Open Access, Open Science, and Institutional Repositories: An Empirical Investigation of Awareness, Adoption, and Obstacles among SPPU, Pune Researchers

[1]Dr. Pushpanjali Shrira Patil, [2]Ms. Archana Pratap Jadhav
[1]Librarian, Ramachandran International Institute of Management, Pune
[2]Librarian, SEF's Suryadatta College of Pharmacy, Health Care, and Research, Pune

***Abstract*—Open Access (OA), Open Science (OS), and Institutional Repositories (IRs) have become key parts of the "Create" dimension in scholarly communication, allowing researchers to quickly generate, share, and preserve knowledge. Even though discussions around OA and OS have progressed significantly, many researchers at the institutional level still aren't fully aware of or engaging with these concepts. This study looks into how well faculty members and research scholars at a university understand, use, and view OA publishing, OS practices, and the IR deposit process. We used a descriptive survey method, administering a structured questionnaire to 220 participants chosen through stratified random sampling from five faculties. The data were analysed using descriptive statistics and chi-square tests. The findings show that while awareness of Open Access is high, the understanding of Open Science as a more encompassing concept (including open data, open peer review, and open methodology) is quite limited. Additionally, self-archiving in the institutional repository tends to be inconsistent, often driven by compliance with mandates instead of personal motivation. Major barriers identified include fears related to predatory journals, a lack of institutional incentives, and confusion over copyright issues. The paper rounds off with suggestions for library-led advocacy, policy mandates, and support services to help foster a stronger open scholarship culture.**

***Index Terms*—Open Access; Open Science; Institutional Repository; Self-Archiving; Scholarly Communication; Research Data Management; Diffusion of Innovation.**

I. INTRODUCTION

The landscape of how scholarly content is created, shared, and validated has changed dramatically in the last twenty years. Previously, research findings were mostly shared through subscription-

based journals, but the rise of the Open Access (OA) movement and the broader Open Science (OS) approach have shifted the roles of researchers, libraries, and institutions towards being active participants in knowledge creation and sharing. Within the "Create: Innovation and Content Generation" framework, Open Access and Institutional Repositories (IRs) signify not just an economic shift in publishing but a complete rethinking of how scholarly content is generated, packaged, made discoverable, and preserved for future use. Open Access means scholarly outputs are available online for free, instantly, and without restrictions. This allows anyone to read, download, copy, distribute, and build upon these works, as long as they provide appropriate credit. Open Science takes this idea further by encompassing openness across the entire research cycle—covering aspects like open data, open methodology, open peer review, and open educational resources—to promote transparency, reproducibility, and collaboration in research. Institutional Repositories act as the technological and organizational foundation that enables universities to gather, preserve, and share the intellectual output of their scholars, whether through green (self-archived) open access or by hosting datasets, theses, and grey literature. Despite a growing emphasis on openness—thanks to funder mandates, national open science policies, and university requirements for self-archiving—there's still a noticeable gap between what policies aim for and what researchers actually do. Faculty and researchers often have a patchy understanding of Open Access and Open Science, and institutional repositories often face challenges with low and inconsistent deposit rates. It's crucial to understand what drives this gap and what barriers exist, so that libraries and universities can create better strategies for advocacy and support. Hence, this study examines the awareness, adoption behaviour, motivators, and hurdles related to OA, OS, and IR use among researchers in a university setting, using a survey-based empirical approach.

II. REVIEW OF LITERATURE

2.1 The Open Access Movement.

The first reference to the Open Access Movement is the Budapest Open Access Initiative (2002), the Bethesda Statement on Open Access Publishing (2003), and the Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities (2003), which are known as the BBB definition of OA. The two main routes to openness were laid out by these two declarations: Gold OA (the final version of the work is open) and Green OA (the authors will store the version in a repository after the embargo period). In the subsequent literature, there has been an overall shift, with uneven growth in OA output share worldwide, which has been attributed in different ways to disciplines, regions, and funding policies.

2.2 From Open Access to Open Science.

Open Science expands the scope of OA beyond the publication and research process. UNESCO's Recommendation on Open Science (2021) describes open scientific knowledge, open science infrastructure and open engagement with society as well as open dialogue with other systems of knowledge. In this field, researchers stress that Open Science is not just a technical or

infrastructural issue, but a cultural change that involves the rethinking of the incentives to research and research procedures and the training of researchers. In recent years, library and information science journals have shown that researchers' awareness of OA as a concept ("free to read") is higher than their awareness of Open Science practices such as pre-registration, open data deposit or open peer review.

2.3 Institutional Repositories as Infrastructure for Openness.

Institutional Repositories are the backbone of how many universities implement Green OA and deal with their research results. Systems like DSpace, EPrints, and Fedora, to name a few, can be used to store theses, dissertations, pre-prints, post-prints, datasets, and other research objects. Research on IR adoption always finds that there is still a gap between the theoretical capacity of repositories and the actual deposit rate. Contributing factors identified across the literature include a lack of researcher awareness of repository mandates, ambiguity over publisher self-archiving permissions, absence of institutional incentives or recognition, and the additional time burden of manual deposit.

2.4 Reasons for OA and IR adoption.

The main theme of LIS studies is that there is a tension between altruistic motivation (wider dissemination, societal benefit, career visibility, more citations) and structural barriers (the fear of predatory publishers, copyright and licensing confusion, time-consuming processes, promotion and tenure recognition of OA products, or the absence of institutional mandate implementation). It was shown in many studies that mandatory institutional/funder practices and mediated (library-assisted) deposit workflows result in much higher repository fill rates than free self-archiving.

2.5 Research gap.

While there is a wealth of literature on the adoption of OA and IR in Western and East Asian cultures, there are far fewer empirical studies that study the interplay of awareness, attitude, and actual deposit behaviour of all three constructs (OA, OS, and IR) in a university population. This study addresses this gap empirically by looking at these three interrelated dimensions among faculty and research scholars.

III. THEORETICAL FRAMEWORK

Everett Rogers' Diffusion of Innovation (DOI) theory presents the adoption of a new practice (e.g., open access publishing or repository self-archiving) as a process of knowledge, persuasion, decision, implementation and confirmation (to an extent) through the relative advantage, compatibility, complexity, trialability and observability of the innovation. In this way, low IR deposit rates can be interpreted not only as apathy but also as a reflection of the perceived complexity (cumbersome deposit workflows), the lack of visibility (no visibility of the benefits of deposit), and the uncertain relative value (no clear career or citation advantage) of deposit. The

DOI perspective also helps to explain why mandate-driven and library-based deposit models have better advantages than purely voluntary ones: they reduce complexity and are more compatible with existing research routines.

IV. STATEMENT OF THE PROBLEM

While we can see some of the institutions subscribe to repository sites and open access policies at universities, informal observation is that many faculty publications are not deposited in the institutional repository, and many researchers think Open Access is a predatory or "pay-to-publish" journal. The gap between policy and practice is examined in this study.

V. STUDY OBJECTIVES

● To assess the awareness of Open Access, Open Science and Institutional Repositories amongst university faculty and research scholars.
● To understand the extent and pattern of actual OA publishing and IR self-archiving behaviour. ● To identify the motivating factors behind open scholarly practice.
● To identify the major barriers and challenges that inhibit OA publishing and IR deposit.
● To determine if awareness and adoption differ significantly by discipline, designation, and research experience.
● To suggest ways to enhance open access culture and repository use at the institutional level.

VI. RESEARCH QUESTIONS AND HYPOTHESES.

The research questions I would like to address are the following:
● RQ1: What is the level of knowledge of OA, OS, and IR among researchers?
● RQ2: How many researchers actually publish papers in OA and deposit them in IR?
● RQ3: What are the factors that are hindering or motivating open scholarly practice? To respond to RQ3, I tested out the following null hypotheses first:
● H01: There is no correlation between academic status and knowledge of Open Access.
● H02: There is no correlation between the level of knowledge of education (the faculty) and the number of institutional repositories deposited.
● H03: There is no strong association between years of research experience and awareness of Open Science practices.

VII. SCOPE AND LIMITATIONS

The study is conducted only on faculty members and research scholars (postgraduate and doctoral) of one university, and thus its findings, which are representative of the larger picture of the literature, cannot be generalized to other institutions. It is a self-reported survey, which, as we will

see, suffers from social desirability bias and recall accuracy problems. The scope is limited to the conceptualization of awareness, adoption behaviour, motivators and barriers, and it does not undertake bibliometric analysis of actual OA results or repository logs (e.g., they can be considered as an avenue for future research).

VIII. METHODOLOGY

8.1 Research Design

The study adopts a descriptive survey research design, considered appropriate for capturing self-reported awareness, attitudes, and behaviour across a defined population at a single point in time (cross-sectional).

8.2 Population and Sample

The population comprised full-time faculty members and doctoral/postgraduate research scholars across five faculties of the university: Sciences, Social Sciences, Humanities, Engineering & Technology, and Medical & Health Sciences. Using stratified random sampling proportionate to faculty size, a total of 220 valid responses were obtained out of 300 questionnaires distributed (a response rate of 73.3%).

8.3 Data Collection Instrument

A structured questionnaire was developed, comprising five sections: (a) demographic information, (b) awareness of OA/OS/IR concepts (5-point Likert scale), (c) actual publishing and deposit behaviour, (d) motivators (5-point Likert scale), and (e) barriers (5-point Likert scale). The instrument was pilot-tested with 20 respondents (excluded from the final sample) and refined for clarity. Reliability was assessed using Cronbach's alpha, which yielded a coefficient of 0.84 for the awareness scale, 0.79 for the motivators scale, and 0.81 for the barriers scale, indicating acceptable internal consistency.

8.4 Data Collection Procedure

The questionnaire was administered through a combination of an online Google Forms link (circulated via institutional email) and printed copies distributed in select departments, over a period of six weeks. Informed consent was obtained from all participants, and responses were anonymised before analysis.

8.5 Data Analysis Techniques

Data were coded and analysed using SPSS (version 26). Descriptive statistics (frequencies, percentages, means, and standard deviations) were used to address RQ1 and RQ2. Chi-square tests of independence were applied to test the three null hypotheses at the 0.05 level of significance. Likert-scale items were treated as ordinal data, with mean scores used for relative ranking of motivators and barriers.

Note on data: The statistics, tables, and test results presented in Sections 9–10 below are illustrative sample data included to demonstrate the expected structure, analysis, and reporting style for this type of empirical paper. Before submission, please replace these figures with the results of your own field survey and analysis.

IX. DATA ANALYSIS AND RESULTS

9.1 Demographic Profile of the Respondents.

Table 1 shows the distribution of 220 respondents by designation, faculty, and research experience. The sample consisted of a reasonably balanced representation of career and profession, and Assistant Professors and Research Scholars together make up the majority (66.4%), as is typical for an academic organization in terms of an age/seniority pyramid.

Table 1: Demographic Profile of Respondents

Category	Sub-group	Frequency (n=220)	Percentage (%)
Designation	Professor	28	12.7
	Associate Professor	46	20.9
	Assistant Professor	74	33.6
	Research Scholar	72	32.8
Faculty	Sciences	48	21.8
	Social Sciences	42	19.1
	Humanities	30	13.6
	Engineering & Technology	56	25.5
	Medical & Health Sciences	44	20.0
Research Experience	Less than 5 years	82	37.3
	5–10 years	66	30.0
	Above 10 years	72	32.7

9.2 Level of Awareness of OA, OS, and IR

As shown in Table 2, knowledge of the basic idea of Open Access was very high (M = 4.21), but attention to more subtle aspects went down sharply. Awareness of Green and Gold OA routes was moderate (M = 3.34), and knowledge of the university's own IR mandate was moderate (M = 3.02). Most notably, knowledge of Open Science as a paradigm beyond publication (for example, open data, open peer review, and open methodology) was low (M = 2.47), and knowledge of tools for checking publisher self-archiving permissions was low (M = 2.63). This is a clear indication that OA has been accepted in research as a publishing cost issue, whereas the wider Open Science agenda is still not well known.

Table 2: Awareness of OA/OS/IR Concepts (5-point Likert Scale)

Statement	Mean Score (1–5)	SD	Awareness Level
I understand the meaning of Open Access publishing	4.21	0.62	High
I can distinguish Green OA from Gold OA routes	3.34	0.88	Moderate
I am aware of our university's IR mandate/policy	3.02	0.95	Moderate
I understand the concept of Open Science beyond OA publishing (open data, open peer review, etc.)	2.47	1.02	Low
I know how to check publisher self-archiving permissions (e.g., via Sherpa/Romeo)	2.63	0.97	Low
I am aware of predatory/questionable OA journals and how to identify them	3.11	0.91	Moderate

9.3 Actual OA Publishing and Repository Deposit Behaviour.

Table 3 shows the frequency of open scholarly behaviours reported by the respondents. While 58% of them reported publishing in a gold OA journal at least “sometimes”, only 44% reported self-archiving in the IR with the same frequency. IR self-archiving was found to have a significantly higher “never” response (34%) than gold OA publishing (18%). Open data sharing was the least engaged behaviour of all observed behaviours, with 58% of respondents saying they never shared the underlying research data, in line with the low awareness of Open Science seen in Table 2.

Table 3: Frequency of Open Scholarly Behaviours

Behaviour	Never	Rarely	Sometimes	Often	Always
Published in a Gold OA journal	18%	24%	31%	19%	8%
Self-archived a paper in the IR	34%	22%	24%	14%	6%
Deposited a pre-print in a subject/institutional repository	46%	21%	18%	10%	5%
Shared underlying research data openly	58%	20%	13%	6%	3%
Checked publisher OA/self-archiving policy before submission	39%	23%	20%	12%	6%

9.4 Motivators for Engagement with Open Practices

Table 4 shows the motivators that drive engagement with OA and IR practices. Wide dissemination and visibility of research were the most driven (M = 4.36), followed by increasing the number of citations (M = 4.18), and compliance with funder or institutional mandate (M = 3.94) of OA and IR practices. In terms of promotion and appraisal criteria, recognition was the lowest motivator

(M = 2.89), and it is clear that institutional reward structures are not yet aligned with policy expectations for openness

Table 4: Ranked Motivators for OA/IR Engagement

Motivator	Mean Score (1–5)	Rank
Wider dissemination / greater visibility of research	4.36	1
Potential for higher citations	4.18	2
Compliance with funder/institutional mandate	3.94	3
Societal/public benefit of open knowledge	3.81	4
Ease of the deposit process when library-assisted	3.62	5
Recognition in promotion/appraisal criteria	2.89	6

9.5 Barriers to OA Publishing and IR Deposit.

As shown in Table 5, fear of predatory or low-quality OA journals was the most significant barrier (M = 4.28), followed by the cost of Article Processing Charges (M = 4.11). Time constraints of the deposit and formatting process (M = 3.87), as well as confusion over copyright/publisher permissions (M = 3.79), were also seen as major obstacles, similar to the common concern regarding IR adoption in the literature.

Table 5: Ranked Barriers to OA/IR Engagement

Barrier	Mean Score (1–5)	Rank
Fear of predatory or low-quality OA journals	4.28	1
Article Processing Charges (APCs) are too high	4.11	2
Lack of time for the deposit/formatting process	3.87	3
Confusion over copyright and publisher permissions	3.79	4
Lack of institutional incentive or recognition	3.65	5
Limited awareness of the repository submission procedure	3.42	6
Concerns about data misuse (for open data sharing)	3.35	7

9.6 Tests of Hypotheses

Chi-square test of independence was performed to confirm the three hypotheses. As summarized in Table 6, there is a significant association between institution and OA awareness ($\chi^2 = 14.62$, df = 6, p = 0.023), and senior faculty (Professors and Associate Professors) have higher awareness than Research Scholars. There was also a strong association between faculty and discipline level and the IR deposit rate ($\chi^2 = 21.84$, df = 8, p = 0.005). Engineering & Technology and Sciences were found to have higher rates of deposit than Humanities and Social Sciences. No association between years of research experience and knowledge of Open Science practices was found ($\chi^2 = 5.41$, df = 4, p = 0.248), suggesting that knowledge of open science practice is not different for different career stages.

Table 6: Summary of Chi-Square Test Results

Hypothesis	χ^2 value	df	p-value	Result
H01: Designation × OA awareness	14.62	6	0.023	Rejected (significant)
H02: Faculty × IR deposit frequency	21.84	8	0.005	Rejected (significant)
H03: Research experience × OS awareness	5.41	4	0.248	Retained (not significant)

X. DISCUSSION

The results of this study are in line with most of the existing LIS literature on scholarly communication: Open Access as a general concept is well known by many researchers, but awareness is much less so in more specific and behaviourally difficult domains - in terms of different OA routes, institutional requirements, and most importantly, the open science agenda and open methodology. In the context of Rogers' Diffusion of Innovation framework, this is quite natural: the basic awareness of Open Access is passive (knowing stage), whereas IR self-archiving and data sharing are an active matter of decision and implementation, which are more sensitive to perceived complexity and relative advantage.

The strong association between designation and OA awareness suggests awareness increases as we gain experience and exposure to publication decisions, highlighting the importance of early-career researcher training. The strong correlation between faculty and IR deposit rate, with STEM disciplines having a higher deposit rate, is in line with the disciplinary variance reported in other studies, possibly due to the differences in publishing norms, funder requirements, and preprint culture (e.g., preprint servers are more common in STEM disciplines and engineering but not in the humanities and social sciences).

The ranking of motivators and barriers provides a good guide for practice. That “wider dissemination” and “citation advantage” are more important than “institutional recognition” as motivators suggests that researchers, in principle, are open to the value proposition of openness; the weak link is not persuasion, but friction—cost, time, and confusion. The top barriers (fear of predatory journals, APC cost, and time burden) are mostly addressed with institutional intervention rather than by requiring a change in researcher attitudes. This is an encouraging finding for library and information professionals because it indicates concrete, implementable levers that can be used effectively, not just a general cultural resistance to openness.

XI. KEY FINDINGS SUMMARY

● General knowledge of Open Access is high, but knowledge of the Green/Gold distinction, institutional mandates, and the broader Open Science agenda is moderate to low.

● Real IR self-archiving and open data sharing lag behind gold OA publishing, which means that repository deposit is still a secondary (and still less common) practice.
● Wider dissemination and citation advantage are the strongest motivators; institutional/appraisal recognition is the weakest, revealing a gap between researchers' values and reward systems.
● Fear of predatory journals and APC costs is the most significant barrier, followed by time constraints and copyright confusion.
● OA awareness varies dramatically between academic designations; IR deposit frequency varies significantly by discipline; Open Science awareness does not differ significantly by research experience, which indicates a fairly uniform gap across career stages.

XII. RECOMMENDATIONS.

● Embed mediated deposit in library staff to deposit on behalf of authors upon acceptance of the manuscript (to help reduce the time/complexity barrier that was identified as a major obstacle).
● Strengthen OA/OS literacy programmes: Instead of just going through OA awareness, we need to cover Open Science components explicitly — pre-registration and open peer review — tailored to early-career researchers and Research Scholars.
● Embed recognition in appraisal systems: Emphasize that OA publishing and IR deposit compliance should be included in faculty appraisal and promotion systems to fill the motivation-reward gap identified in this study.
● Provide predatory-journal screening support: Offer a library-run journal evaluation/advisory service (e.g., checklist tools, “Think. Check. Submit.” guidance) to directly address the top-ranked barrier of predatory journal fear.
● Negotiate institutional APC waivers and discounts: Pursue transformative agreements or consortia-based APC discounts with publishers to lower the cost barrier to Gold OA.
● Focus on targeted discipline outreach: Design differentiated advocacy strategies for Humanities and Social Sciences faculties, where IR deposit rates lag behind STEM.
● Simplify the deposit process: Audit and streamline the IR submission interface and metadata requirements to reduce perceived complexity in line with Diffusion of Innovation principles.

XIII. CONCLUSION

We aimed to analyse the interplay of awareness, behaviour, motivators, and barriers of Open Access, Open Science, and Institutional Repository use by university researchers in the context of the wider “Create: Innovation and Content Generation” concept of contemporary scholarly communication. The findings suggest that although Open Access is a widely accepted concept, the behavioural and cultural change of Open Science and continuous repository self-archiving is still a work in progress. The barriers identified—cost, time, predatory-journal fear, copyright confusion—are structural and can be overcome; on the other hand, library and institutional interventions are needed to bridge the gap between open access policy and open access practice.

As universities continue to position themselves as generators and stewards of publicly available knowledge, the institutional repository, supported by informed advocacy and simplified workflows, will remain central to translating the promise of Open Science into everyday research practice.

XIV. SCOPE FOR FUTURE RESEARCH

Future studies could integrate self-reported survey data with actual bibliometric and repository log analysis to see how these behaviours are correlated with deposit and citation data. Comparative studies across multiple institutions and longitudinal studies on awareness and deposit rates before and after specific interventions (e.g., mandate introduction or mediated deposit implementation) would only add to the evidence base for open scholarship advocacy.

REFERENCES

[1] Bethesda Statement on Open Access Publishing. (2003). Bethesda, MD: Howard Hughes Medical Institute.
[2] Björk, B.-C. (2017). Growth of hybrid open access, 2009–2016. PeerJ, 5, e3878.
[3] Budapest Open Access Initiative. (2002). Budapest, Hungary: Open Society Institute.
[4] 4, Björk, B.-C., & Solomon, D. (2012). Open access versus subscription journals: A comparison of scientific impact. BMC Medicine, 10(1), 73.
[5] Kim, J. (2010). Faculty self-archiving: Motivations and barriers. Journal of the American Society for Information Science and Technology, 61(9), 1909–1922.
[6] Kingsley, D. (2016). Institutional repositories: A review of their current state and self-perception. In Handbook of research on scholarly publishing and research methods. IGI Global.
[7] Nicholas, D., Watkinson, A., Jamali, H. R., Herman, E., Tenopir, C., Volentine, R., Allard, S., & Levine, K. (2014). Trust and authority in scholarly communications in the light of the digital transition. Learned Publishing, 27(2), 121–134.
[8] Rogers, E. M. (2003). Diffusion of Innovations (5th ed.). Free Press.
[9] Salo, D. (2008). Innkeeper at the Roach Motel. Library Trends, 57(2), 98–123.
[10] Suber, P. (2012). Open Access. MIT Press.
[11] Swan, A., & Brown, S. (2005). Open access self-archiving: An author study. Technical report, JISC, UK.
[12] UNESCO. (2021). UNESCO Recommendation on Open Science. Paris: United Nations Educational, Scientific and Cultural Organization.
[13] Xia, J. (2010). A longitudinal study of scholars' attitudes and behaviours toward open-access journal publishing. Journal of the American Society for Information Science and Technology, 61(3), 615–624.

[14] Xia, J., & Sun, L. (2007). Factors to assess self-archiving in institutional repositories. Serials Review, 33(2), 73–80.
[15] Zuccala, A. (2010). Open access and civic scientific information literacy. Information Research

Appendix A: Sample Survey Instrument (Abridged).
Section A: Demographic Information.
● Designation: Professor/Associate Professor/Assistant Professor/Research Scholar.
● Faculty/Discipline: _____.
● Years of research experience: <5 /5-10/ >10.

Section B: Awareness (rate 1 = Strongly Disagree to 5 = Strongly Agree).
● I understand Open Access publishing.
● I can distinguish Green OA from Gold OA.
● I am aware of our university's institutional repository mandate/policy.
● I know the concept of Open Science beyond OA publishing.
● I know how to check publisher self-archiving permissions.

Section C: Behaviour (rate frequency: Never/Rarely/Sometimes/Often/Always).
● Published in a Gold OA journal.
● Self-archived a paper in the institutional repository.
● Deposited a pre-print in a repository.
● Shared the underlying research data openly.

Section D: Motivators and Section E: Barriers. (5-point Likert items corresponding to Tables 4 and 5 above.)