate their repositories.

Future research could focus on proposing a commonly agreed set of indicators for repositories and a comparative analysis of these indicators on the repositories in Singapore.

REFERENCES

Arlitsch, K., O'Brien, P., Mixter, J. K., Clark, J. A., & Sterman, L. (2016). Ensuring discoverability of IR content. In B. B. Callicott, et al (Eds.), Making institutional repositories work. West Lafayette, IN: Purdue University Press.

Armstrong, M. (2014). Institutional repository management models that support faculty research dissemination. *OCLC Systems & Services, 30* (1), 43-51. DOI: 10.1108/OCLC-07-2013-0028

Baro, E. E., & Nwabueze-Echedom, A. U. (2023). An evaluation of institutional repository development in African universities. *IFLA Journal 49* (1), 18-38. DOI: 10.1177/03400352221089

Bashir, S., Gul, S., Bashir, S., Nisa, N.T., & Ganaie, S.A. (2022). Evolution of institutional repositories: Managing institutional research output to remove the gap of academic elitism. J*ournal of Librarianship and Information Science 54* (3), 518-531. DOI: 10.1177/09610006211009592

Baughman, S., Roebuck, G., & Arlitsch, K. (2018). Reporting practices of institutional repositories: Analysis of responses from two Surveys. *Journal of Library Administration, 58* (1), 65-80. DOI: 10.1080/01930826.2017.1399705

Bell, S., Fried Foster, N., & Gibbons, S. (2005), Reference librarians and the success of institutional repositories. *Reference Services Review, 33* (3), 283-290. DOI: 10.1108/00907320510611311

Bruns, T., & Inefuku, H. W. (2016). Purposeful metrics: Matching institutional repository metrics to purpose and audience." In In B. B. Callicott, et al (Eds.), *Making institutional repositories work*. West Lafayette, IN: Purdue University Press.

Carr, L., & Broady, T. (2007). Size isn't everything: Sustainable repositories as evidenced by sustainable deposit profiles. *D-Lib Magazine, 13* (7), 1-19. DOI: 10.1045/july2007-carr

Demetres, M. R., Delgado, D., & Wright, D. N. (2020). The impact of institutional repositories: A systematic review. *Journal of the Medical Library Association, 108* (2), 177-184. DOI: 10.5195/jmla.2020.856

Eysenbach, G (2006). Citation advantage of open access articles. *PLoS Biology, 4* (5), 692-698. DOI: 10.1371/journal.pbio.0040157

Gibbons, S. (2004). Defining an institutional repository. *Library Technology Reports*. American Library Association, pp. 6-10. https://journals.ala.org/index.php/ltr/article/view/437

Giesecke, J. (2011) Institutional repositories: Keys to success. *Journal of Library Administration, 51*, (5-6), 529-542. DOI: 10.1080/01930826.2011.589340

Hwang, S. Y., Elkins, S., Hanson, M., Shotwell, T., & Thompson, M. (2020). Institutional repository promotion: Current practices and opinions in Texas academia. *New Review of Academic Librarianship, 26* (1), 133-150, DOI: 10.1080/13614533.2019.1587483

Kirkman, N., & Haddow, G. (2021). Publisher embargoes and institutional repositories: A case study of journal articles subject to an Australian funder mandate. Information Research 26 (3). DOI: 10.47989/irpaper907

Langham-Putrow, A., Bakker, C., & Riegelman, A. (2021). Is the open access citation advantage real? A systematic review of the citation of open access and subscription-based articles. *PLoS ONE, 16*, e0253129. DOI: 10.1371/journal.pone.0253129

Lynch, C. (2003). Institutional repositories: Essential infrastructure for scholarship in the digital age. *Portal: Libraries and the Academy 3* (2), 327-336. DOI: 10.1353/pla.2003.0039

Mark, T., & Shearer, K. (2006). Institutional repositories: A review of content recruitment strategies. *World Library and Information Congress*, August 20-24, Seoul. DOI: 10.11575/PRISM/29080

Marsh, R. M. (2015). The role of institutional repositories in developing the communication of scholarly research. *OCLC Systems & Services 31* (4), 163-195. DOI: 10.1108/OCLC-04-2014-0022

Piwowar, H., Priem, J., Larivière, V., Alperin, J. P., Matthias, L., Norlander, B., Ashley, F., Jevin, W. & Haustein, S. (2018). The state of OA: A large-scale analysis of the prevalence and impact of Open Access articles. *PeerJ, 2018* (2), e4375. DOI: 10.7717/peerj.4375

Serrano-Vicente, Rocio, Melero, Remedios, & Abadal, Ernest (2018). Evaluation of Spanish institutional repositories based on criteria related to technology, procedures, content, marketing and personnel, *Data Technologies and Applications, 52* (3), 384-404. DOI: 10.1108/DTA-10-2017-0074

Swan, A. (2008). The business of digital repositories. *In* K. Weenink, L. Waaijers & K. van Godtsenhoven (eds.) *A DRIVER's Guide to European repositories.* Amsterdam: Amsterdam University Press.

Thomas, G. (2007). Evaluating the impact of the institutional repository, or positioning innovation between a rock and a hard place. *New Review of Information Networking, 13* (2), 133-146. DOI:10.1080/13614570802105992

Westell, M. (2006). Institutional repositories: Proposed indicators of success. *Library Hi Tech 24* (2), 211-226. DOI: 10.1108/07378830610669583

About the author:

YEO Pin Pin is the Head of Research Services at the Singapore Management University (SMU) Libraries. She oversees the Research Publications in the SMU research management system, institutional repository, data repository, and oral history. She is also involved in measuring the citation impact of SMU publications.