# Advancing Open Repositories in Ireland: A Survey and Strategic Recommendations for National Progress

**Ollscoil na Gaillimhe**
**University of Galway**

## Lead Authors

### Christopher Loughnane
University of Galway https://orcid.org/0000-0002-7789-6939

### Cillian Joy
University of Galway https://orcid.org/0000-0003-2470-0995

## Co-Authors

Fran Callaghan Maynooth University
Stephanie Ronan Marine Institute
Christoph Schmidt-Supprian Trinity College Dublin
Audrey Drohan University College Dublin
Donna Ó Doibhlin University College Cork
Dan Holden, Queen's University Belfast
Seán Harnett Research@THEA / Atlantic Technological University
Breda Connell Research@THEA / SETU Carlow Campus
Andrew Simpson RCSI University of Medicine and Health Sciences
Pádraig Manning Digital Knowledge Service – Health Library
Ireland, Health Service Executive
Caleb Derven University of Limerick
Niamh Brennan Representative for Ireland's OpenAIRE National
Open Access Desk located at Trinity College Dublin

# Contents

# Executive Summary

The National Open Research Forum (NORF)[1] Survey of Open Repositories in Ireland, conducted from April 24 to May 30, 2023, presents a comprehensive overview of the status of open repositories in Ireland. The survey, involving respondents from educational, governmental, and research institutions, aimed to assess various aspects of repository provision and management across key performance indicators.

Findings show metadata alignment is viewed as a critical need for repositories to support ambitious open research goals and projects across Ireland, which currently lags compared to some European peers in terms of coordinated national infrastructure, such as the Netherlands Research Portal[2], and HAL[3] and Recherche Data Gouv[4] in France. The survey found that repository metadata practices in Ireland are fragmented and require comprehensive national cooperation and coordination. Although there are concerted attempts to align individually with international metadata standards, such as OpenAIRE[5], alongside Plan S requirements[6], further work and support on the national level is urgently needed. The prevalence of Creative Commons licenses for content indicates widespread adoption and accessibility. Also, most repositories assign persistent identifiers; however, standardisation, expansion and integration in this area are also required to fully realise the potential workflow and efficiency benefits[7]. Disparities in staffing and resource allocation were observed, highlighting the necessity for dedicated personnel and robust resource allocation for efficient repository management and open requirements.

Most repositories were found to rely on external hosting providers rather than the institution themselves, highlighting the need for specialised services often beyond the remit and resources of the institution, and the current reliance upon commercial platforms for repository technical sustainability. The report also identifies visibility, funding, staffing, and ongoing technical operations concerns, all of which affect the ability of repositories to sustain critical open access standards and objectives. Overall, the findings of this survey align with the 2021 National Open Access Landscape Report, which found gaps and issues with the national open repository network[8].

The report concludes with recommendations for collaborative initiatives, training, metadata standardisation and national alignment, and regular assessments, providing a solid foundation for enhancing open repository networks in Ireland and addressing challenges in serving academic and research communities. Shared national or regional infrastructure and advocacy are critical solutions, along with closer coordination, training, and technical support. Areas of concern include the fragmented landscape of educational, governmental, and research repositories, with the data indicating that some repositories are struggling to maintain sustainability while others are better supported by their individual institutions. The data suggests increasing collaborative initiatives, training programs, standardised metadata adoption, and regular assessments to enhance Irish open repository progression on a par with European peers. The data is vital for addressing challenges and fostering collaboration to serve academic and research communities effectively. While this report offers detailed information drawn from the data, readers are advised to consult the full anonymised dataset for additional information that could not be included (Survey Dataset).

---

[1] https://norf.ie/
[2] https://netherlands.openaire.eu/
[3] https://hal.science/
[4] https://recherche.data.gouv.fr/en
[5] https://www.openaire.eu/
[6] https://www.coalition-s.org/plan_s_principles/

[7] See 'Persistent Identifiers: Addressing the challenges of global adoption': https://www.coar-repositories.org/news-updates/persistent-identifiers-addressing-the-challenges-of-global-adoption/
[8] See section 2.1.2 of the NORF National Open Research Landscape Report: https://repository.dri.ie/catalog/5q485c938

# Strategic Context

The National Open Repositories project[9] is a two-year initiative supported by the NORF 2022 Open Research Fund and led by the University of Galway. The project's primary goal is to strengthen Ireland's open repository network by addressing repository metadata fragmentation and other challenges impeding progress towards 100% open research goals for Ireland. This falls under Theme 2 – 'Achieving 100% open access to research publications' – of the 'National Action Plan for Open Research 2022-2030'[10]. NORF aims to align with the European Commission on open science, including the ongoing development of the European Open Science Cloud (EOSC). The report states, under Theme 2, that:

> By 2030 Ireland will have implemented a sustainable and inclusive course for achieving 100% open access to research publications. Provisions put in place to support a diverse open access publishing ecosystem and the retention of authors' rights will ensure Irish researchers have the freedom to choose from a range of quality options for making their research open access. An interoperable and robust network of repositories will contribute to making Irish research more visible and accessible locally and internationally. Benchmarking, drawing on a transparent national open access monitoring mechanism, will place Ireland amongst the leading countries internationally in terms of achieving 100% open access.

Concurrently, there is a growing consensus to establish and reinforce a broader strategic imperative to transform scholarly publishing through the principles of open research. The vision sees a resilient, globally interconnected repository infrastructure integrated into a next-generation scholarly communication ecosystem[11]. The ultimate goal is to empower researchers with unparalleled access to the entire corpus of research, fostering innovation, equity, and diversity in open publishing[12].

The current landscape holds promising opportunities, with funders endorsing repositories for their pivotal role in ensuring equitable access to research[13]. However, the lack of alignment among Irish open repositories challenges the effective functioning and representation of research output and the national network.

Despite facing financial constraints and limited visibility, open repositories play a crucial role in promoting equity and diversity in open publishing. They offer an alternative route to global open access, addressing concerns about the shift from paywalled access. Redirecting investments towards local publishing through open repositories can catalyse sustainability, innovation, and the development of community-driven research ecosystems. Transformative Agreements risk favouring commercial markets, potentially hindering community-based open access, and perpetuating a gradual and incomplete transition[14]. To avoid a siloed transition, a comprehensive embrace of repositories is essential. These repositories, viewed not as a parallel system but as vital public research infrastructure, facilitate equitable and diverse access to content. The collaborative nature of open repositories, hosted by research institutions, supports flexible and cost-effective alternatives to the existing for-profit publishing models. Additionally, open repositories are well-positioned to adapt to the next generation of publishing, incorporating mechanisms for open peer review and related open practices. The envisioned transformation of scholarly publishing necessitates a united effort to align repositories nationally, promote equity, diversify content access, and resist profit-centric publishing models. Embracing responsible and sustainable practices enables the scholarly community to optimise its infrastructure for the greater benefit of research, innovation, and global problem-solving.

---

[9] https://www.universityofgalway.ie/openrepositories/
[10] https://norf.ie/national-action-plan/
[11] https://www.coar-repositories.org/news-updates/openaire-liber-sparc-europe-and-coar-launch-joint-strategy-to-strengthen-the-european-repository-network/
[12] National Framework on the Transition to an Open Research Environment': https://enterprise.gov.ie/en/publications/national-framework-on-the-transition-to-an-open-research-environment.html

[13] Towards responsible publishing: a proposal from cOAlition S https://www.coalition-s.org/towards-responsible-publishing/
[6] https://www.coalition-s.org/plan_s_principles/

# Key Findings

The survey encompassed a diverse group of repositories affiliated with educational, governmental, research, and other institutions. The findings reveal significant trends and concerns, offering implications for future developments in open repositories. Although proportionally, most respondents are from educational institutions, representing a comprehensive cross-section of Ireland's educational sector, governmental and research institutions are also fully involved in open repository efforts. The diverse nature of institutional repositories in Ireland is a strength that can be built upon. The prevalence of Dublin Core for metadata schemas demonstrates its wide adoption, while the commitment to OpenAIRE compliance varies, suggesting the need for greater standardisation across Irish repositories to ensure interoperability and accessibility. Fourteen respondents (48%) adhered to some iteration of OpenAIRE guidelines. Most of those indicated compliance with OpenAIRE versions 1, 2, or 3, while a smaller fraction (10%) reported alignment with version 4. Some respondents acknowledged compliance but expressed uncertainty about the specific version. Seven respondents (24%) indicated uncertainty regarding compliance, citing reasons such as reliance on vendors for metadata compliance or ongoing assessments. Additionally, another seven respondents (24%) stated that their repositories were not compliant.

The utilisation of DSpace as the predominant platform, hosting nearly half of the surveyed repositories, underscores the importance of commercial hosting software. In addition, repositories currently rely on external hosting providers to provide technical infrastructure. This reliance, 76%, emphasises the importance of engaging with specialised commercial services to ensure repository sustainability. Resource discrepancies and staffing level variations highlight the need to increase dedicated personnel to ensure efficient management and support for repository managers.

A wide range of roles are associated with repository management. However, disparities in staffing levels and technical support arrangements mean that some repositories cannot fully engage with the development needed to further open research goals adequately. Deposit policies vary, with mediated workflows being predominant for quality control. Persistent identifiers, such as handles or DOIs, are widely assigned, emphasising resource traceability and findability; this area also requires more significant expansion and integration to support the open research agenda and other NORF-funded initiatives in this area, notably the National Open Access Monitor[15].

Repository size and content diversity are significant, indicating the broad range of academic and research materials made openly available through repositories. Repositories offer diverse content types, with predominant text, datasets, and images complemented by Creative Commons licenses for increased accessibility. Repository holdings also display linguistic diversity, with English and Irish as primary languages. Only a minority of repositories have implemented preservation policies, highlighting the crucial need for robust strategies to ensure long-term accessibility, particularly for content stored on commercial platforms. Developing strategies for content preservation is crucial to ensure robust open infrastructures and long-term viability in terms of accessibility, especially for content otherwise stored only on platforms that could rescind access at any later date.

Many respondents express confidence in repository sustainability but cite concerns such as lack of visibility, underfunding skilled staff availability, and ongoing technical operations. Solutions proposed include shared national or regional infrastructure, advocacy for repositories' role and importance, closer coordination, training for managers, and establishing a technical support community of practice.

---

[14] See Shearer, et al. 'Fostering Bibliodiversity in Scholarly Communications: A Call for Action!': https://zenodo.org/records/3752923. See also: https://www.coar-repositories.org/news-updates/transformative-agreements-are-not-the-key-to-open-access/

# Recommendations and Future Considerations

The data suggest several potential recommendations, summarised here:

Encourage collaborative initiatives among repositories to foster knowledge exchange and resource sharing.

Provide training programs and technical support for repository managers.

Promote the adoption of standardised metadata schemas linked to international best practices alongside persistent identifiers.

Conduct regular assessments to address sustainability concerns and ensure the long-term viability of open repositories.

The data provides a valuable foundation for enhancing Irish open repositories, fostering collaboration, and addressing the challenges repositories face in serving academic and research communities across Ireland.

# Introduction

## Background to the project

The survey forms one output[16] of a two-year initiative[17], supported by the National Open Research Forum 2022 Open Research Fund and led by the University of Galway, involving a consortium of ten institutional and organisational partners from across Ireland[19]. The project's primary objective is to enhance Ireland's network of open repositories. The project involves auditing the existing Irish open repository landscape, testing a standardised approach, and presenting a national roadmap. The proposed solution addresses metadata and support challenges at the source repository level rather than creating new technical infrastructure. This work aims to harmonise the currently disparate national open repository policies and standards, promoting interoperability by establishing standard operating procedures, guidelines, and services for coordinating Irish open repositories. The project is led by Dr Cillian Joy, Head of Open and Digital Research at the University of Galway Library, supported by the Project Manager, Dr Christopher Loughnane, also at the University of Galway Library. The Project Board, including Dr Joy, comprises subject matter experts in open access, open repositories, digital archives, and library leadership[20].

Project Board members serve as strategic advisers, meeting monthly to offer invaluable guidance and expertise to ensure the success of each project phase. They play a crucial role in providing insights into the evolving landscape of open repositories and publishing and contributing to developing effective strategies to engage the repository community throughout the project as it evolves. Their advisory role extends beyond mere oversight, encompassing proactive engagement with emerging trends, technological advancements, and community needs. By leveraging their collective experience, the Project Board ensures that the project initiatives align with best practices, guiding innovative solutions tied to current trends and expectations and enhancing the accessibility and impact of scholarly knowledge as the project seeks ways to strengthen Irish repositories and their network capabilities.

# Methodology

The Survey of Open Repositories in Ireland examines various aspects of open repositories in Ireland. The survey sought to gather insights into the characteristics of participating repositories, their technical infrastructure, staffing, metadata practices, content types, sustainability, and concerns. The survey was open for responses from April 24 to May 30, 2023. The survey consists of 40 questions about the participant's repository and additional questions on Plan S compliance. Participants were provided with a preview of the survey questions using the provided PDF link: Survey Questions PDF. The primary goal is to understand the Irish repository landscape better. The survey includes questions adapted from the recent survey conducted by OpenAIRE, LIBER, SPARC Europe, and COAR as part of their Joint Strategy to Strengthen the European Repository Network. We, therefore, acknowledge the contribution of questions from the recent survey undertaken by OpenAIRE, LIBER, SPARC Europe, and COAR.

Participant data was securely stored and handled following the University of Galway privacy policy: Privacy Policy. The only personal data collected are the participants' names and email addresses, and with consent, this information will be added to a private Irish open repository network inventory while an inventory of open repository names and URLs is published publicly[21].

# Repository Participants

The survey targeted repositories previously registered by the project as open repositories on the island of Ireland. Details collected included the name of the repository, its location, and its affiliated home organisation, whether an academic, governmental, or research institution. The population included various institutions, such as universities, technological universities, specialised educational institutions, governmental institutions, and research entities. A range of roles associated with repository management were identified among respondents.

[21] https://www.universityofgalway.ie/openrepositories/inventory/

# Technical and Metadata Information

The survey gathered data regarding the technical aspect of repository management, including the software platform used, whether hosted locally or externally, frequency of updates, and machine readability and harvesting practices. Respondents were asked about metadata schemas available to depositors and OpenAIRE compliance. Integration with author identifiers, such as ORCID, was assessed, including other measures which enhance author identification and research impact tracking.

# Services and Staffing, Content, Deposit Policies

Staffing levels and repository roles were examined to understand the allocation of resources for repository management. The survey gathered data on the number of staff members dedicated to the repository, their roles, and FTE staffing levels related to particular tasks and functions, including repository management, technical support, and metadata and content curation. The survey collected data on the types of content held in repositories, ranking content types by prevalence. Content languages were also analysed. The survey also inquired who could deposit into the repository and deposit workflow practices, distinguishing between mediated and direct deposits. Deposit policies and workflow are essential for content control and validation, so highlighting the process was necessary.

# Sustainability Concerns and Solutions

After checking that the repository had been registered as an Irish open repository, the survey established the particulars of each repository, including the name and the role of the respondent, the name of the repository, its location, and the affiliated home organisation, including classification as an academic, governmental or research institution.

# Survey Results

After checking that the repository had been registered as an Irish open repository, the survey established the particulars of each repository, including the name and the role of the respondent, the name of the repository, its location, and the affiliated home organisation, including classification as an academic, governmental or research institution.

# Repository Participants

Most respondents (27 of 29 or 93%) are based in Ireland, while the remaining two are based in Northern Ireland. Nineteen (66%) repositories classified themselves as belonging to educational institutions, five (17%) as governmental, three (10%) as research, and two (7%) classified themselves as 'other'. However, after correcting misclassification, 20 (69%) were identified as educational, seven (24%) were governmental, and two (7%) were research. The diversity of educational entities, from universities to colleges to technological institutions, underscores the comprehensive integration of open repositories within Ireland's educational landscape. Additionally, the involvement of governmental institutions, such as the Houses of the Oireachtas and healthcare agencies, demonstrates a multifaceted approach to open research dissemination at the governmental level.

Educational respondents included institutions from across the entire range of Ireland's third-level education sector, which is highly varied across all the regions of Ireland. The sector comprises traditional universities, newer technological universities amalgamated from Institutes of Technology, and various smaller colleges and specialised educational institutes such as the Royal College of Surgeons. Government institutions with open repositories included Teagasc, the Marine Institute, Houses of the Oireachtas, the Health Research Board, the Health Service Executive, and Tusla Child and Family Agency. A repository of note, acting both as a research-performing organisation and national trusted digital repository for Ireland's humanities, cultural heritage, and social sciences data, the Digital Repository of Ireland (DRI) serves higher education institutions, cultural heritage institutions (the GLAM sector of galleries, libraries, archives, and museums), government agencies, and county councils. While funded by the Department of Further and Higher Education, Research, Innovation and Science (DFHERIS) via the Higher Education Authority (HEA) and the Irish Research Council (IRC), it is managed by the Royal Irish Academy, Maynooth University and Trinity College Dublin.

Respondents had a wide variety of roles associated with the repositories, including Administrator, Senior Associate Publisher, Systems Librarian and other technical roles, Head of Library, Associate Directorate for Research and Teaching, Co-Designer and Administrator, Repository Librarian, Repository Manager, and IR Manager. 27 of 29 (93%) of repositories are funded by their host institution, one (3%) is funded by a government body, and the last repository responded: "External project funding; Fees for depositors; Our main funding is Government funding from the [Government body] and [Government body]. This is complemented by project funding and membership fees for depositors". 23 of 29 repositories (79%) are registered with OpenDOAR, a global-level registry, while 10 of 29 (34%) indicate that they are also

registered with a national-level registry (Figure 1). Other registries that were indicated include Re3Data, FAIRsharing and OpenAIRE. Four (14%) respondents also replied that they are part of a national-level service or network. One (3%) noted that they comply with the now-defunct national-level, multi-institution repository RIAN[22]. One (3%) noted that, while they are not currently part of a national-level service or network, they are part of the newly instituted Technological University Network (TU-NET) portal (https://tunet.openaire.eu/), an affiliation of the Technological Universities in Ireland provided by OpenAIRE.

**Question 10 Select which directory the repository is registered with?**



| | OpenDOAR | Re3Data | National repository directory |
|---|---|---|---|
| | FAIRsharing | Don't know | None | Other |
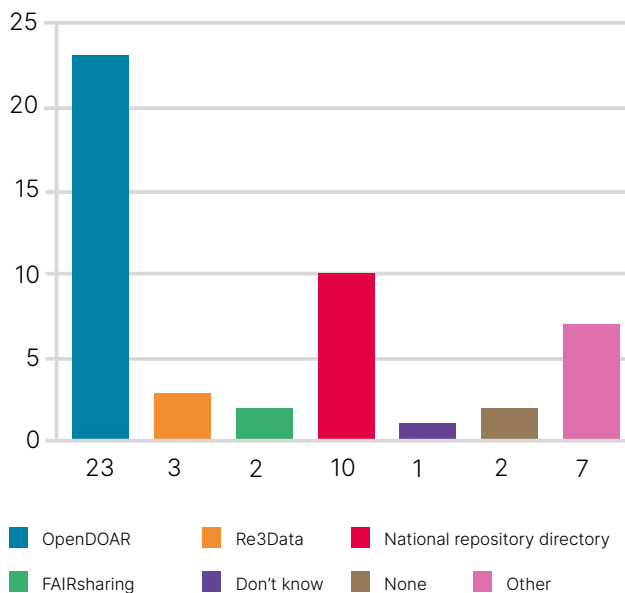
*Figure 1 Question 10 Select which directory the repository is registered with?*

The diversity of educational entities, from universities to colleges to technological institutions, underscores the comprehensive integration of open repositories within Ireland's educational landscape.

[22] See http://roar.eprints.org/3981/

# Technical Information

Only 7 of 29 (24%) repositories are hosted locally by their institution, while 22 (76%) rely on external hosting providers. Some locally hosted repositories still rely on commercially provided or managed platforms such as services and managing Dspace, Digital Commons and ePrints, while others have developed their own platforms in-house. Commercial platforms for local and externally hosted repositories include DSpace, ePrints, Figshare, Knowvation, datAdore, Digital Commons, and Pure (Figure 2). DSpace is the most popular platform, hosting 14 out of 29 repositories (48%). The prevalence of external hosting (76%) suggests a reliance on specialised services, possibly due to technical expertise or resource constraints. DSpace's popularity (48%) underscores its reputation as a reliable platform for open repositories. Many platforms have been heavily customised by different repositories over the years for local needs. Two (6%) also indicated that their repository is CoreTrustSeal certified, while one said they are certified but did not know which certification.

**Question 12 What software platform does the repository use?**
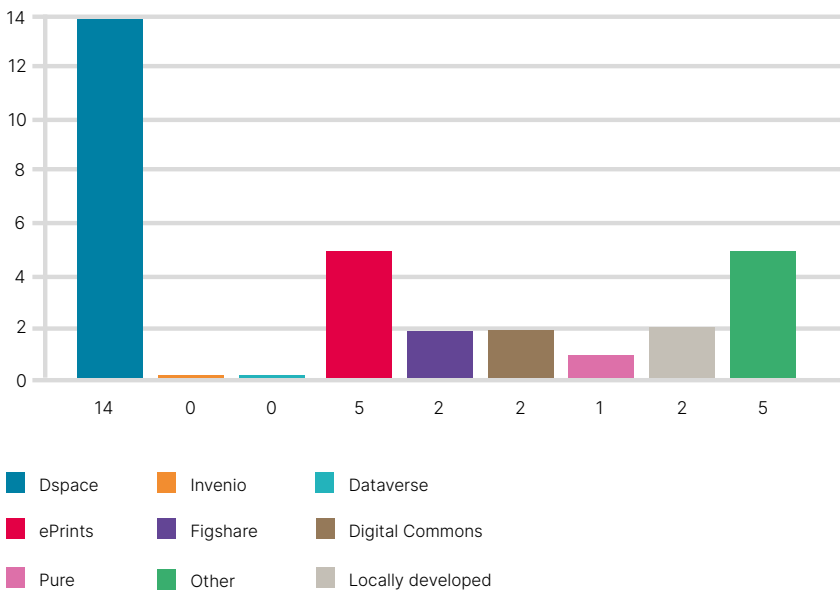


*Figure 2 What software platform does the repository use?*

Of these platforms, 11 were updated in 2023, 3 in 2022 and 2 from 2020-2021. 6 respondents did not know when the platform was last updated, and only a couple had not been updated later than 2019. Eleven respondents indicated the platform would again be updated in 2023, 3 in 2024-2025, 12 did not know the next update planned, while three responded 'other'. The commitment to regular updates (38% in 2023) signifies ongoing efforts to enhance repository functionality, metadata compliance and development. However, the presence of platforms without updates since 2019 emphasises the need for further support and maintenance across the repository landscape.

# Services and Staffing

Varied staffing levels and technical support arrangements indicate significant resource disparities among repositories. Proper technical and managerial support is vital for ensuring these repositories' seamless operation and maintenance. We asked how many staff members (in total Full-Time Equivalents) are employed to operate/manage the repository. This question highlighted significant variance in repository resources and support, ranging from zero to 0.1 and 0.2 on the lower end to 10 at the higher end. At 10 FTE, the DRI understandably topped the category with its wide-ranging, cross-Ireland remit. The number of repositories with less than one FTE of any grade was significant (7 of 29 or 24%). Other respondents indicated that the staff working on the repository were also split in their duties, with one, for example, responding: 'a team of three where repository administration is one small part of overall role.'

When we asked what the roles were for those staff members:

- Four (14%) repositories indicated no dedicated repository manager.

- Eight (28%) had managers working on the repository 1-25% of their time, with one only working on it occasionally when upgrades are needed.

- Six (21%) had managers working on the repository for 26-50% of their time

- Three (10%) had managers working on it 51-75% of their time.

- Three (10%) had managers working on it, 76-99%.

- Only 5 (17%) repositories have a full-time repository manager.

Repository managers often split their time between other librarian functions. One manager working at a government institution, for example, said that their role consisted of a 'solo librarian' working three days in the library (FTE.6), fully responsible for the repository and all library tasks while working one additional day as a Data Protection Officer. Ten repositories (34%) have zero FTE dedicated to technical support. Most of these are externally hosted, so this is to be expected. However, one locally hosted repository also indicated that they have zero staff dedicated to technical support. 13 (45%) have 1-25% FTE dedicated to technical support. Two (7%) have 76-99%, while 3 (10%) have full-time dedicated technical support staff.

Staff working on metadata and content management is similarly varied. Eight (28%) repositories have a staff member dedicated at 100% FTE to metadata and content management roles, down to two (7%) with zero staff dedicated to a similar role. Overall, 15 (52%) of repositories had a staff member working 50% or less on metadata and content management. When asked 'Does the repository provide any other services?' and asked to specify, 12 (41%) answered affirmatively and provided responses, including:

- Archive/discipline-specific archive (3 respondents)

- Yes, as a solo librarian, full library management on a part time basis.

- Journal publishing, data portal, [Redacted] Academic Press

- Depositing from CRIS [current research information system] on behalf of researchers; CRIS publication data clean up

- It is intended that the repository will also become a knowledge sharing portal

- Publishing of journals and events

- We provide advise (sic) on setting up journals on OJS which is integrated into [Repository]

- eThesis Submission system: all [Institution] research theses are compulsorily submitted by students via the CRIS/[Repository] interface (using [Redacted] – Student Information System) data), archived theses sent electronically to the [Redacted] for printing/binding;[Redacted] Archive: provides additional resources to enhance the accessibility of research for people with visual and aural disabilities ;[Redacted]: national voluntary open access legal deposit service - part of [Repository] with an individual domain name ([Redacted]); [Redacted]: collaboration of Depts of Classics, History and History of Art to provide a large and important collection of OA images and accompanying VRA Core metadata for teaching and learning ([Redacted url])

- Digital Preservation, Aggregation to platforms such as Europeana, access to training and education.

- Digitisation services, metadata templates, copyright clearance for collections, consultation services

- Peer review

From these responses, it is evident that repositories function to underpin a distinct and diverse open research publishing ecosystem beyond their stereotypical depositing functions.

From these responses, it is evident that repositories function to underpin a distinct and diverse open research publishing ecosystem beyond their stereotypical depositing functions.

# Content and Metadata

When asked about metadata schemas available to depositors, 27 (93%) said Dublin Core, and 6 (21%) said DataCite, with other responses including VRA Core, EAD, MARCXML, and MODS (Figure 3). The prevalence of Dublin Core reflects its widespread adoption of metadata standardisation. The use of other schemas, however, such as DataCite, showcases repository flexibility in accommodating diverse data types.

**What metadata schemas are available to depositors?**
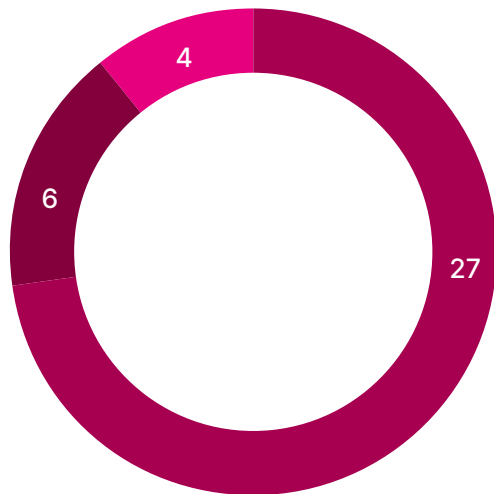


Dublin core | Datacite | Other

*Figure 3 What metadata schemas are available to depositors?*

While some repositories are fully compliant with specific versions of OpenAIRE Guidelines, others are in the process of assessing compliance or transitioning to new systems. The commitment to align with standards and the diverse compliance status reflects the need for standardisation across the board. When asked whether their repository is compliant with some version of OpenAIRE Guidelines for Literature, Institutional, and Thematic Repositories, 14 of 29 (48%) replied that they are compliant with some version of OpenAIRE guidelines, with the majority compliant with OpenAIRE version 1, 2, or 3 (Figure 4). Three (10%) are compliant with version 4. Others noted that, while compliant, they were unsure of

the version. Seven (24%) said that they were unsure about compliance, with some saying, for example, that their vendor may comply, while others are still assessing compliance. Seven (24%) said that they were not compliant. Two also said that they are compliant with data guidelines, two with CRIS (current research information system) guidelines, and one is compliant with software guidelines. The data reveals a variety of responses regarding OpenAIRE compliance and an area to be assessed more thoroughly when designing a community metadata approach.

**Is the repository compliant with the OpenAIRE Guidelines?**



OpenAIRE Guidelines | OpenAIRE Guidelines | Data guidelines
CRIS Guidelines | Software guidelines | Does not comply
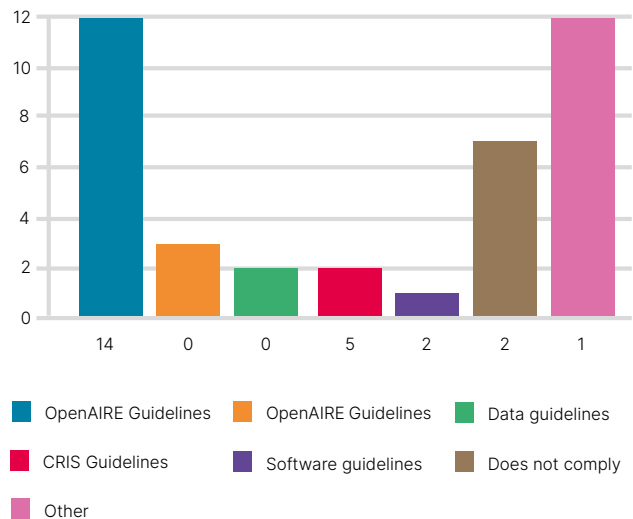Other

*Figure 4 Is the repository compliant with the OpenAIRE Guidelines?*

When asked, 'Is high-quality metadata (including open access status, reuse licence, and funder & grant ID) openly available for each Item?' 21 (69%) answered yes, and 8 (31%) answered no. Qualifying comments to those who answered yes included:

- CC licenses are visible on document files that are uploaded, as are embargo periods; funder grant IDs are not available.

- Author-name normalisation, keywords, abstracts, open access status, reuse license, and funder and grant ID.

- We collect all of the data mentioned in Q 46. However, the availability of that data in any given resource differs wildly and is often incomplete.

- [Repository] is OpenAIRE compliant and that includes fields for funder and grant ID, OA status and rights along with the CC licence.

- Metadata includes descriptive metadata, licence, funder info, etc. Grand IDs can be entered, but as not all of our data comes from research projects (e.g. cultural heritage collections) this is not mandatory. Objects can have some access restrictions, if so this is indicated in the metadata, the majority are open access.

- DC available for each.

- Rights statements are used, Creative Commons licensed where appropriate, funder and grant ID published where applicable.

Those who answered no to the above question on high-quality metadata were asked to qualify why not. Six answered 'resources', one answered 'cost', and one answered 'policy' (Figure 5).

**If you answer 'No' to Q46 (High quality metadata), please select a reason**



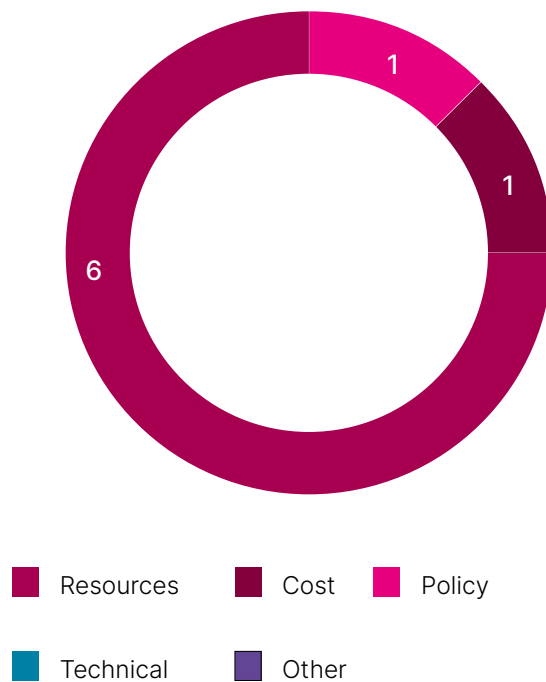Resources    Cost    Policy

Technical    Other

*Figure 5 Reason for metadata quality issues*

Those who answered no were asked to rank the importance of high-quality metadata on a 5-point scale. Six (75%) rated the importance at 5 while 2 (25%) rated it at 4, demonstrating the importance to repositories who are currently struggling to deliver high-quality metadata (Figure 6).

**If you answer 'No' to Q46(High quality metadata), please rank a (1=Not important 5= Very Important) how important you see this functionality**
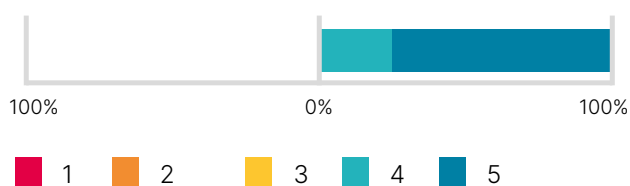


*Figure 6 Importance of high quality metadata*

Comments and qualification by those who answered yes included:

- All of the relevant data is available through an open OAI/PMH service.

- Creative Commons license link is used.

- Our IR is successfully harvested by international repositories.

- Stored in the database as xml.

- We follow the Dublin Core metadata standards.

- [Repository] uses the usual Dspace protocols for machine readability of its metadata and processes uploaded text-based content for full text searching.

- Metadata can be downloaded as XML or accessed as JSON via a variety of APIs. We are currently exploring FAIR Signposting headers, RO-Crate, etc.

Respondents were also asked 'Is machine readable metadata (including open access status, reuse licence, and funder & grant ID) embedded in the item itself?' 11 (38%) answered yes, while 18 (62%) answered no. Those who answered no were asked to rank the importance of machine-readable metadata on a 5-point scale. Four answered 5, seven answered 4, five answered 3, and two answered 2 (Figure 7).

**If you answer 'No' to Q50 (Machine readable metaverse), please rank (1=Not important 5= Very Important) how important you see this functionality**
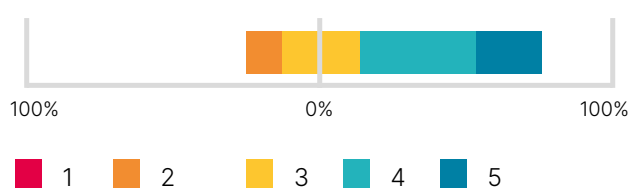


*Figure 7 Importance of machine-readable metadata*

For those who answered no, the reasons included technical (7 respondents), cost (1 respondent), policy (1 respondent), resources (7 respondents) and other (2 respondents) (Figure 8).

**If you answer 'No' to Q50 (machine readable metadata) please select a reason**
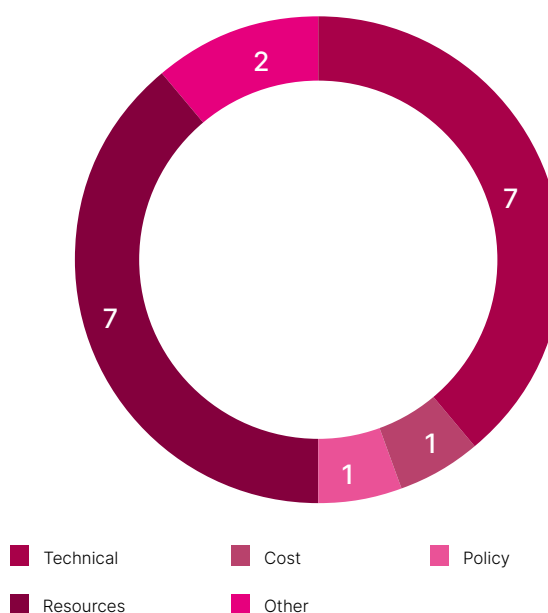


*Figure 8 Reason for machine-readable metadata issues*

28 of 29 (97%) offer one or more variant of Creative Commons license, including CC-BY-NC-ND and CC-BY. The one governmental repository that does not offer any license is considering Creative Commons. One respondent wrote 'CC-BY-NC-ND by default, CC-BY when appropriate (Plan S, some Gold OA)', while another said:

> 'Principally Creative Commons licences & Open Data Commons licences. We also allow depositors to specify rights reserved, possibly also with certain exceptions in the form of an Educational Use statement from RightsStatements.org or a statement that an item is in copyright, rights are reserved, but that it is an orphan work and has been registered with the EU IPO Orphan Works Database.'

When asked which author IDs the repositories link to, 12 (41%) responded that they link to ORCID while 16 (55%) do not link to any author IDs. Integration with ORCID IDs enhances author identification and research impact tracking across institutions and needs further promotion and integration. One repository said that they link to 'Other IRIS' through integration with their Vidatum CRIS. Another repository said they link to ORCID, VIAF, LOC-NAF, and ULAN. One repository indicated that while they do not link, they are planning ORCID links and that there is currently a link to the CRIS author profile, which contains the ORCID ID. Another repository said that although they collect ORCID IDs, they do not currently map to ORCID.

When asked if the repository assigns persistent identifiers to the resources, 16 (55%) assign handles, 12 (41%) assign DOIs, and four (14%) assign both. Five (17%) assign neither DOIs nor Handles. One wrote that 'DOIs are assigned to all resources published under a journal with an ISSN on the publishing end of the platform. DOIs previously assigned can be added to resources across the repository.' While assigning handles and DOIs is common practice for many repositories, ensuring further resource traceability and long-term accessibility is needed here and is an area for further support. Follow-up questions sought to assess PIDs' importance and barriers to implementation. 21 (72%) of respondents said that PIDs are created for each item in the repository (Figure 9).

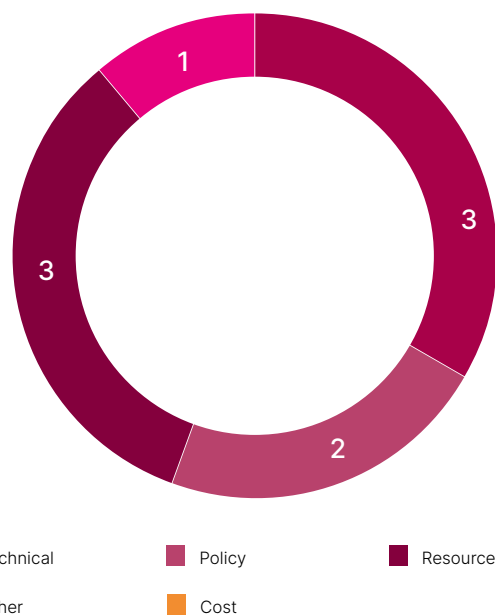## Arepersistent identifiers (PIDs)(DOIs, handles, ect.) created for each item deposited in the repository
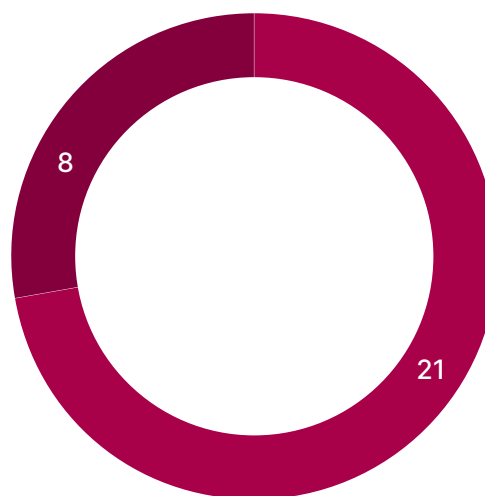


Yes   No

*Figure 9 Are PIDs created for each item*

For the 8 (28%) that do not assign PIDS, the reasons included technical (3 respondents), policy (2 respondents), resources (3 respondents) and other (1 respondent) (Figure 10).

## If you answer 'No' to Q42 (PIDs) please select a reason



Technical     Policy     Resources
Other         Cost

*Figure 10 Reason for not creating level PIDs*

Most respondents who answered no, however, did see PIDs as necessary for repositories (Figure 11).

**If you answer 'No' to Q42 (Machine readable metaverse), please rank (1=Not important 5= Very Important) how important you see this functionality**
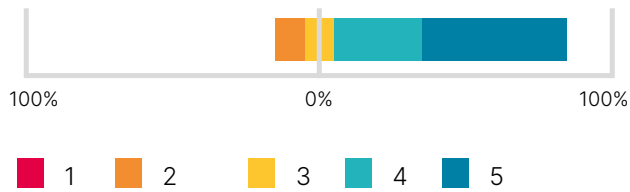


100%        0%        100%

■ 1    ■ 2    ■ 3    ■ 4    ■ 5

*Figure 11 Importance of item level PIDs*

In terms of repository size, item count reveals the varying sizes of different repositories, ranging from hundreds to approximately 700,000 items. Irish repositories encompass a wide range of content and serve different purposes within the academic and research landscape, with differing age and maturity levels. Responses indicate that the most prevalent content types in the repository are text, datasets, images, interactive resources, and learning objects (Figure 12). While other content types like design, cartographic material, patents, software, and sound are also present, they tend to rank lower in prevalence. The diversity of content types suggests that the repository caters to a wide range of academic and research materials. In terms of content type, ranked order puts Text, Dataset, and Image into first, second and third place. 27 (93%) ranked text as their first choice, and the remaining two ranked text as their second choice. While datasets were not ranked as any repository's first choice, it was ranked highly as second, third and fourth choice. Images were ranked as the first choice for two repositories, although they ranked third overall.

**What are the predominant types collected in the repository? (subset of the COAR Resource type vocabulary)(order to rank)**



1. Text
2. Dataset
3. Image
4. Learning object
5. Cartographic material
6. Design
7. Interactive resource
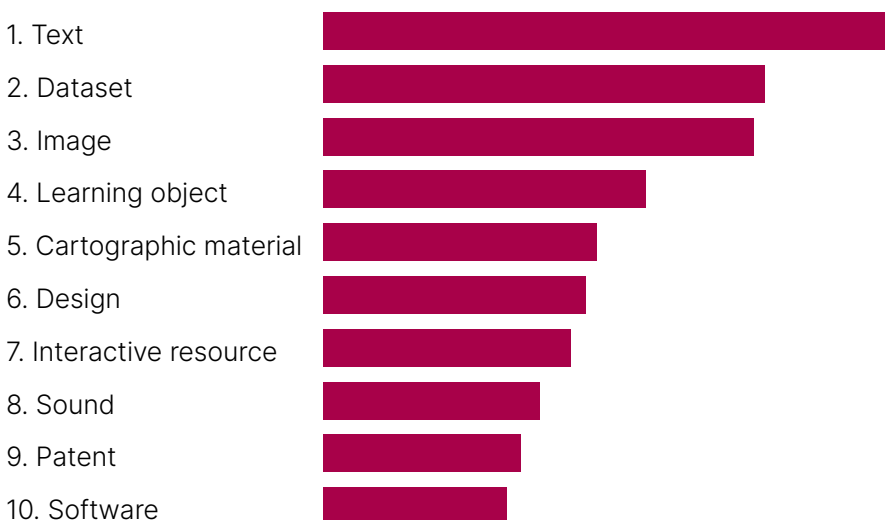8. Sound
9. Patent
10. Software

*Figure 12 Content types in the repository*

Although English is the predominant language for metadata and content in the repository, Irish is the second most prevalent language (second choice for 21 (72%) and third choice for 8 (28%) of repositories (Figure 13). While other languages are present, they tend to be less common in the repository's content.

**What are the predominant language(s) of the metadata and content in the reposiitory? (order to rank)**



*Figure 13 Predominant language of the repository*

When asked who can deposit into the repository, 18 (62%) answered 'persons associated with the institution.' 9 (31%) answered 'repository administrators', while 4 (14%) answered 'persons with relevant domain content.' Other answers included 'staff and faculty', 'anyone working in the [governmental service]', 'persons funded by the [funding agency]', 'authorised persons affiliated with Member Institutions', 'anyone' and 'users who register via a user access management system (in practice staff from government departments and agencies who wish to lay documents)'. Some also have close integration with the institutional CRIS where users deposit directly, which is then routed to the repository. The data demonstrates a variety of deposit policies across different repositories. Some repositories have open policies, allowing a broad range of contributors, while the majority restrict deposits to specific groups such as administrators, affiliated persons, or authorised individuals. These policies reflect the institution-specific author affiliations of most Irish repositories.

When asked to select one or more answers about deposit workflow, 23 of 29 (79%) selected 'Mediated deposit', 14 (48%) selected 'Users deposit directly to the repository,' 11 (38%) selected 'Metadata validation,' 9 (31%) selected 'Content validation,' and 9 (31%) selected 'Copyright validation' (Figure 14). The prevalence of mediated deposits (79%) suggests a robust quality control mechanism, ensuring accurate

metadata and content validation. The diverse responses indicate tailored approaches that accommodate various researcher needs and preferences.

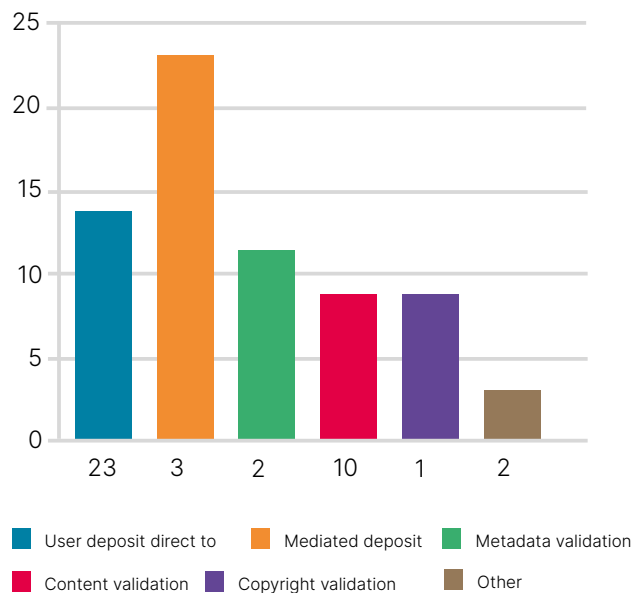**What is the deposit workflow for the repository?**



*Figure 14 Deposit workflow in the repository*

Some respondents selected all or mostly all categories, while some selected only one, such as metadata validation. Others also qualified their answers, with one noting that deposits are routed to them through the CRIS, while another wrote:

> 'Automatic metadata validation is carried out on all deposits, and in some cases a staff member will also perform some further metadata validation and potentially some content validation - this tends to be rolled into pre-deposit education though, rather than being a standard part of the deposit workflow. Only in some cases will Repository staff perform metadata and content validation during or post-deposit. This may be done for certain depositors who require additional support.'

While most respondents selected answers rather than giving detailed responses, some variety of mixed deposit workflow is likely to be the norm rather than the exception.[23]

[23] This has been confirmed through later interviews with a number of repository managers.

Only 10 (34%) of repositories have implemented preservation policies, indicating that a minority are taking a proactive approach to safeguarding digital assets and ensuring their long-term accessibility and usability (Figure 15). 19 (66%) have no preservation policy, suggesting they may need to develop strategies for preserving their content to mitigate preservation and accessibility risks, especially if the content is stored only on commercial publishing platforms. In this context, it is crucial for repositories to have preservation policies in place to protect their digital holdings, especially given the importance of long-term open access to research and scholarly materials originating from the research work of their staff and affiliates.

## Does the repository have a preservation policy?
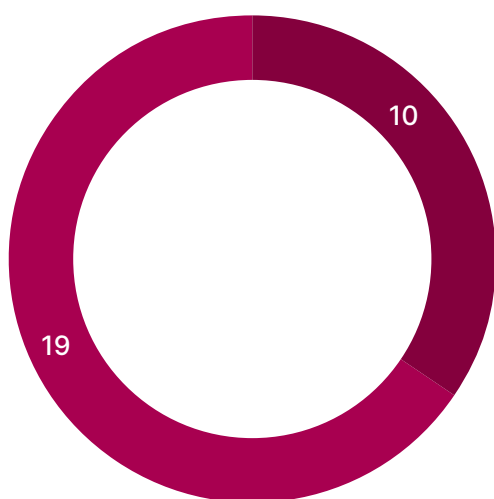


Yes    No

*Figure 15 Does the repository have a preservation policy?*

Regarding usage statistics and analytics, 26 (90%) rely on local repository statistics, 12 (41%) utilise Google Analytics, 6 (21%) use OpenAIRE statistical data, one (3%) makes use of PlumX Metrics in addition to other statistics, and one (3%) is still deciding (Figure 16).

## Please select the type of usage statistics or other analyatics services the repository offers:
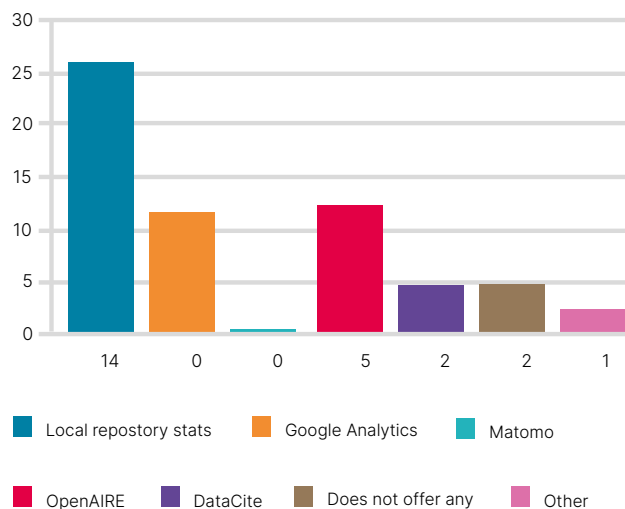


Local repostory stats    Google Analytics    Matomo

OpenAIRE    DataCite    Does not offer any    Other

*Figure 16 Type of usage analytics services*

When asked 'Does the repository provide reliable, continuous service (less than 24hrs unplanned downtime per annum)?' 28 (97%) of 29 answered yes. For the one who answered no, the reason given was that the repository is "largely self-managed".

When asked 'Does the repository service offer helpdesk functionality, with a monitored email address as a minimum?' 27 (93%) answered yes (Figure 17).

**Does the repository service offer helpdesk functionality, with a monitored email address as a minimum?**



2 27
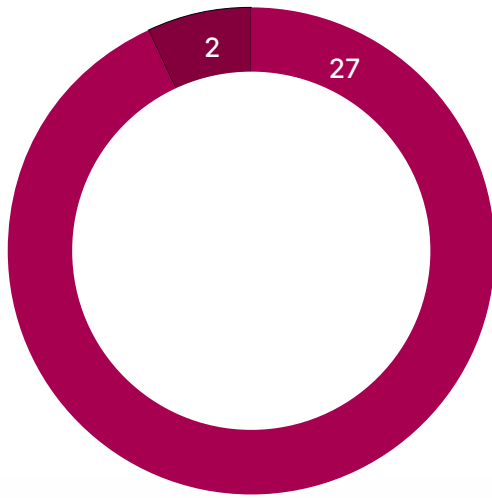
■ Yes ■ No

*Figure 17 Helpdesk funcationality*

Those who answered no gave the reason of 'policy' and "Very basic IR".

Respondents were asked 'Is there bulk upload functionality, allowing the automated bulk upload of both article and metadata from depositors (e.g. SWORD)?' 20 (69%) said yes while 9 (31%) said no. For those who qualified their yes responses, these included.

- Authors can upload a BIBTEX file.

- CSV upload, generally done by FTEs.

- Bulk upload is possible via Scopus, Pubmed, Pure and ORCID. However, the functionality is not reliable and in the case of ORCID it is not done via best practice per ORCID 2 way authentication recommendations.

- SWORD integration in place for upload from our universities CRIS system We are testing bulk uploads from Scopus and Pubmed in Dspace 7 but it's early days as we have just upgraded to DSpace 7.

- SWORD 2 has been used by [Repository] to automatically bulk upload 11K items and metadata from the [Redacted] Project; manual bulk uploads are also available (via .csv); OAI PMH harvesting is also possible but not currently used.

- Upload via our Cris system (Vidatum).

- We provide a bulk upload service which accepts metadata as XML. We do not use SWORD or any other deposit protocol.

Reasons for not providing bulk upload functionality included technical (4 respondents), policy (2 respondents), and resources (2 respondents) (Figure 18).

Other responses included "Quantity has not required this to date," "Loss of access to server, over-complication of bulk import process by vendor," and "Content is curated. Only team members ingest content. All content is done via bulk upload through the Ingest workflow."

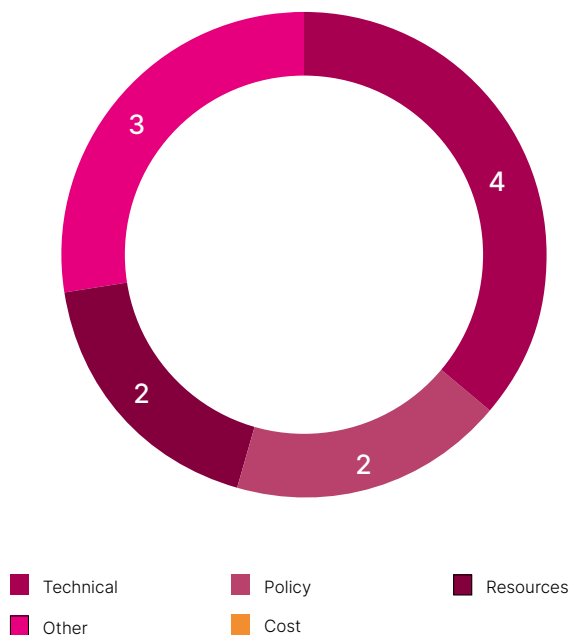**If you answer 'No' to Q62 (bulk upload), please select a reason**



Technical
Policy
Resources
Other
Cost

*Figure 18 Bulk uploads?*

Respondents were asked 'Can the full text be stored in a machine-readable format (such as e.g. JATS/XML)?' Thirteen (45%) answered yes, while 16 (55%) answered no. Reasons why not included technical (4), policy (2), and resources (2). Other responses included

- All of the above. Not currently available in DSpace. We'd be interested in a 3rd party service provider for JATS/XML (with the provisor that such a national service would not be axed without consultation like Rian).

- Not yet a priority, the majority of our digital objects are not publications and this is therefore not a necessary feature (the majority of our data are images). We would consider it in future.

- Technical aspects outside our remit.

When asked to rank the importance of this feature on a 5-point scale, four answered 5, three answered 4, seven answered 3, and one answered 2.

Respondents were asked 'Can the repository support openly accessible data on citations (according to the standards by the Initiative for Open Citations - I4OC or similar).' Seven said yes, 8 said no, and 14 said 'I don't know.' Qualified answers included "DC citation identifier is established" and "Figshare provide support by hosting publications. It is also possible to extract data using the Figshare API." When asked to rank the importance of this feature on a 5-point scale, three said 5, four said 4, two said 3 and one said 2.

Respondents were also asked 'Are standard interoperability mechanisms such as OAI-PMH and OpenAPIs provided (that allow harvesting of both metadata and full text)?' 27 (93%) said yes while 2 (7%) said no. Qualified answers from those who answered yes, included:

- DSpace APIs (x2, versions <1.6.x and 7), OAI-PMH, SWORD v1 and v2; SQL and FTP used for ETDs.

- OAI-PMH supported and used by 3rd parties such as JSTOR to harvest data and assets.

- We expose OAI-PMH feeds for a variety of formats (e.g. Dublin Core, Europeana Data Model) Full text is not included directly in the OAI-PMH feed but links to the digital assets are included.

- OAI-PMH standard adhered to in DSpace.

- OAI-PMH is provided. IIIF Images API, IIIF Presentation API, and the unAPI protocol are available. Other APIs are in development stages.

For those who said no, one noted that this feature was "coming soon, in technical development,' while the other said resources as the reason.

Only 10 (34%) of repositories have implemented preservation policies, indicating that a minority are taking a proactive approach to safeguarding digital assets and ensuring their long-term accessibility and usability

# Sustainability and Concerns

When asked the question, 'How sustainable do you consider the repository for the next 3 years?' three (10%) respondents answered that their repository is currently 'not sustainable', 8 (28%) answered 'somewhat sustainable', while 18 (68%) answered 'very sustainable'. The majority of respondents expressed a high level of confidence in the sustainability of their repositories. This suggests they believe their repositories are well-equipped to endure challenges and continue their operations effectively. Some respondents indicated a moderate level of confidence in the sustainability of their repositories. This group likely has some concerns or uncertainties about certain aspects of sustainability but still believes their repositories can manage challenges to a certain extent. A minority of respondents expressed a lack of confidence in the sustainability of their repositories. This indicates that these repositories might face significant challenges or lack the resources necessary for long-term operation. While it is heartening to see the majority of respondents view their repository as very sustainable in the near term, some repositories are struggling to continue, indicating potential areas that might need improvement or additional community and resource support to ensure long-term viability.

Respondents were asked to indicate their level of concern for a range of issues, including content recruitment, ongoing technical operations, upgrading software to new versions, metadata curation, employing skilled staff, lack of visibility, and underfunding (Figure 19). Many respondents were 'very concerned' or 'somewhat concerned' with all areas, particularly lack of visibility, underfunding, employing skilled staff and ongoing technical operations. While there are varying degrees of concern expressed for each category, upgrading software to new versions and metadata curation ranked slightly lower in terms of areas of concern, with an even balance between those who were and were not concerned, although a desire for more effective and up-to-date repository management practices still seems to be essential.
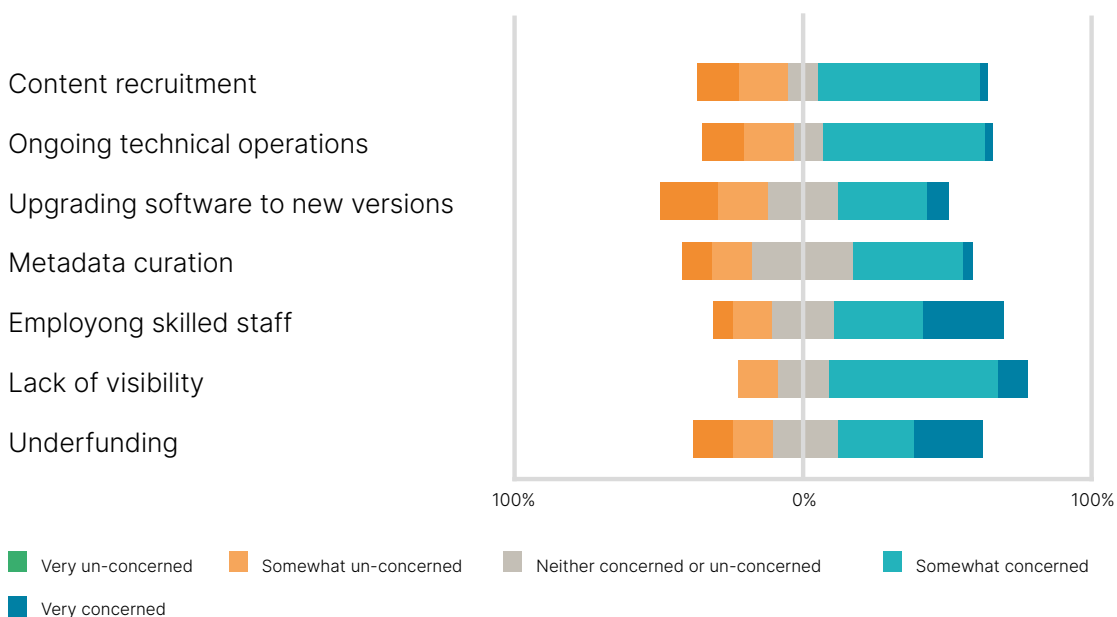


*Figure 18 Bulk uploads?*

Respondents were asked to rank, in terms of importance, the following solutions to the previous concerns cited: closer national or regional coordination of repositories; shared national or regional infrastructure for repositories; community of practice for technical support; training for repository managers; advocacy on the role and importance of repositories; other. Although the order below is ranked, the answers were diverse, suggesting the importance of implementing many or all of these solutions where possible, as each repository and its staff have differing and diverse needs depending on their particular situation.

1. Shared national or regional infrastructure for repositories: Several respondents ranked This activity the highest in importance. It indicates a strong preference for developing and utilising shared infrastructure to enhance the effectiveness and interoperability of repositories.

2. Advocacy on the role and importance of repositories: Advocacy is considered the second most important activity, suggesting that stakeholders recognise the need to promote and raise awareness about the significance of repositories in the academic and research community as a viable route for open research publications.

3. Closer national or regional coordination of repositories: Closer coordination at the national or regional level is considered the third most important activity, showing that respondents see the benefits of aligning and coordinating repository efforts on a larger scale.

4. Training for repository managers: Training for repository managers is ranked fourth in importance, highlighting the value of skills development for those responsible for managing and curating repository content.

5. Community of practice for technical support: Building a community of practice for technical support is ranked fifth in importance. This underscores the need for collaborative platforms for technical and subject experts to share knowledge and best practices.

6. Other: "Other" activities received mixed rankings and likely encompass a range of additional initiatives or approaches that respondents deemed relevant but may not fit into the specified categories.

While it is heartening to see the majority of respondents view their repository as very sustainable in the near term, some repositories are struggling to continue, indicating potential areas that might need improvement or additional community and resource support to ensure long-term viability.

Although responses are ranked, the data reflects respondents' priorities and preferences regarding activities and solutions addressing repository-related issues. While shared infrastructure and advocacy are considered highly important, the focus and ranking of activities differed among stakeholders involved in repository management and development.

At the end of the survey, respondents were asked for any other comments, which included:

- The ranking questions force choices that don't apply, so Q22. 'What are the predominant content types…' Q41 is more representative. Q38. 'Rank from the top, the following activities in terms of importance in addressing the issues in the previous question (Q. 37)' is almost completely unrepresentative of our views.

- We will need to procure a new repository (under procurement law) in 2024.

- On the question of reservation policy, we do not have one currently but are in the process of joining a Private LOCKSS Network, or PLN, as part of the Digital Commons network.

- The major challenge is that of lack of staff resources and no source of funding to support this. That issue was identified in the NORF Action Plan and its Landscape report. Open access is a cultural change and needs a strong human interface which is sorely lacking in this country.

- We have only recently upgraded and moved to a hosted solution so lots of issues raised in this survey are currently being worked through.

- We are not primarily a publications repository, so some of the questions relating to full-text might not be entirely relevant to us.

- [Government Body Publishing Platform] is a publishing platform (akin to a journal), rather than a repository meaning some questions are not relevant or answered.

- Many of the technical features mentioned in the survey would probably be implementable if staffing allowed. As it is, despite being a repository manager much of my time is devoted to data entry and submission vetting. A community of support and good practice would be a great help.

- Talking to academics about authors versions and pre/post prints is an uphill struggle and the message is often lost on them. The green open access repository needs to be made a part of the living scholarly communication lifecycle. I tell authors that if they can't get APCs (which they all understand), then they can still meet their OA obligations to funders by taking the green route. Repositories need to somehow become a part of the scholarly communications lifecycle, rather than being a kind of graveyard of old articles.

# Conclusion

The Survey of Open Repositories in Ireland provides a comprehensive and revealing snapshot of the current state of open repositories in Ireland. The survey, undertaken with input from diverse educational, governmental, and research institutions and repository representatives, found key findings that highlight the strengths and challenges inherent in the existing landscape. This survey forms a pivotal part of a larger two-year initiative, supported by the National Open Research Forum 2022 Open Research Fund, aiming to enhance Ireland's network of open repositories. The focus on a standardised approach and the proposal to address metadata and support challenges at the repository level demonstrate a strategic emphasis on practical solutions without requiring extensive new technical infrastructure.

Concerns over visibility and staffing identified the need for greater coordination in training and metadata compliance collectively emphasise the urgency of addressing these critical aspects of repository provision and management via national coordination. Metadata alignment conforming to international standards is critical for repositories to support open research goals across Ireland. Fragmented metadata efforts hinder greater national coordination and monitoring of open research progress via monitoring mechanisms. Staff shortages underscore the need for trained, connected and dedicated personnel, indicating a need for concerted efforts to ensure efficient repository management at a national level.

The implications drawn from the survey findings underscore the diversity of efforts and resources among educational, governmental, and research institutions, revealing varying levels of sustainability and support for repositories. The recommendations, including collaborative initiatives, training programs, standardised metadata adoption, and regular assessments, serve as a first step toward the roadmap for advancing Irish open repositories in alignment with European peers. In moving forward, the identified trends and concerns provide valuable insights for shaping the future of open repositories in Ireland. The commitment to collaborative initiatives, training programs, and standardised metadata and technical practices will address the challenges highlighted in this report and contribute to the overall advancement and harmonisation of the Irish open repository network. By implementing these recommendations, Ireland can establish a more robust and interconnected infrastructure, effectively disseminating academic and research resources to benefit its scholarly community and beyond.

# Recommendations and Future Considerations

The data in this report suggests several steps that should be taken to improve concerns and issues in Irish open repositories.

**Collaborative Initiatives:** Encourage collaborative efforts among repositories, fostering knowledge exchange and resource sharing to enhance technical capabilities and expertise.

**Training and Support:** Provide training programs and technical support for repository managers, ensuring they have the necessary skills to manage and maintain repositories effectively.

**Standardisation:** Promote the adoption of standardised metadata schemas and persistent identifiers across repositories, enhancing interoperability and data discoverability.

**Continuous Assessment:** Conduct regular assessments to identify challenges repositories face and implement targeted interventions to address sustainability concerns, ensuring the long-term viability of open research initiatives.

# Appendix

# Potential Further Uses for the Data

1. **Managerial Role and Percentage Allocation:**

- How is the role of managers distributed among different functions, and what percentage of their managerial responsibilities does it constitute?
- How does this distribution vary across institutions and types, affecting compliance and other issues?
- To what extent does this align with the institution's stated remit?

2. **Institutional Characteristics:**

- Consider the type of institution, whether hosting is internal or external, and the percentage of Full-Time Equivalent (FTE).
- Examine the correlation between institution type, metadata schema, and Plan S/OpenAIRE compliance.
- Investigate the relationship between institution type and the support provided to the repository.

3. **Repository Hosting and Interoperability:**

- Explore how the type of institution and internal or external hosting affects interoperability.
- Investigate who can deposit content and analyse its impact on the repository's purpose and remit.
- Consider the relationship between institution type, Persistent Identifiers (PIDs), and author IDs.

4. **Metadata**

- Examine deposit workflows in relation to institution type.
- Analyse the differences between mediated and direct deposit and metadata types and local or external hosting. Does metadata validation correlate to either of these?
- Analyse metadata schemas and their appropriateness based on institution and hosting type.

- Investigate whether the hosting type and software platform impact metadata input/automation and preservation policies.
- Explore the reasons behind the absence of preservation policies in 14 out of 22 repositories.
- Assess the level of OpenAIRE compliance across repositories and determine if this is influenced by local or external hosting.

5. **Content Type and Platform Upgrades:**

- Investigate the correlation between content types and institutions.
- Examine institutions requiring platform upgrades and barriers to upgrades.
- Consider the role of the NORF-funded Open Access Monitoring Project in repository audits, particularly through metadata compliance and Plan S-related inquiries.

6. **Barriers to Metadata Compliance:**

- Analyse barriers to compliance, including funding, technical challenges, staffing, and time constraints.
- Examine the variation in barriers among different types of institutions and explore how their remit and role influence these barriers.
- Consider the impact of funding sources or parent bodies on compliance barriers.

7. **Institutional Areas of Concern & Importance:**

- Investigate the specific areas of concern for different institution types.
- Explore which areas are deemed essential to address concerns and how these priorities differ across institution types.
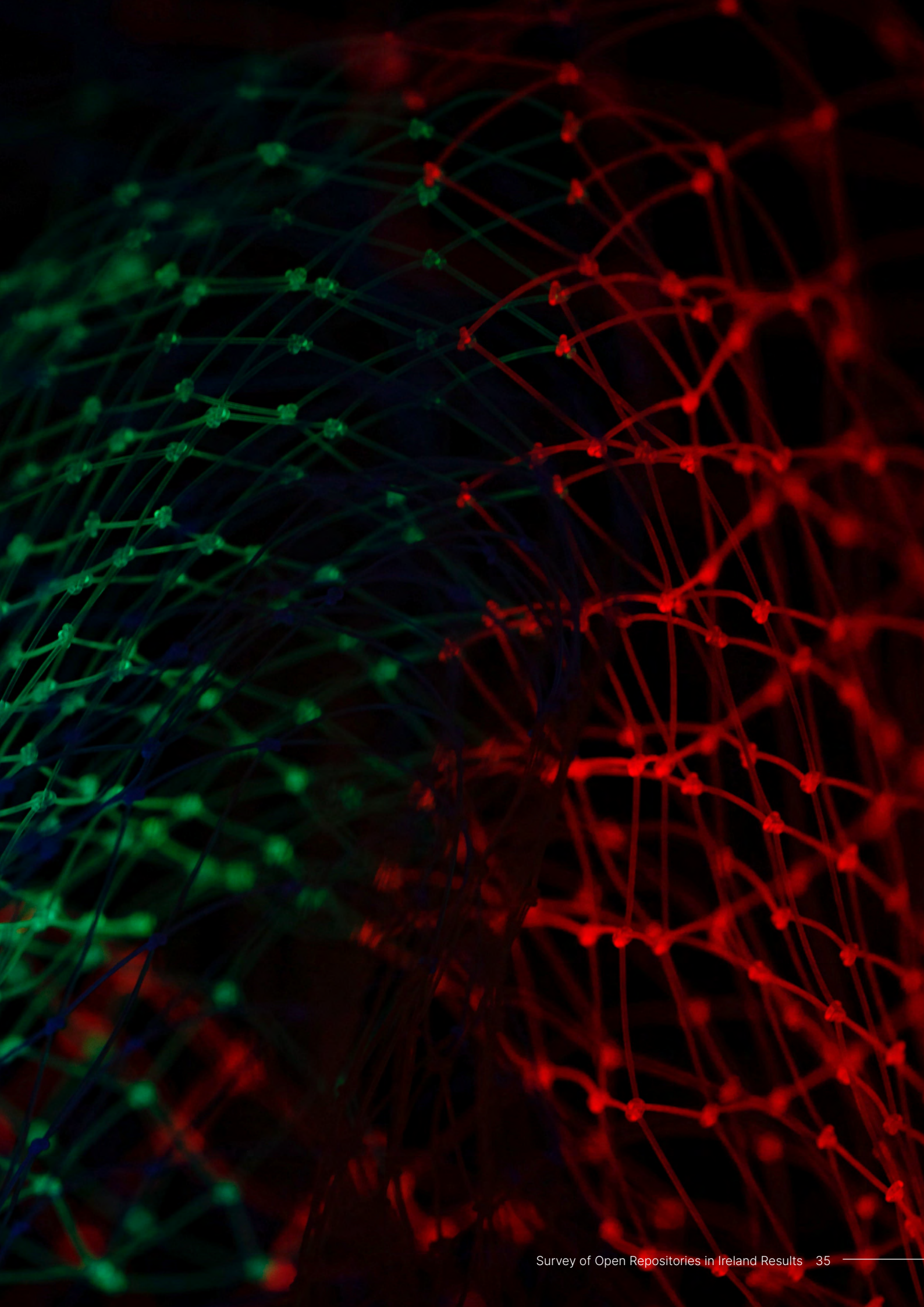
# Survey Questions

The complete list of survey questions can be found at: **Survey Questions PDF.**

# Survey Dataset

The anonymised survey dataset is published alongside the survey report.
**https://doi.org/10.5281/zenodo.10377235**