



National Open Access Monitor Advisory Group Meeting, 27th October 2023

Present:

- Advisory Group: Fran Callaghan (MU); Edie Davis (SFI); Caleb Derven (UL); Eoin Kenny (HEAnet); Kevin Kiely (TCD); Andrew Simpson (RCSI)
- OpenAIRE: Natalia Manola; Ioanna Grypari; Leonidas Pispiringas; Harry Dimitropoulos
- IReL: Susan Reilly (Chair), Catherine Ferris (Project Manager)

Susan formally welcomed Edie Davis, who is replacing Suz Garrard on the Advisory Group.

1. **Minutes** of meeting 18th September 2023:
<https://zenodo.org/doi/10.5281/zenodo.8405472>. Approved 4th October
2. **Matters Arising and Action Points:**

Matters Arising: none.

Action Points:

- a) **OpenAIRE:** to provide screenshots showing how deliverables meet tender specifications (focusing on those detailed in the comments to the draft Inception Report).
Action: OpenAIRE to provide for the next Advisory Group meeting.
- b) **OpenAIRE:** to deliver detailed documentation for RFOs and RPOs on what constitutes measures like findability, open access etc.
Action: OpenAIRE to provide for the next Advisory Group meeting.
- c) **OpenAIRE:** to deliver detailed instructions for researchers on all aspects of the ORCID functionality: claiming, how this impacts these research outputs on the Monitor, logging in with ORCID and logging in with other credentials etc.
Action: OpenAIRE to provide for the next Advisory Group meeting.
- d) **Advisory Group:** to define procedure for how dashboard managers will be assigned.

Procedure: Catherine to ask, on the Monitor mailing list, for people to volunteer as the dashboard manager for their organisation/entity, and Catherine to supply that list to OpenAIRE. Regarding consent: many people have chosen to be identified in their consent forms. For those contributors that have chosen not to be identified in their consent forms, Catherine will specifically ask for permission to share their name and contact details with OpenAIRE, and to share their name and contact details with others affiliated with their organisation/entity on request. This will ensure the process is open and transparent, with validation.

OpenAIRE advised that there is no deadline for receipt of list of dashboard managers; that dashboard managers will have access to OpenOrgs for the deduplication of organisation names, which will impact how an organisation is represented in the Monitor; and that it takes a month for changes in OpenOrgs to be reflected on the Monitor.

Action: Catherine to initiate the process and communicate the timeline to impact the draft Monitor on launch in January.

e) **Catherine:** to provide list of RPOs and RFOs to OpenAIRE

Draft list (Supplement A) sent on 4th October 2023. Stakeholder Survey: Organisational Identity currently open to validate/update/correct this list.

Action: Catherine to provide the results of the stakeholder survey with the final confirmed list to OpenAIRE and Zenodo on 10th November.

f) **Catherine:** to provide OpenAIRE with a list of stakeholder PIDS to enable OpenAIRE to query the OpenAIRE graph and identify Irish research outputs.

Stakeholder Survey: Organisational Identity currently open to capture this information.

Action: Catherine to provide the results of the stakeholder survey with the final confirmed list to OpenAIRE and Zenodo on 10th November.

g) **Catherine:** to provide OpenAire with examples of peer-reviewed journal articles with DataCite DOIs

Example file of peer-reviewed conference proceedings with DataCite DOIs collated by Kevin sent on 25th September (Supplement B).

OpenAire confirmed that the list of conference proceedings is sufficient (and not journal articles specifically).

3. **Review OpenAIRE's draft of the *National Open Access Report, Ireland***, provided by OpenAIRE on 9th October (Appendix 1). Outcome of the meeting is to provide feedback to OpenAIRE on any areas in the Report, in particular (if applicable) areas which do not meet the Advisory Group's expectation of the NORF and tender requirements, and to answer any questions OpenAIRE have at this stage.

- Ioanna: presentation on draft *National Open Access Report, Ireland* by OpenAIRE (Appendix 2)
- Discussion

To note: This Report is a key deliverable of the National Open Access Monitor Project. This is the first draft, with the final draft expected at month 6. The Report will also go out to stakeholders, to invite their feedback, once the Advisory Group has completed the first review.

Action: OpenAIRE to update the draft *National Open Access Report, Ireland* by 15th November and provide to Catherine for distribution.

The Advisory Group agreed:

- The draft Minutes will be supplied by Tuesday 31st October; Advisory Group and OpenAIRE to review, update and approve by 7th November; no response will be taken as approval of the minutes.
- The draft *National Open Access Report, Ireland*, annotated with comments during the review process, is to be included as Appendix 1 to these Minutes.
- The OpenAIRE presentation is to be included as Appendix 2 to these Minutes.
- That OpenAIRE make the following updates to the Report and the Monitor:

- The domain representation should allow for better visibility in the underrepresented domains.
 - DataCite DOIs should be accommodated in the future, as many repositories assign DataCite DOIs.
 - Diamond OA to be presented in the Report and the Monitor under “gold”, not as an “APC business model”.
 - Bronze should be captured and tracked in the Report and the Monitor, but not identified as “open access” as it does not align with the BOAI definition of open access.
 - That publishing in hybrid journals under transformative agreements meets Plan S compliance and should be represented in the Plan S section of the Report.
 - HRB and IRC have not endorsed Plan S and should not be included in the Plan S section of the Report.
 - That the Report should acknowledge Research Classification Ireland <https://www.gov.ie/en/policy-information/e1a41-research-classification-ireland/> in the domain FOS section.
 - That the Report include guidance for stakeholders on how they can enable the Monitor to report on “immediate” open access (specific metadata requirements etc).
 - All other comments and questions in the annotated draft *National Open Access Report, Ireland* document (Appendix 1) are approved to be actioned by OpenAIRE.
- To review how green is captured and represented in the Report and the Monitor. Green under the Unpaywall definition does not distinguish between green with an open license or green without an open license (both are “green OA”). The BOAI definition infers the requirement of a license. For review: if green should be identified as “open access” if it does not have an open license. Advisory Group confirmed that it is common practice in Irish repositories to assign a default open license to items deposited in the repository.
- Action:** OpenAIRE to prepare mockups of visualisations for what it would look like to (a) not identify green as “open access” without an open license and (b) identifying all green as “open access”, but distinguishing between green with open license and green without open license.
- Action:** Advisory Group to give feedback on the OpenAIRE mockups and agree on how green will be represented in the Report and Monitor.



Commented [CF1]: OpenAIRE have identified some formatting issues in this word version of the document, which we're using to capture comments. Please refer to the PDF if required: [National Open Access Monitor Draft Report.Updated121023.pdf](#)

Commented [CF2]: This document was updated on 12th October and supersedes the document shared on Tuesday.

National Open Access Monitor, Ireland

Draft Report

First Insights and Path Forward

October 9, 2023



Executive Summary

This draft report provides an initial assessment of Ireland's open access (OA) landscape, with a primary focus on the Baseline Analysis and Data Evaluation. It offers valuable insights into the current state of OA in the country, emphasizing data quality and coverage, while also explaining the construction and methodological steps taken. In addition to these assessments, this draft also outlines the next steps for the OpenAIRE team and the optional strategies to be followed by key stakeholders to ensure accurate monitoring.

In the final version of the report, we will enhance this foundation by comprehensively addressing high-level infrastructure challenges. Our primary goal will be to provide practical, actionable recommendations for achieving continuous and robust OA monitoring in Ireland. The final report will offer a more extensive analysis, drawing upon additional insights gathered through further investigations.

Contents

Executive Summary	2
Contents	3
List of Figures	4
List of Tables	6
Abbreviations	7
Glossary	9
1 Introduction	11
2 Baseline Analysis	11
2.1 Scientific Production	11
2.2 Open & FAIR	17
2.2.1 Access rights	18
2.2.2 Open Access routes	23
2.2.3 Unrealised & Immediate Open Access	26
2.2.4 FAIRness	28
2.3 Plan S	33
3 Methodology	42
3.1 OpenAIRE Graph: Foundation of the Monitor	42
3.2 The Publication Set of the Monitor	43
3.3 Data Disambiguation Techniques	49
3.4 Enrichment via Text Mining	51
3.5 Indicators	52
3.6 Additional Aspects	54
4 Data Evaluation	56
5 Conclusion	62
References	63

List of Figures

Figure 1: Irish Peer Reviewed Publications over time	13
Figure 2: Irish Peer Reviewed Publications by FoS (level 1)	14
Figure 3: Irish Peer Reviewed Publications by FoS (levels 1 & 2)	14
Figure 4: Irish Peer Reviewed Publications by hosted data source type	16
Figure 5: Irish Peer Reviewed Publications by SDGs	16
Figure 6: Irish Peer Reviewed Publications by Access Rights over time	19
Figure 7: Irish Peer Reviewed Publications by Access Rights by FoS (level 1)	19
Figure 8: Irish Peer Reviewed Publications by Access Rights by FoS (level 2)	20
Figure 9: Irish Peer Reviewed Publications by Access Rights by Data Source Type	22
Figure 10: Irish Peer Reviewed Publications by Access Rights by SDGs	22
Figure 11: Repository mediated vs publisher mediated OA Irish Peer Reviewed Publications over time	24
Figure 12: Repository mediated vs publisher mediated OA Irish Peer Reviewed Publications by FoS (level 1)	25
Figure 13: Repository mediated vs publisher mediated OA Irish Peer Reviewed Publications by FoS (level 2)	25
Figure 14: Unrealised Open Access Irish Peer-reviewed Publications over time	27
Figure 15: Unrealised Open Access Irish Peer-reviewed Publications by FoS (level 1)	27
Figure 16: Unrealised Open Access Irish Peer-reviewed Publications by FoS (level 2)	28
Figure 17: Irish Open Access Peer Reviewed Publications with a license over time	29
Figure 18: Irish Open Access Peer Reviewed Publications with a license by FoS (level 1)	29
Figure 19: Irish Open Access Peer Reviewed Publications with a license by FoS (level 2)	30
Figure 20: Irish Open Access Peer Reviewed Publications with a license by Publication Type	31
Figure 21: Irish Open Access Peer Reviewed Publications with a license by Data Source Type	31
Figure 22: Irish Open Access Peer Reviewed Publications by journal business model, over time	34
Figure 23: Irish Open Access Peer Reviewed Publications by journal business model, by FoS (level 1)	35
Figure 24: Irish Open Access Peer Reviewed Publications by journal business model, by FoS (level 2)	35
Figure 25: Total APCs (EUR) of Irish OA Peer Reviewed Publications over time	36
Figure 26: Average APCs per publication (EUR) of Irish OA Peer Reviewed Publications over time	37

Figure 27: Top 20 Journals by Total APCs (EUR) of Irish OA Peer Reviewed Publications	38
Figure 28: Top 20 Journals by Average APCs per publication (EUR) of Irish OA Peer Reviewed Publications	39
Figure 29: Total APCs (EUR) of Irish OA Peer Reviewed Publications by FoS (level 1)	40
Figure 30: Average APCs per publication (EUR) of Irish OA Peer Reviewed Publications by FoS (level 1)	40
Figure 31: The Monitor and the OpenAIRE Graph Pipeline	43

List of Tables

Table 1: Outline of overall production of Irish publications	12
Table 2: Irish Peer Reviewed Publications by type	15
Table 3: Number and share of Irish Peer Reviewed Publications by Access Rights	18
Table 4: Number of Irish Peer Reviewed Publications by Access Rights by Publication Type	20
Table 5: Number and share of Repository mediated vs publisher mediated OA Irish Peer Reviewed Publications	23
Table 6: Number and share of All vs Immediate vs Unrealised Open Access Irish Peer-reviewed Publications	26
Table 7: Number and share of Open Access Irish Peer-reviewed Publications with a CC license.....	28
Table 8: Licence normalisation	32
Table 9: Number and share of Irish Peer-reviewed Publications with a PiD	33
Table 10: Irish Open Access Peer Reviewed publications by Journal Business Model	34
Table 11: Total and average APCs of # Irish OA peer-reviewed publications	36
Table 12: Number and share of Grant supported Irish Peer Reviewed publications by a Plan S Funder	41
Table 13: Number of Grant supported Irish Peer Reviewed publications by Open Access Route	41
Table 14: Share of Grant supported Irish Peer Reviewed publications by Open Access Route	42
Table 15: RPOs & their publications in the Monitor.....	44
Table 16: Irish Institutional Data Sources & Alignment with OpenAIRE	44
Table 17: Irish Funders that have joined OpenAIRE.....	46
Table 18: General Representation of Irish Funders in OpenAIRE (excluding SFI)	46
Table 19: Irish Peer Reviewed Publications	47
Table 20: Irish publications with DOI not assigned by Crossref and NOT included in current set.....	47
Table 21: Mean and standard deviation in # of references by FoS (level 1).....	47
Table 22: Irish peer-reviewed publications by PID type	48
Table 23: Licence normalisation	50
Table 24: Construction methodology of indicators.....	52
Table 25: Metadata Analysis.....	56

Abbreviations

APC	Article Processing Charge
CC	Creative Commons
DOI	Digital Object Identifier
EC	European Commission
EOSC	European Open Science Cloud
FAIR	Findable, Accessible, Interoperable, Reusable
FoS	FoS
Graph	OpenAIRE Graph
IReL	The consortium of Irish research libraries
Monitor	National Open Access Monitor, Ireland
MU	Maynooth University
NLP	Natural Language Processing
NORF	National Open Research Forum
OA	Open Access
OFR	Open Funder Registry
OS	Open Science
ORCID	Open Researcher and Contributor ID
PID	Persistent Identifier
RFO	Research Funding Organisation
RPO	Research Performing Organisation
SDG	Sustainable Development Goal
SFI	Science Foundation Ireland
VoR	Version of Record
ROR	Research Organization Registry
ISNI	International Standard Name Identifier



GRID Global Research Identifier Database

Glossary

TERM	DEFINITION
ARTICLE PROCESSING CHARGE (APC)	The fee charged by publishers in order to publish a research publication in an open access journal. These charges are meant to cover the costs of publication and ensure the work is freely accessible to all.
RESEARCH OUTPUTS/PRODUCTS	The four different types of research products in the OpenAIRE Graph: Publications, Research data, Research software, Other research products.
OPEN ACCESS	We use the Budapest Open Access Initiative definition of "open access": "By "open access" to this literature, we mean its free availability on the public internet, permitting any users to read, download, copy, distribute, print, search, or link to the full texts of these articles, crawl them for indexing, pass them as data to software, or use them for any other lawful purpose, without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself." ¹
JOURNAL BUSINESS MODELS	
OA (GOLD)	A journal that publishes only in OA.
SUBSCRIPTION	A journal that charges for access to its articles.
HYBRID	A subscription journal where some of its articles are open access.
TRANSFORMATIVE	<i>A Transformative Journal is a subscription/hybrid journal that is actively committed to transitioning to a fully Open Access journal. In addition, a Transformative Journal must gradually increase the share of Open Access content; and offset subscription income from payments for publishing services (to avoid double payments).</i> ²

¹ <https://www.budapestopenaccessinitiative.org/read/>

² <https://www.coalition-s.org/transformative-journals-faq/>

JOURNAL APC BUSINESS MODELS

DIAMOND OA	A fully OA journal that does not charge article processing charges (APCs).
-------------------	--

OA TYPES/COLOURS³

GREEN OA	Green articles are published in toll-access journals, but archived in an OA archive, or "repository". These repositories may be discipline-specific (like ArXiv) or institutional repositories operated by universities or other institutions. Green articles may be published versions or preprints, and can have any licence or no licence.
-----------------	---

BRONZE OA	Bronze articles are free to read on the publisher's website, without a licence that grants any other rights. There may be a delay between publication and availability to read, and often articles can be removed unilaterally by the publisher.
------------------	--

HYBRID OA	Hybrid articles are free to read at the time of publication, with an open license. These are usually published in exchange for an article processing charge, or APC.
------------------	--

GOLD OA	Gold articles have all the same characteristics as Hybrid articles, but are published in all-Open Access journals, which are in turn called "Gold journals", or just "OA journals".
----------------	---

ACCESSIBILITY - INTEROPERABILITY

ACCESSIBLE	A publication is accessible if the text file can be fetched via a valid URL in its metadata. ⁴
-------------------	---

INTEROPERABLE	A publication is considered interoperable if its full-text is in a machine-readable format, allowing machines to process and understand the content.
----------------------	--

³<https://support.unpaywall.org/support/solutions/articles/44001777288-what-do-the-types-of-oa-status-green-gold-hybrid-and-bronze-mean>

⁴ I.e., if a publication does not include a valid URL, we cannot assess accessibility.

Commented [CF3]: For Advisory Group: I propose that Diamond is moved up under Journal Business Model. Listing Diamond under an "APC" Business Model when by definition there's no APC doesn't align. Do you agree?

Commented [CF4]: For Advisory Group: to my reading, how these types/colours are treated in this report needs further clarity.

The key priority is that "open access", as per the BOAI definition means that an open license must be attached. The NORF National Action Plan states "Open access publications must be accompanied by an open licence."

I propose that only research outputs with open licenses should be considered as "open access" in this report (and the Monitor) to align with these definitions.

While Unpaywall definitions of green + bronze do not need licenses to be categorised as green/bronze by Unpaywall, I think it's important that when the Report (and Monitor) refers to green OA, it's referring to green OA with license (or at least is specific in calling out green OA without a license), and that Bronze OA is not considered as "open access" for the purposes of the Report (and the Monitor).

Do you agree?

Commented [CF5R4]: @edie.davis@sfi.ie can you advise - from the SFI point of view, for a researcher to be OA compliant with SFI policy by the green route (depositing in a repository), is there an expectation that the research output is accompanied by an open license in the repository? Would depositing in an repository, but not assigning an open license, be compliant with the SFI OA policy?

Commented [ed6R4]: Yes, an open license is expected. The SFI OA policy is Plan S-aligned, so publications must be available under a CC-BY license (unless an exception has been agreed to allow another license).

1 Introduction

The National Open Access (OA) Monitor, Ireland stands as a pivotal initiative to comprehensively assess the state of open access research and scholarly publishing within the country. This report has been developed specifically to accompany the launch of the monitor, offering stakeholders an in-depth understanding of the methodologies employed, the data evaluated, and the preliminary findings gleaned.

Our primary aim is twofold: to present a thorough baseline analysis of the current Open Access (OA) landscape in Ireland, and to articulate the data evaluation processes we've instituted. While this report offers a comprehensive initial assessment, the subsequent final report will delve into more extensive evaluations, addressing infrastructural challenges and drawing from broader insights.

2 Baseline Analysis

This section presents OpenAIRE's initial baseline analysis of Open Access (OA) in Ireland. The analysis encompasses an examination of Ireland's present publishing output, both from an overarching perspective and at a domain-specific level. The FoS classification system⁵, as defined by OpenAIRE, serves as the basis for this domain-level assessment, categorizing publications into distinct domains⁶.

The "Scientific Production" section provides contextual information, detailing the volume and nature of research outputs. In the "Open and Fair" section, the accessibility and fairness of disseminated research are explored, assessing the breadth of availability of research outputs. The "Plan S" section is dedicated to the evaluation of how Irish research publications align with the guidelines established by cOAlition S⁷.

While this analysis offers a foundational overview, subsequent refinements to the data are anticipated. Brief indications of the measures for these refinements are provided, with a comprehensive evaluation set forth in the Data Evaluation section.

2.1 Scientific Production

This section serves as a foundational exploration of scientific production in Ireland, providing crucial context for the subsequent analysis. It offers an insightful overview of the

⁵ <https://explore.openaire.eu/fields-of-science>

⁶ See Section 3.4 for a description.

⁷ <https://www.coalition-s.org/>

Commented [CF7]: For Advisory Group: propose changing "Scientific" here (and throughout) to Scholarship or Research. As per feedback on the Stakeholder Webinar, the use of Science/Scientific infers a STEM focus that is not intended and implies that AHSS is not prioritised. Do you agree?

Commented [ed8R7]: Agree

Commented [an9R7]: Yes agree

Commented [le10R7]: OK

Commented [Ca11R7]: Agree

country's peer-reviewed publications, including both their absolute numbers and a granular breakdown across various dimensions. By examining the distribution of publications over time, scientific disciplines, publication types, and data source types, we gain a comprehensive understanding of Ireland's research landscape. This comprehensive analysis lays the groundwork for a more in-depth investigation into the state of OA within the Irish scholarly community, ensuring that the subsequent sections are firmly rooted in the broader scientific context.

Table 1: Outline of overall production of Irish publications

# Irish publications	345,316
# and share of Irish Peer Reviewed Publications	244,156 (70.7%)
# and share of Irish Open Access publications	226,025 (65.45%)
# and share of FoS coverage	78,436 (22.71%)

Commented [CF12]: Please provide time period

Commented [le13R12]: No constraints applied.

Commented [CF14]: This table uses both commas and periods as decimal points. Request use of periods as decimal points throughout the Report.

Commented [le15R14]: We will update it as it is in the whole report with period as decimal point.

We note that a substantial 70% of Ireland's research output consists of peer-reviewed publications, indicating a high level of academic and scientific quality. We anticipate achieving even more favorable outcomes following enhancements to the methodology for identifying peer-reviewed publications.

To enhance the coverage of FoS, we are presently implementing our FoS classifiers on Irish Publications, aiming for a more comprehensive representation in the National Open Access Monitor of Ireland.

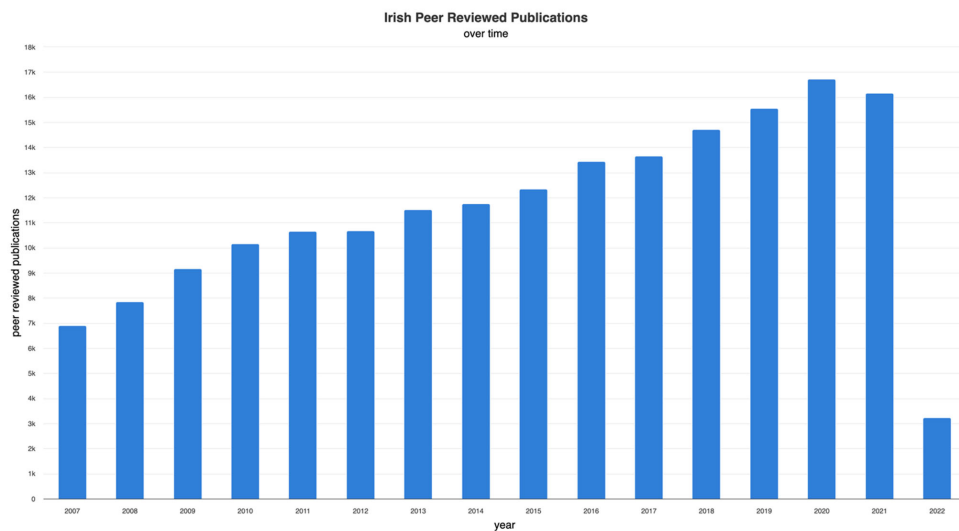


Figure 1: Irish Peer Reviewed Publications over time

In this chart, we observe a consistent, gradual upward trend in peer-reviewed Irish publications over time. To delve deeper into this trend, the following two charts provide a detailed breakdown by Field of Science domains (FoS level 1) and topics (FoS level 2). This finer-grained perspective can shed light on the disciplines that boast larger numbers of peer-reviewed publications.

The delay in the availability of 2022 Irish peer-reviewed publications, as of October 2023, can be attributed to the natural timeline associated with academic publishing. OpenAIRE is actively addressing this delay by focusing on two main strategies: refining data harvesting from Irish institutional sources, enhancing text mining for affiliations specifically for Ireland and reviewing the methodology for establishing peer-review status.

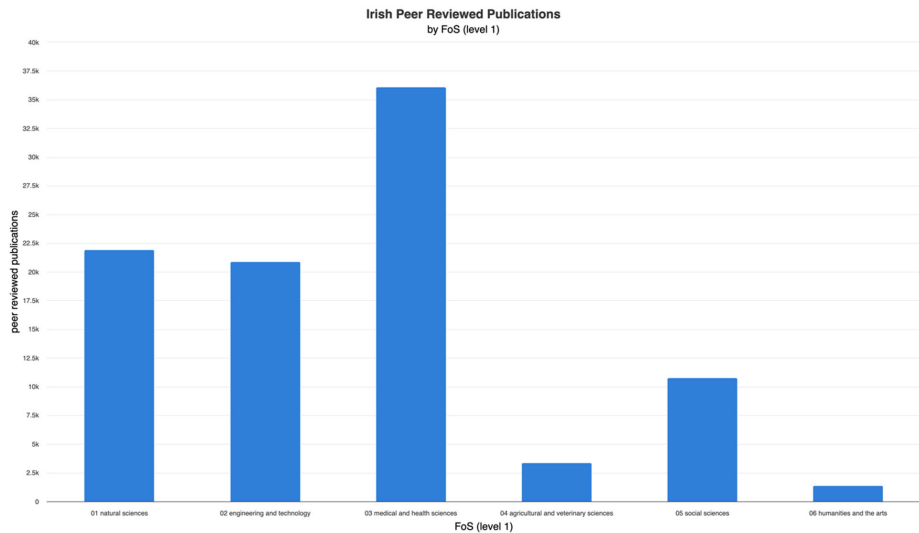


Figure 2: Irish Peer Reviewed Publications by FoS (level 1)

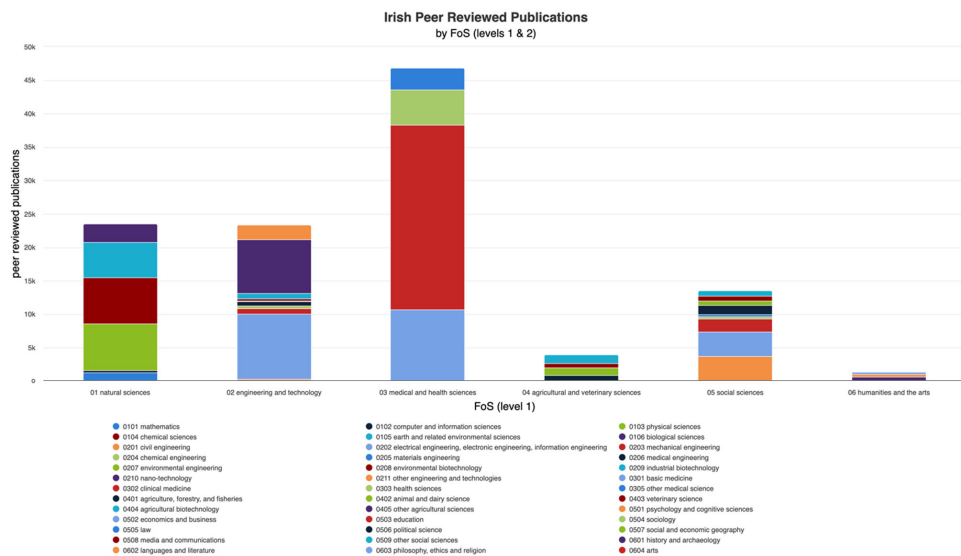


Figure 3: Irish Peer Reviewed Publications by FoS (levels 1 & 2)

These charts present the count of Irish Peer-Reviewed Publications categorized by Field of Science (FoS) domain (level 1) and topic (level 2). It becomes apparent that the Medical and Health Sciences domain, along with its Clinical Medicine and Basic Medicine topics, commands the highest numbers.

To enhance the coverage of FoS, we are presently implementing our FoS classifiers on Irish Publications, aiming for a more comprehensive representation in the National Open Access Monitor of Ireland.

Table 2: Irish Peer Reviewed Publications by type

	# Irish peer-reviewed publications	Share of # Irish peer-reviewed publications
Article	202,997	82,27%
Book	461	1,05%
Conference object	22,218	51,33%
Contribution for newspaper or weekly magazine	1	0,00%
Dataset	1	0,00%
Part of book or chapter of book	18,931	7,75%
Unknown	22	0,00%

Commented [CF16]: Is this a typo? 51.33% seems too high

Commented [le17R16]: Yes, it 9.1%.

Within this table, it is evident that the predominant share of peer-reviewed publications comprises articles, with conference papers and books/book chapters following in sequence.

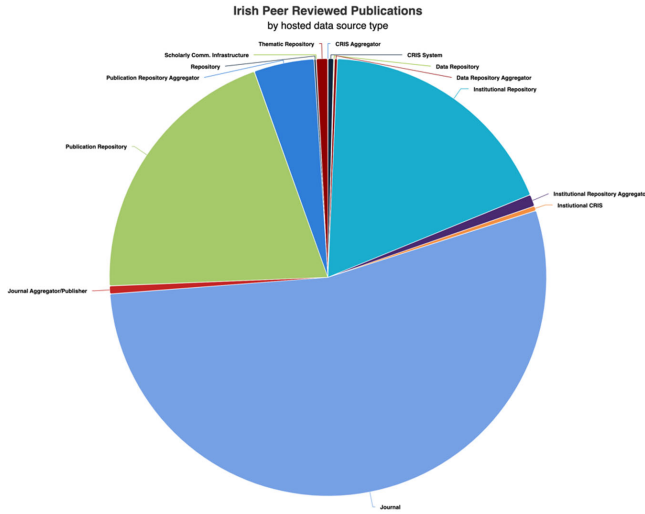


Figure 4: Irish Peer Reviewed Publications by hosted data source type

According to this chart, Journals serve as the primary data source for hosting the majority of peer-reviewed publications, with Institutional and Publications Repositories coming in as the subsequent sources in line.

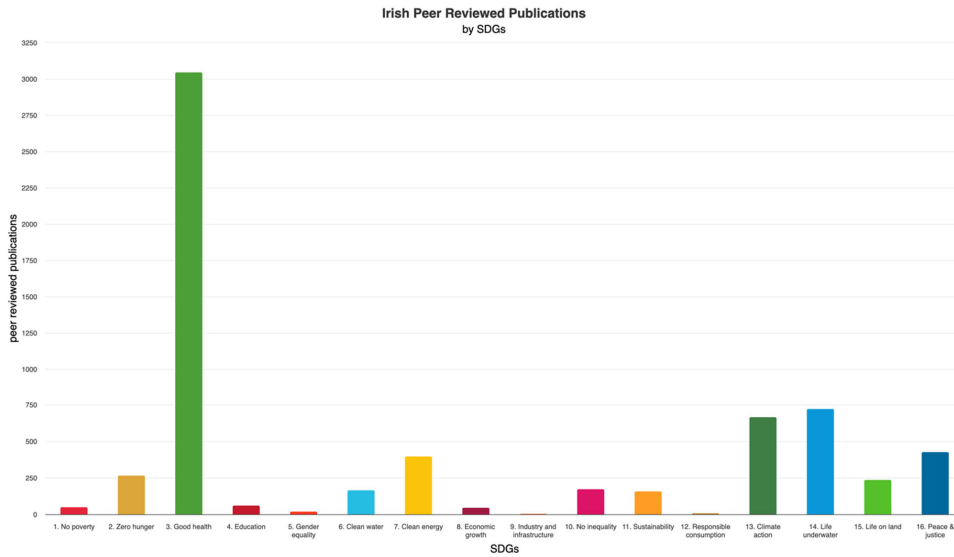


Figure 5: Irish Peer Reviewed Publications by SDGs

In the chart displaying the breakdown by Sustainable Development Goals (SDGs), one can discern the societal impact of Irish peer-reviewed publications, with "Good Health" occupying the leading position, followed by "Climate Action" and "Life Underwater."

To enhance the SDGs coverage of FoS, we are presently implementing our SDG classifiers on Irish Publications, aiming for a more comprehensive representation in the National Open Access Monitor of Ireland.

2.2 Open & FAIR

In this section, we delve into a set of essential indicators that illuminate the openness and fairness of publications within Ireland's Open Access (OA) landscape. These indicators encompass various dimensions, including access rights, OA routes, embargoed access, licensing, persistent identifiers (PIDs), and the accessibility and interoperability of OA peer-reviewed publications.

By analysing these indicators, we gain valuable insights into the extent to which scholarly outputs are openly accessible and align with principles of fairness and transparency. This analysis forms a crucial foundation for understanding the current state of OA in Ireland and identifying areas where improvements and interventions may be necessary. These indicators offer a comprehensive view of the multifaceted dimensions of openness and fairness in scholarly communication across domains, topics, publication types, and data source types.

OpenAIRE's PDF aggregation system has meticulously reviewed approximately 44K URLs extracted from the metadata of the 127K Irish OA peer-reviewed publications. Among these URLs, we have successfully retrieved 16K PDFs (~12,1%). Our ongoing efforts involve inspecting the remaining 28K URLs to further enhance our coverage.

At this juncture, there is no necessity to proceed either with the utilization of the Metadata Validator to evaluate the completeness and FAIRness of metadata within records from repositories and OA Journals or with the OpenAIRE Broker service of enriching the original data sources. This is because the outcomes obtained would not accurately reflect the status of the records under examination. To ensure a more dependable assessment, we will revisit the implementation when we have a greater number of data sources (including repositories, CRIS, and OA Journals) registered in OpenAIRE and in compliance with the latest versions of the OpenAIRE Guidelines (v.3 and above). This decision will have no impact on the timelines/deliverables of the Monitor's release or the overall data quality within the Monitor. [the OpenAIRE Broker](#), to address enrichment of original data sources.

Commented [CF18]: Is this a typo?

Commented [le19R18]: Yes, deleted.

2.2.1 Access rights

Table 3: Number and share of Irish Peer Reviewed Publications by Access Rights

Access Rights	# Irish peer-reviewed publications	Share of # Irish peer-reviewed publications
Open Access	127,221	52,11%
Embargo	26	0,01%
Restricted	890	0,36%
Closed Access	70,534	28,89%
not available	45,485	18,63%

The present chart illustrates the distribution of Irish Peer-Reviewed Publications based on their access rights.

To delve deeper into this trend, the following chart provides a chronological breakdown. We can observe an upward trend in the uptake of open access, with an increase of over 20% from 2014 to 2021.

The delay in the availability of 2022 Irish peer-reviewed publications, as of October 2023, can be attributed to the natural timeline associated with academic publishing. OpenAIRE is actively addressing this delay by focusing on two main strategies: refining data harvesting from Irish institutional sources, enhancing text mining for affiliations specifically for Ireland and reviewing the methodology for establishing peer-review status.

Commented [CF20]: For Advisory Group: do you think that restricted and not available are meaningful in this Report? Or would it be better to include restricted and not available as closed access?

Commented [ed21R20]: I think Restricted is a meaningful category, especially for repository managers; restricted access can be time-limited in repositories.

Commented [CF22R20]: What's the difference in meaning between restricted and embargoed?

Commented [le23R20]: https://vocabularies.coar-repositories.org/documentation/access_rights/

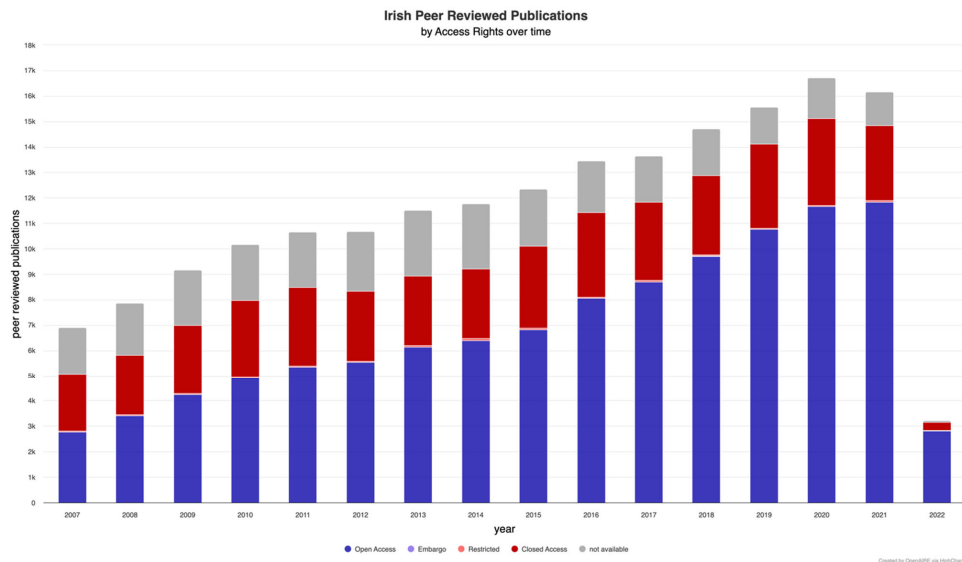


Figure 6: Irish Peer Reviewed Publications by Access Rights over time

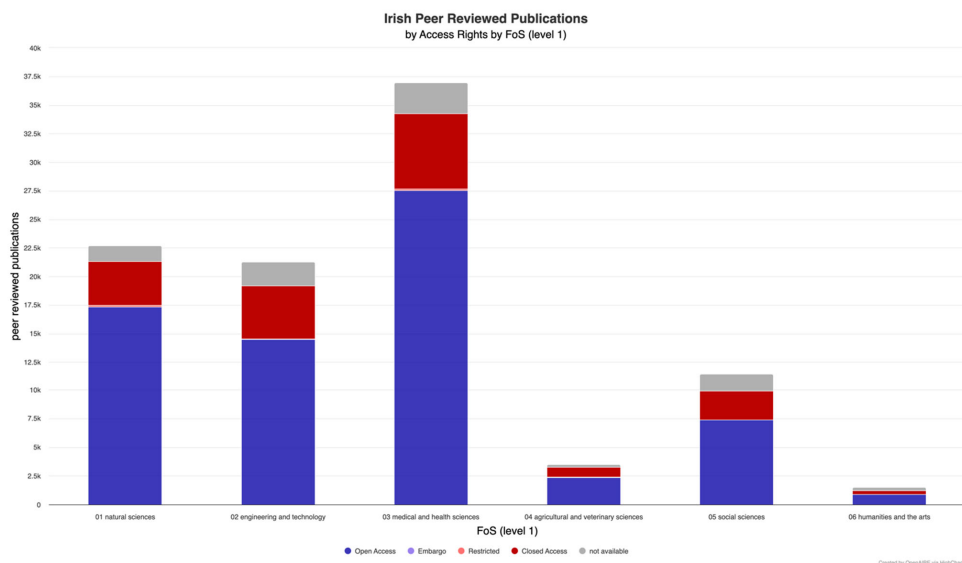


Figure 7: Irish Peer Reviewed Publications by Access Rights by FoS (level 1)

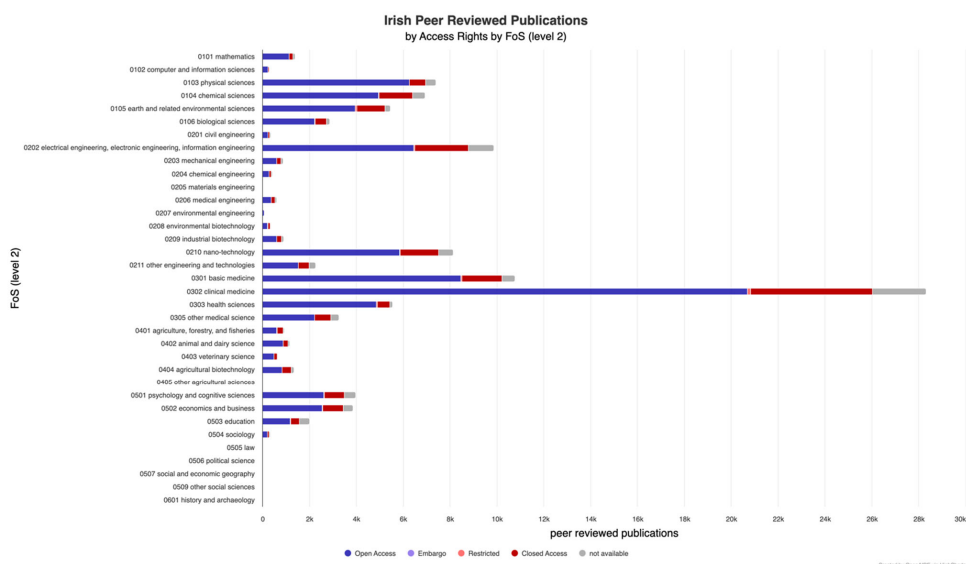


Figure 8: Irish Peer Reviewed Publications by Access Rights by FoS (level 2)

To delve deeper into this trend, these two charts provide a detailed breakdown according to Field of Science domains (FoS level 1) and topics (FoS level 2). This finer-grained perspective can shed light on particular disciplines that might be influencing the overall decline. It offers valuable insights for shaping future OA strategies, as exemplified by the notable increase in Open Access peer-reviewed publications within the Medical and Health Sciences domain, along with the Clinical Medicine topic.

To enhance the coverage of FoS, we are presently implementing our FoS classifiers on Irish Publications, aiming for a more comprehensive representation in the National Open Access Monitor of Ireland.

Table 4: Number of Irish Peer Reviewed Publications by Access Rights by Publication Type

	Open Access	Embargo	Restricted	Closed Access	not available
Article	114,943	21	804	59,774	27,455
Book	145		5	75	236

Conference object	9,052		31	1562	11,573
Contribution for newspaper or weekly magazine	1				
Dataset	1				
Other ORP type	3				
Other literature type	1,742	1	15	327	21
Part of book or chapter of book	3,848	5	50	8,827	6,201
Preprint	4				
Thesis	2				3
Unknown	13				

In this table, it's clear that the majority of Open Access peer-reviewed publications are articles, with conference papers and books/book chapters being the subsequent categories in order.

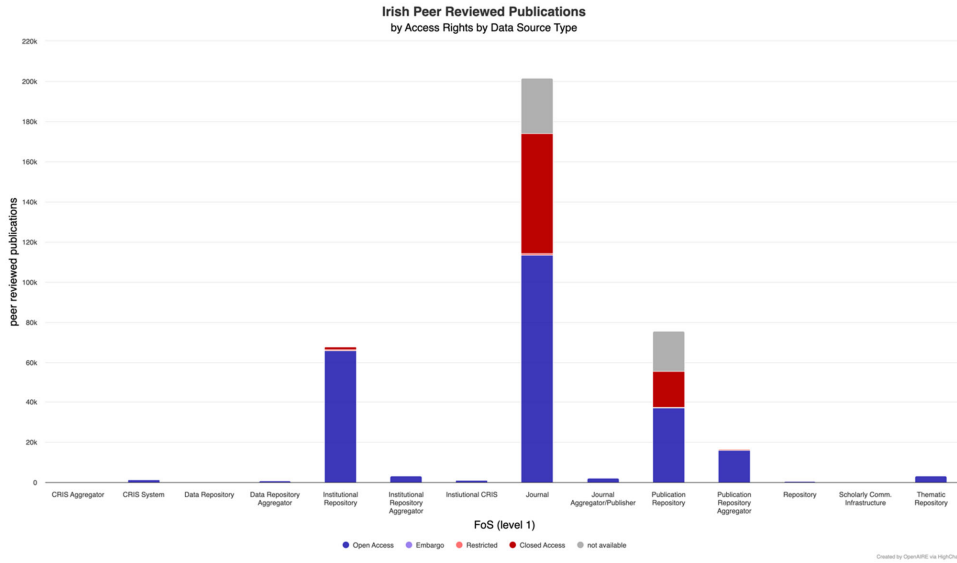


Figure 9: Irish Peer Reviewed Publications by Access Rights by Data Source Type

Commented [CF24]: What does "restricted" and "not available" mean here in the context of journals?

Commented [le25R24]: The access rights shown are those of the publications as merged from multiple data sources. NA means that we have a record found from a journal without its access rights shared.

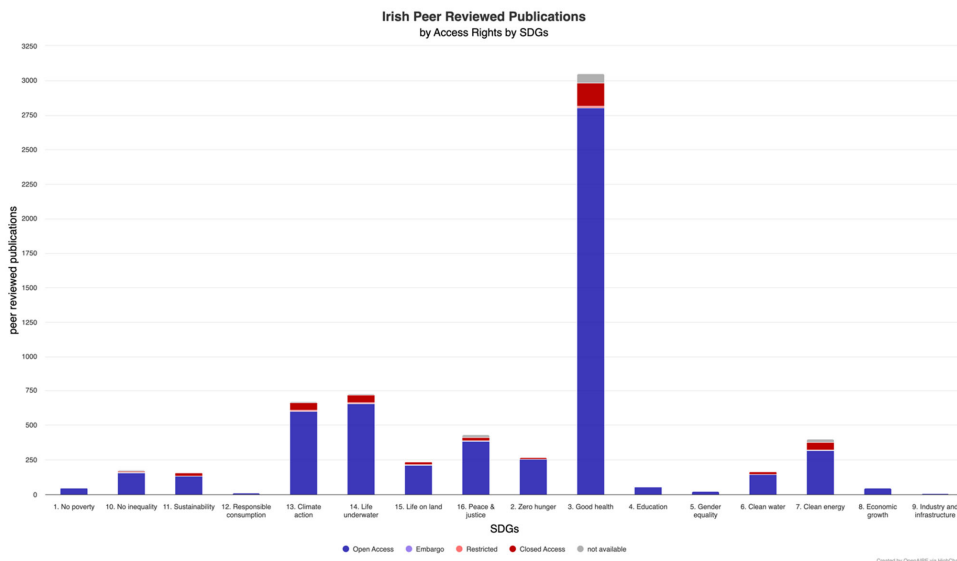


Figure 10: Irish Peer Reviewed Publications by Access Rights by SDGs

As indicated by the two charts presented above, Journals emerge as the foremost data source for hosting the bulk of Open Access peer-reviewed publications, closely followed by Institutional and Publications Repositories. Meanwhile, in the chart depicting the distribution by Sustainable Development Goals (SDGs), one can discern the societal influence of Irish Open Access peer-reviewed publications. Notably, "Good Health" claims the top spot, with "Climate Action" and "Life Underwater" following suit in terms of impact.

To enhance the coverage of SDGs, we are presently implementing our SDG classifiers on Irish Publications, aiming for a more comprehensive representation in the National Open Access Monitor of Ireland.

2.2.2 Open Access routes

Table 5: Number and share of Repository mediated vs publisher mediated OA Irish Peer Reviewed Publications

Open access route	# Irish OA peer-reviewed publications	Share of # Irish OA peer-reviewed publications
Repository mediated OA (Green)	28,277	22,31%
Publisher mediated OA (Gold)	34,592	27,30%
Publisher mediated OA (Hybrid)	19,966	15,76%
Publisher mediated OA (Bronze)	10,880	8,59%

As delineated in Table 4, which provides a detailed breakdown of open access, distinguishing Repository-mediated OA (Green) from Publisher-mediated OA (Gold, Hybrid, Bronze), we note that approximately one-quarter of Irish Open Access publications fall under Repository-mediated Open Access, while nearly three-quarters are categorized as Publisher-mediated Open Access.

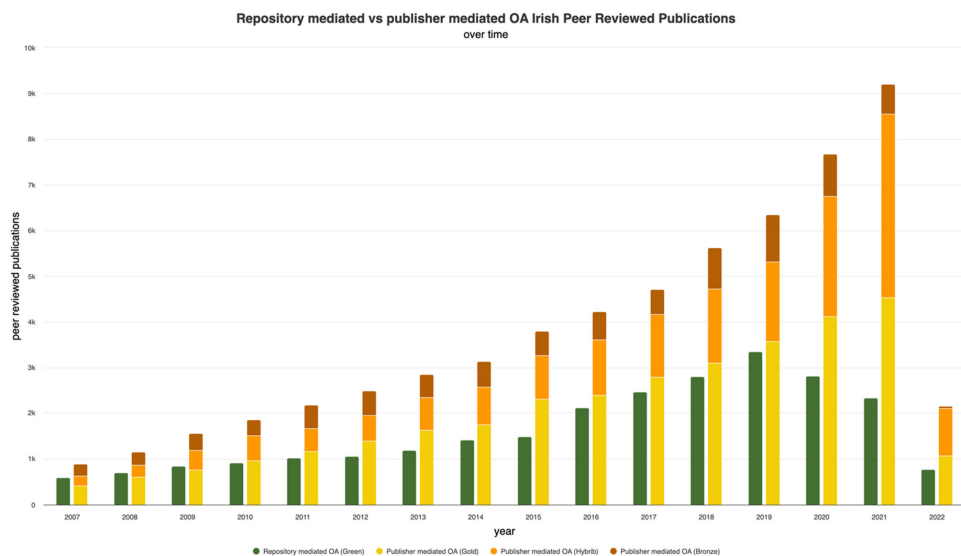


Figure 11: Repository mediated vs publisher mediated OA Irish Peer Reviewed Publications over time

Commented [CF26]: Can this data be presented as a table too please? Over time? The IReL OA Agreements began in 2020 and it's not evident from how the data is presented here the impact that they've had.

In the year-wise distribution chart that follows, we also discern a rising trend in both Repository-mediated and Publisher-mediated Open Access. Notably, Publisher-Mediated Open Access has increased by over 50% from 2014 to 2021. This trend aligns with the direction of the Irish research landscape towards greater openness and underscores the effectiveness of mechanisms established by the global Open Science movement and related initiatives.

To delve deeper into this trend, the following charts provide a detailed breakdown by Field of Science (FoS). This finer-grained perspective can shed light on particular disciplines that might be contributing to the overall rise in Open Access peer-reviewed publications, offering valuable insights for shaping future OA strategies.

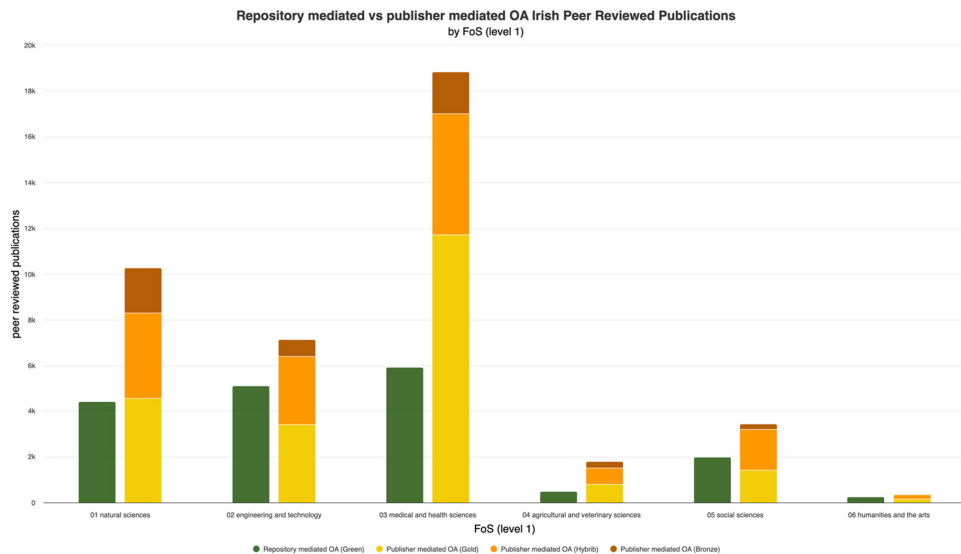


Figure 12: Repository mediated vs publisher mediated OA Irish Peer Reviewed Publications by FoS (level 1)

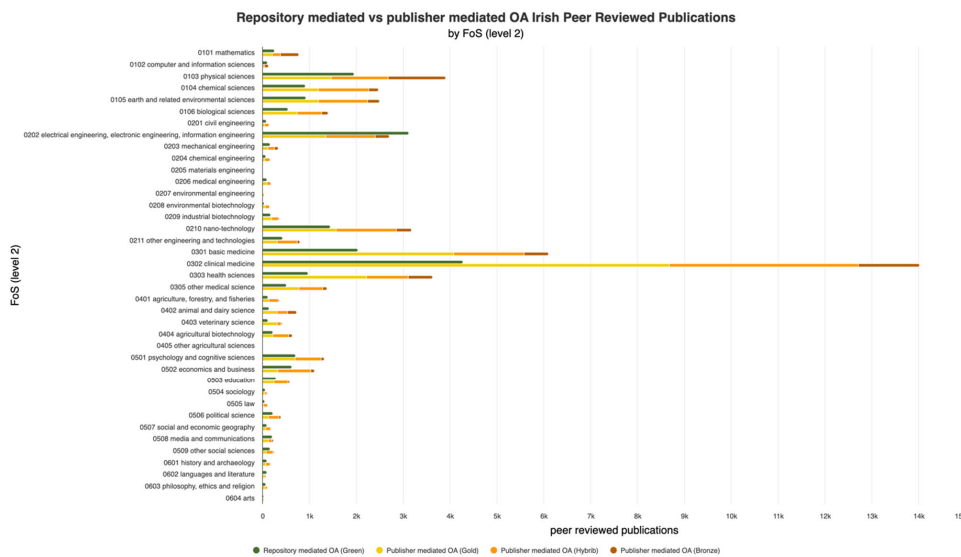


Figure 13: Repository mediated vs publisher mediated OA Irish Peer Reviewed Publications by FoS (level 2)

To enhance the coverage of FoS, we are presently implementing our FoS classifiers on Irish Publications, aiming for a more comprehensive representation in the National Open Access Monitor of Ireland.

2.2.3 Unrealised & Immediate Open Access

As shown in Table 5, which outlines the count and proportion of Immediate and Unrealized Open Access Irish Peer-Reviewed Publications, it is apparent that there are a significant 16,000 publications falling under the "unrealized open access" category. These publications, while initially designated as Closed Access, possess an Open Access version archived in a repository. This identification is achieved by calculating the Green OA publications that do not fall within the Gold, Hybrid, or Bronze categories. Half of the Irish Open Access peer-reviewed publications are readily available in an Open Access route (Immediate OA).

Table 6: Number and share of All vs Immediate vs Unrealised Open Access Irish Peer-reviewed Publications

All vs Immediate OA	# Irish OA peer-reviewed publications	Share of # Irish OA peer-reviewed publications
All OA	127,221	100,00%
Immediate OA (Gold, Hybrid, Bronze)	65,438	51,44%
Unrealised OA (Green and not Gold/Hybrid/Bronze)	16,634	13,07%

Commented [CF27]: Table 6?

Commented [le28R27]: Yes, it is 6.

Commented [CF29]: For Advisory Group. I question the use of "unrealised open access" and "immediate open access" here.
1) Green OA falls under the BOAI definition of OA, and is therefore not "unrealised". Propose that "unrealised" is defined as closed or bronze open access. 2) "Immediate" open access is a different designation which relates to the date of publication + the date the item is made open access and should be assessed together with embargoed content. As per the tender requirements "The Monitor **MUST** identify embargoed and immediate open access." Do you agree?

Commented [le30R29]: This has been fixed in the presentation

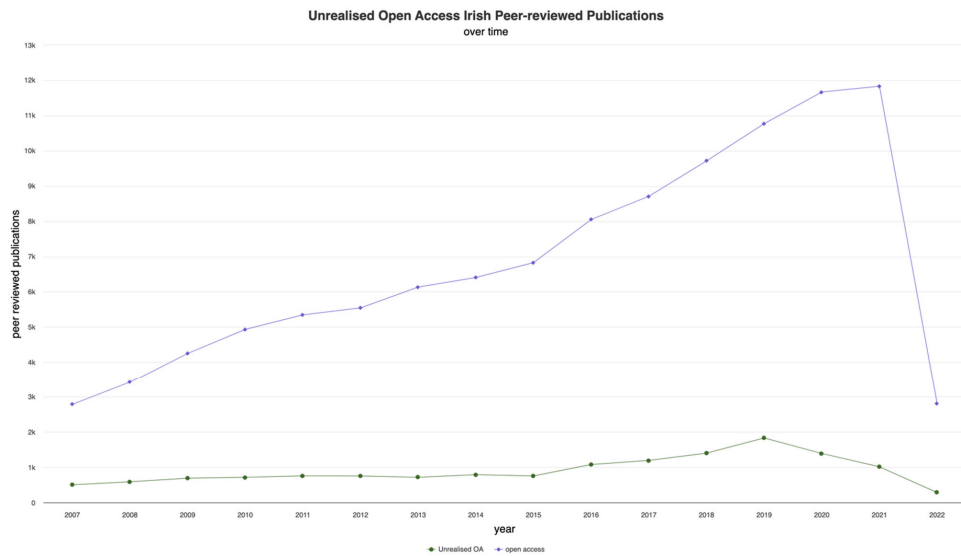


Figure 14: Unrealised Open Access Irish Peer-reviewed Publications over time

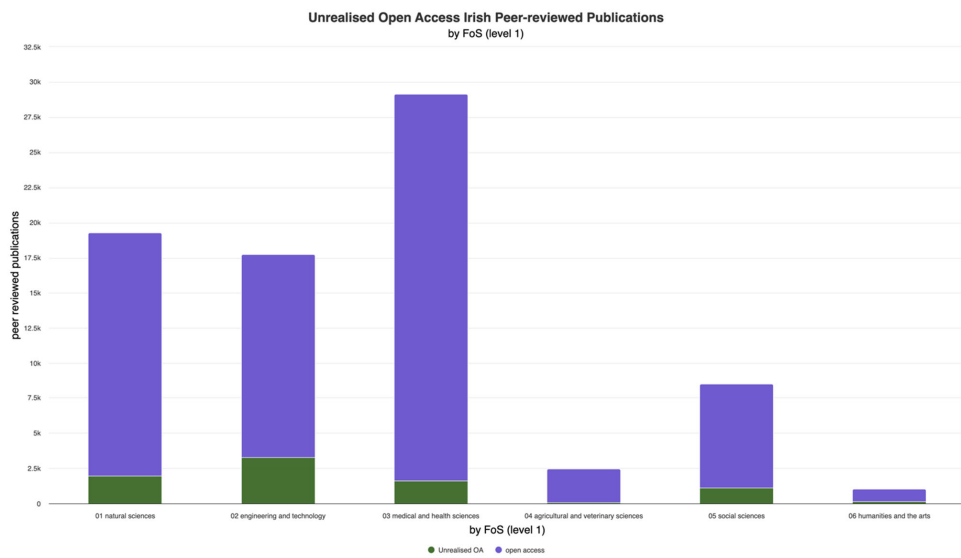


Figure 15: Unrealised Open Access Irish Peer-reviewed Publications by FoS (level 1)

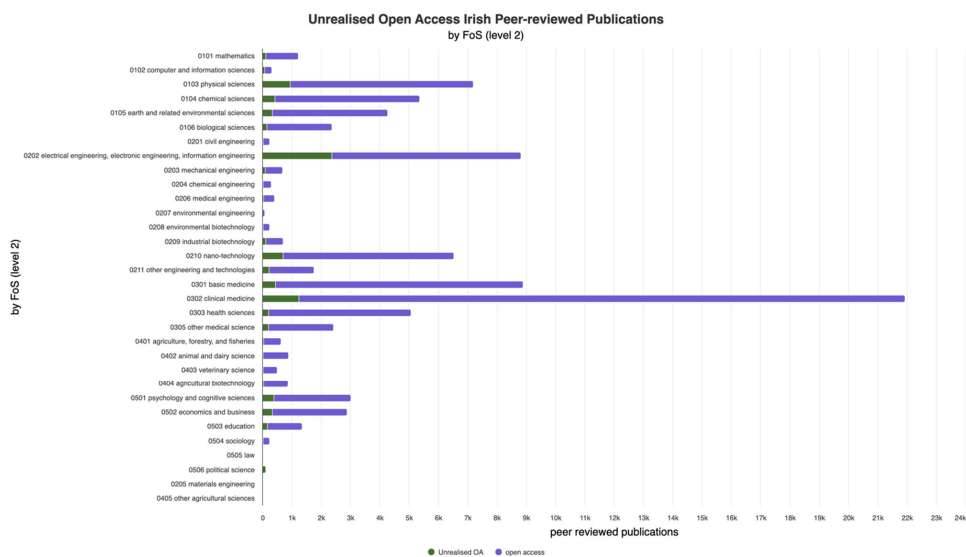


Figure 16: Unrealised Open Access Irish Peer-reviewed Publications by FoS (level 2)

To enhance the coverage of FoS, we are presently implementing our FoS classifiers on Irish Publications, aiming for a more comprehensive representation in the National Open Access Monitor of Ireland.

2.2.4 FAIRness

Table 7: Number and share of Open Access Irish Peer-reviewed Publications with a CC license

	# Irish OA peer-reviewed publications	Share of # Irish OA peer-reviewed publications
with CC license	191,929	78,61%
without CC license	52,227	21,39%

Commented [CF31]: Please include an explanation for how OpenAIRE considers FAIRness = CC license.

Commented [le32R31]: We will include it.

Commented [CF33]: Is there a potential to expand this to include the other OA licenses that Unpaywall use to determine if an article is Open Access:
<https://support.unpaywall.org/support/solutions/articles/44002063718-what-is-an-oa-license->

Commented [le34R33]: Licenses are not easily mapped to CC licenses. It requires legal expertise.

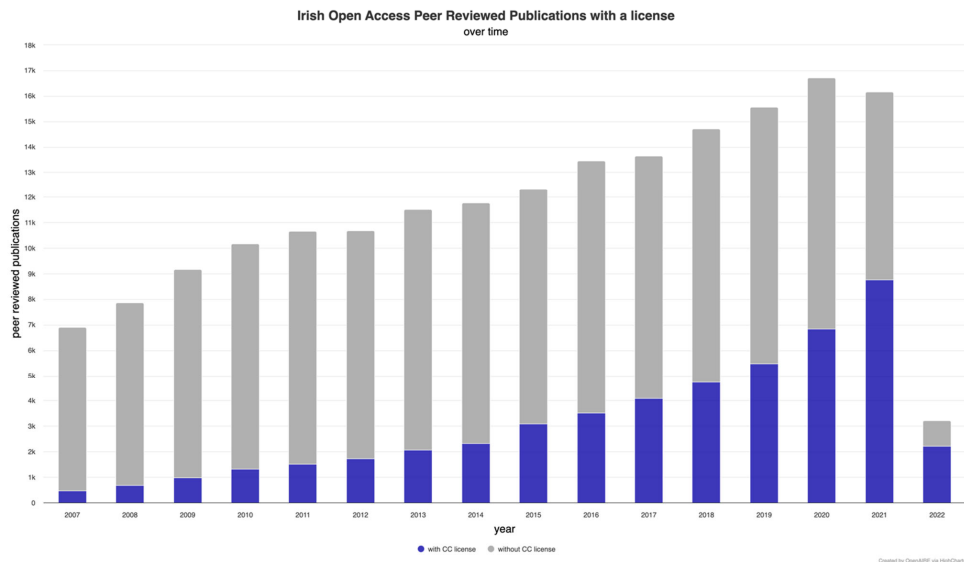


Figure 17: Irish Open Access Peer Reviewed Publications with a license over time

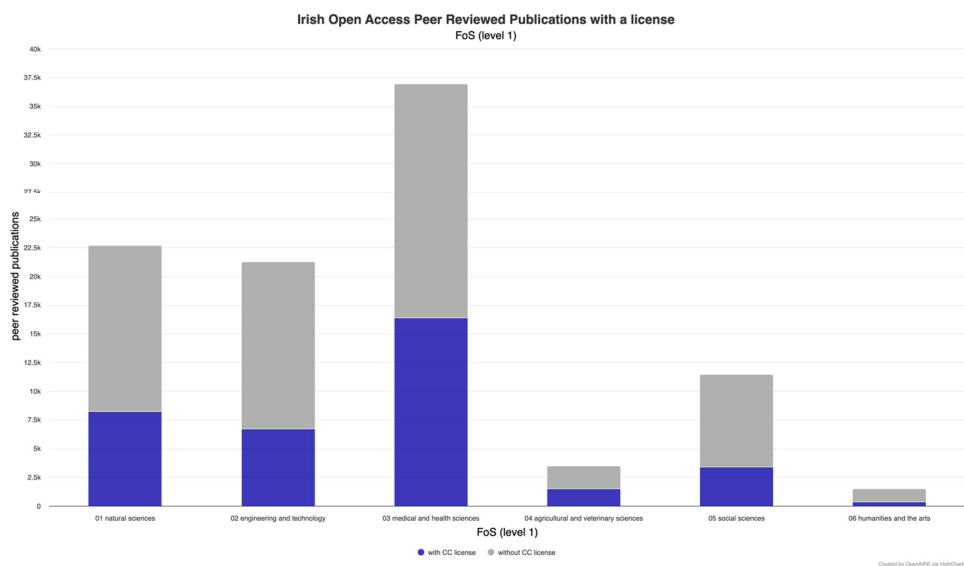


Figure 18: Irish Open Access Peer Reviewed Publications with a license by FoS (level 1)

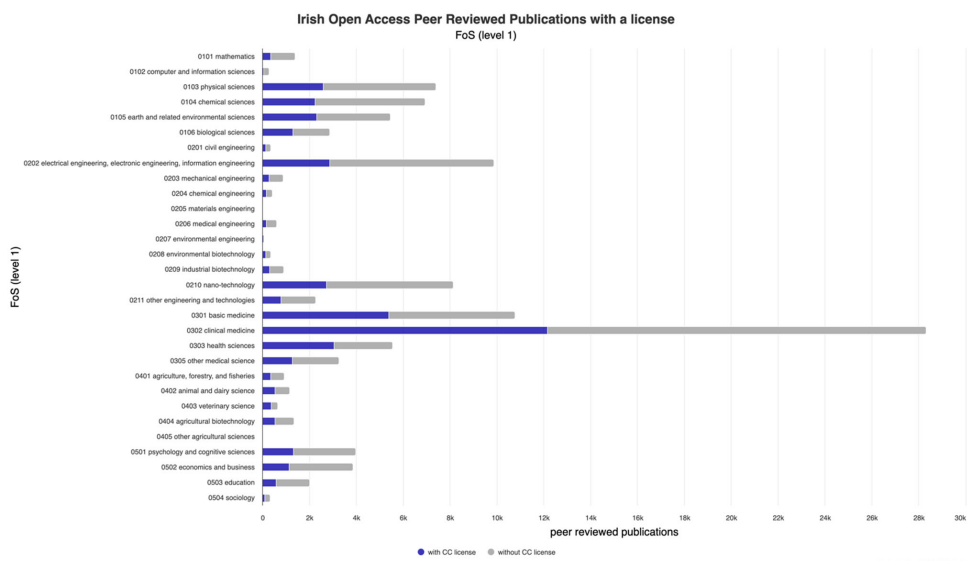


Figure 19: Irish Open Access Peer Reviewed Publications with a license by FoS (level 2)

To enhance the coverage of FoS, we are presently implementing our FoS classifiers on Irish Publications, aiming for a more comprehensive representation in the National Open Access Monitor of Ireland.

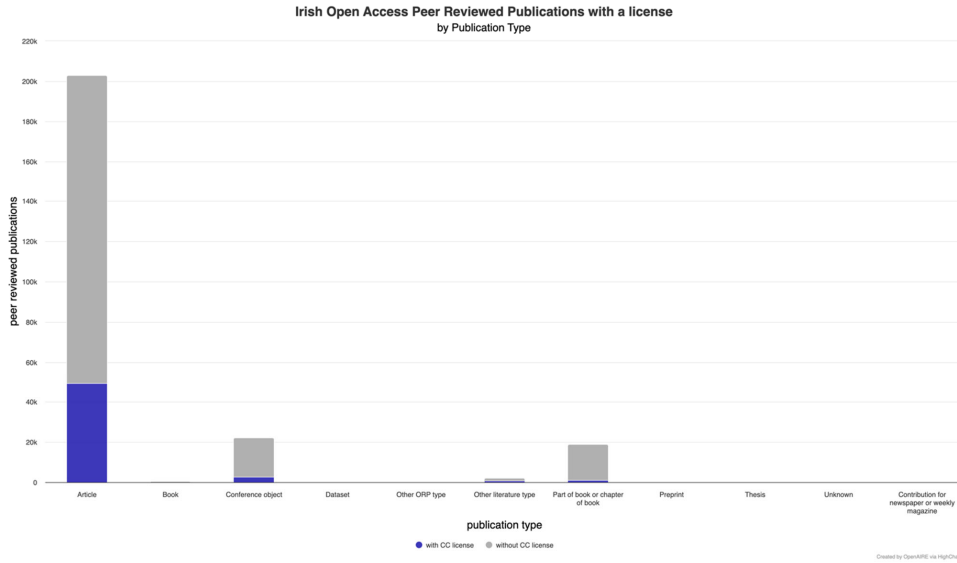


Figure 20: Irish Open Access Peer Reviewed Publications with a license by Publication Type

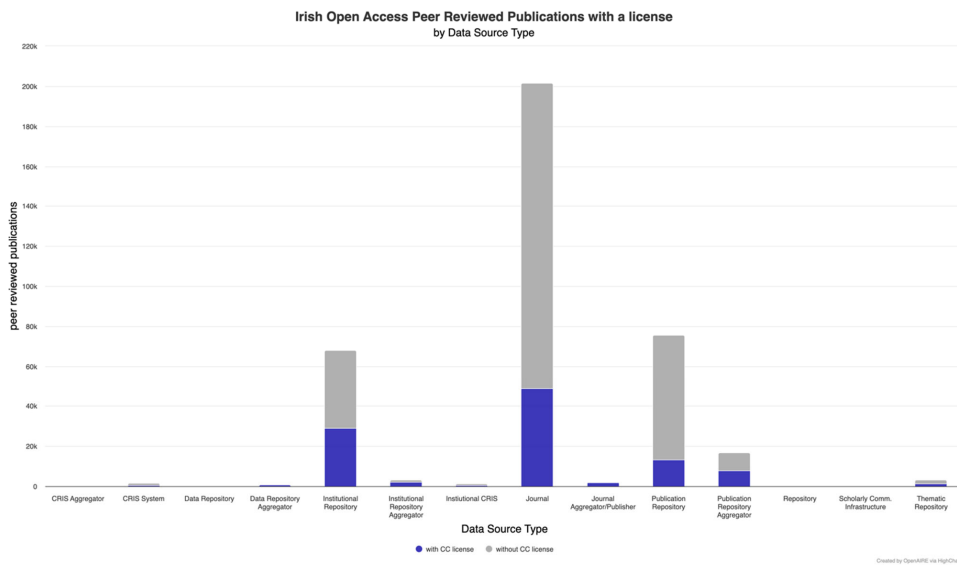


Figure 21: Irish Open Access Peer Reviewed Publications with a license by Data Source Type

In the preceding charts, it becomes evident that a significant segment of Irish OA publications does not feature a CC license. This highlights a scenario where one could contend that there is a necessity for initiatives aimed at training researchers in the art of negotiating their research publications under an Open Access license.

We are currently enhancing the Open License coverage of the Irish Publications and improving the license normalization algorithm, aiming for a more comprehensive representation in the National Open Access Monitor of Ireland.

The table below provides an overview of the progress made in grouping and normalizing licenses. It is important to note that a merged record of a publication may include multiple licenses.

Table 8: Licence normalisation

Licence	# Irish OA peer-reviewed publications
CC-BY	30,561
CC-BY-NC-ND	13,679
CC-BY-NC	4,759
CC-BY-NC-SA	3,755
CC-0	664
CC-BY-SA	375
CC-BY-ND	97
CC-BY-ND-SA	22
Elsevier TDM	10,584
Elsevier OA	8,135
Wiley VOR	4,123
Springer TDM	5,618
Wiley TDM	4,611

Commented [CF35]: This seems to be a very significant finding. How is it possible that research outputs are open access without open licenses? Using the BOAI definition of OA, this doesn't align. Only research outputs with open licenses should be considered as open access for the purpose of this report. Please provide more information to explain this section.

Commented [le36R35]: To be discussed in the presentation.

Commented [CF37]: For Advisory Group: propose that this statement isn't included. (a) no reference is made to the IReL-negotiated publishing agreements, where the license is negotiated by IReL and not the researcher (b) only research outputs with OA licenses should be included in this Report, therefore how this conclusion was reached is ambiguous (c) are there research outputs in institutional repositories without open licenses - if that's the case, then that could be called out here, but that's not what this section currently implies or what this statement is currently saying. Do you agree?

Commented [CF38]: Please provide further context in the Report on TDM licenses. Are readers to infer that TDM licenses are applied to non-OA content only (as an open license would allow for TDM by default)?

Commented [le39R38]: This is license-specific.

arXiv non-exclusive distribution	3,055
Oxford Academic reuse	2,802
Sage TDM	2,595
CambridgeCORE	2,454
IEEE	2,377
APS	1,545
ASM TDM	1,042
% Irish OA peer-reviewed publications with non-normalised licenses	11.63%

Table 9: Number and share of Irish Peer-reviewed Publications with a PiD

	with PiD	without PiD
# Irish peer-reviewed publications	244,156	0
Share of # Irish peer-reviewed publications	100%	0%

2.3 Plan S

Plan S⁸ represents a pioneering initiative with the primary objective of ensuring the immediate and universal accessibility of publicly funded scientific research. It stands as a significant milestone in the broader open-access movement within scientific publishing.

At its core, Plan S embodies a set of fundamental principles that dictate researchers funded by cOAlition S members must publish their research findings through open-access journals or platforms. Alternatively, they can choose to make their articles available in recognized open repositories without any embargo period. This commitment to openness encompasses diverse routes to compliance, including publishing in open-access journals,

⁸ <https://www.coalition-s.org/about/>

depositing articles in recognized repositories, and participation in transformative journals transitioning towards full open access.

Moreover, Plan S places emphasis on transparency and accountability regarding publication fees. It seeks to establish limits on charges levied by publishers for open-access publishing. Additionally, it underscores the importance of licensing, mandating that articles be published under an open license, preferably the Creative Commons Attribution licence(CC BY), which enables others to reuse the works with appropriate credit.

In this section, we explore the indicators associated with Plan S compliance, offering insights into the progress and adherence to these transformative principles within the Monitor.

Commented [CF40]: This section doesn't include compliance with Plan S when "Authors publish Open Access in a subscription journal under a transformative arrangement." Transformative Journals do not equal Transformative Agreements. Please update this section to include research published under transformative agreements.

Commented [le41R40]: Fixed in presentation.

Table 10: Irish Open Access Peer Reviewed publications by Journal Business Model

	# Irish OA peer-reviewed publications
Full OA with APCs	31,224
Diamond	146
Transformative Journal	5,467

Commented [CF42]: For Advisory Group: the PlanS Transformative Journal concept was launched in April 2020 (<https://www.coalition-s.org/transformative-journals-faq/>) - propose that the Report (and Monitor) only assigns a Transformative Journals as being a Transformative Journal since they were given that designation by PlanS and not prior to that date. This section seems to include publications in now-current Transformative Journals, but for all time. Do you agree?

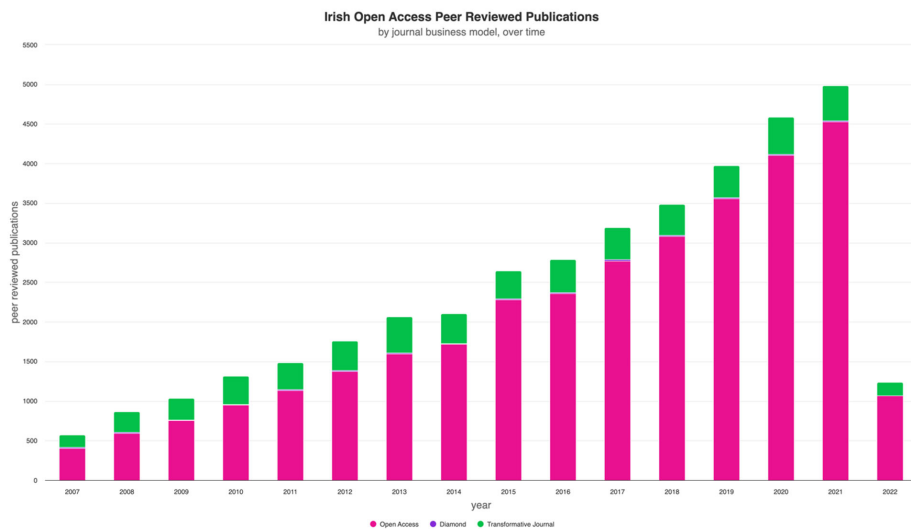


Figure 22: Irish Open Access Peer Reviewed Publications by journal business model, over time

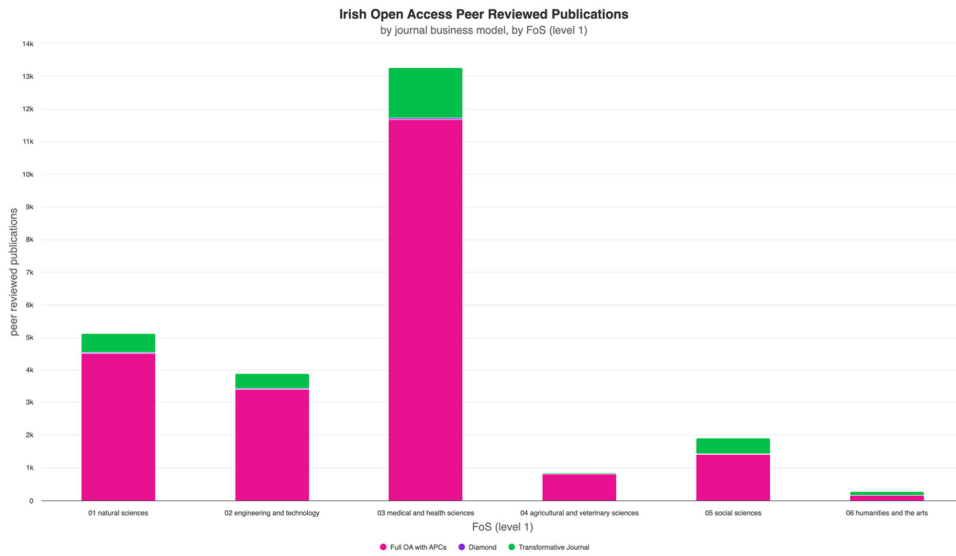


Figure 23: Irish Open Access Peer Reviewed Publications by journal business model, by FoS (level 1)

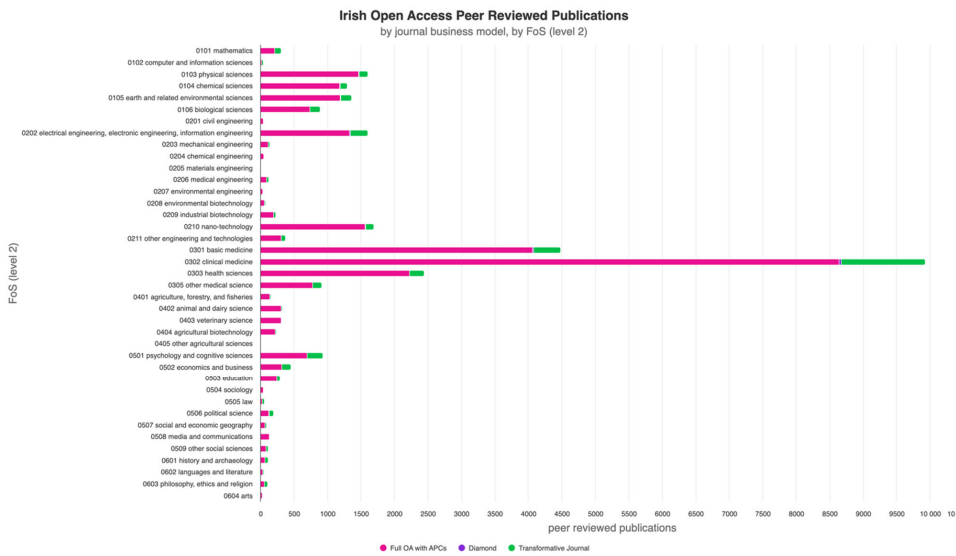


Figure 24: Irish Open Access Peer Reviewed Publications by journal business model, by FoS (level 2)

The Irish Open Access Peer Reviewed Publications by journal business model table and charts reveal that the Full OA with APCs business model holds a substantial proportion comparing to the Diamond and Transformative model.

Table 11: Total and average APCs of # Irish OA peer-reviewed publications

	# Irish OA peer-reviewed publications
APCs for all affiliated publications	€1,775,412
Average APCs per publication (For affiliated publications with reported APCs)	€2,233.22

The table presented above provides information on the overall and average APCs associated with Irish Open Access peer-reviewed publications. At present, the APCs are computed in accordance with the described methodology, combining the APCs reported by the institution with those for publications featuring co-authors affiliated with other institutions (APCs not funded by the institution). We are currently in the process of refining this calculation to specifically identify APCs reported solely by the institution. This enhancement will be integrated before the launch of the Monitor pilot.

Commented [CF43]: This methodology isn't clear. Please provide more information.

Commented [le44R43]: We will do it.

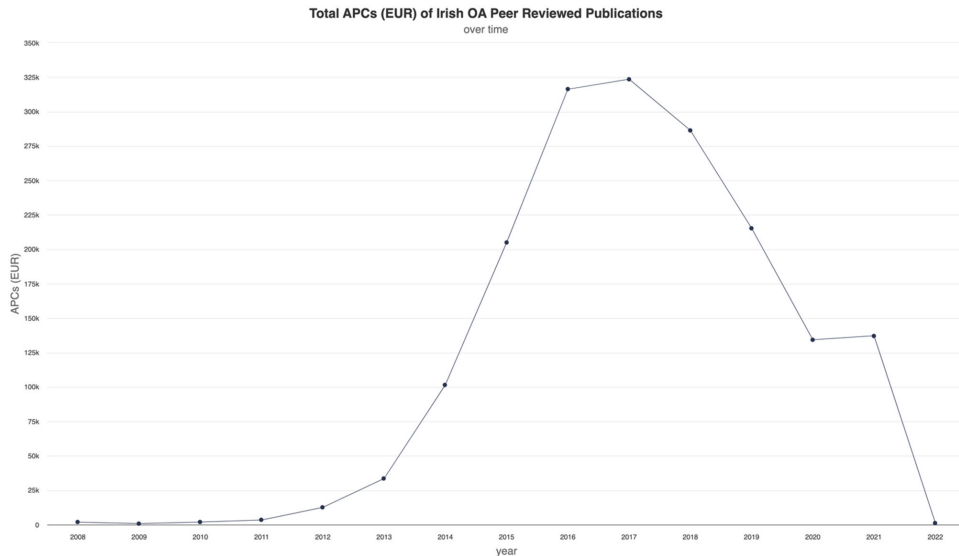


Figure 25: Total APCs (EUR) of Irish OA Peer Reviewed Publications over time

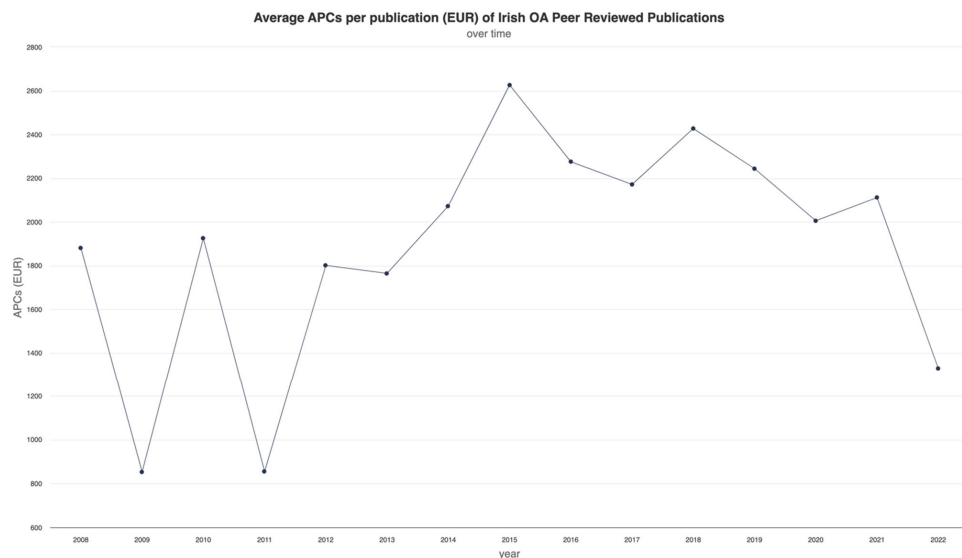


Figure 26: Average APCs per publication (EUR) of Irish OA Peer Reviewed Publications over time

Examining the breakdown of Total APCs by year, a noteworthy pattern emerges, indicating a substantial upsurge commencing in 2012, reaching its zenith during 2016-2017. The cumulative APCs and the average APCs per publication present another potential metric for evaluating the effectiveness of open access policies within RPOs or RFOs. A prevailing assumption suggests that the APC level serves as an indicator of the prestige or impact level of an Open Access journal. However, previous analyses have demonstrated that this assumption does not hold true, as high APCs do not consistently align with the high impact of specific articles. Empirical studies have revealed that exorbitant fees charged by certain journals for publishing articles do not necessarily correlate with increased citation rates.

The delay in the availability of 2022 Irish peer-reviewed publications' APCs, as of October 2023, can be attributed to the natural timeline associated with academic publishing. OpenAIRE is actively addressing this delay by focusing on two main strategies: refining data harvesting from Irish institutional sources, enhancing text mining for affiliations specifically for Ireland and reviewing the methodology for establishing peer-review status. Furthermore, we are in the process of aggregating and conducting additional assessments of APC data sourced from the DOAJ public data dump and OpenAPC. It's worth noting that the OpenAPC platform depends entirely on institutions supplying their APC data. Decentralised APCs paid by faculties or individual authors are not covered in the database.

Upon reviewing the two charts below, which feature the Top 20 Journals ranked by Total APCs and Average APCs per publication, we can discern the distribution of APC expenditures among the journals most frequently utilized by Irish researchers for Open Access publishing.

Commented [CF45]: For Advisory Group: from my reading, a meaningful metric here would be APC expenditure by publisher (rather than, or in addition to, journal). Would you also find this useful?

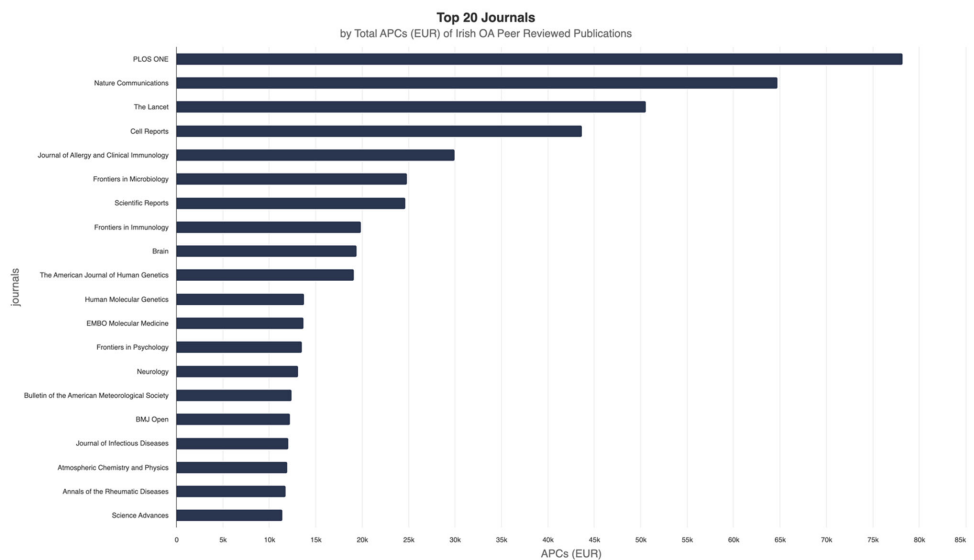


Figure 27: Top 20 Journals by Total APCs (EUR) of Irish OA Peer Reviewed Publications

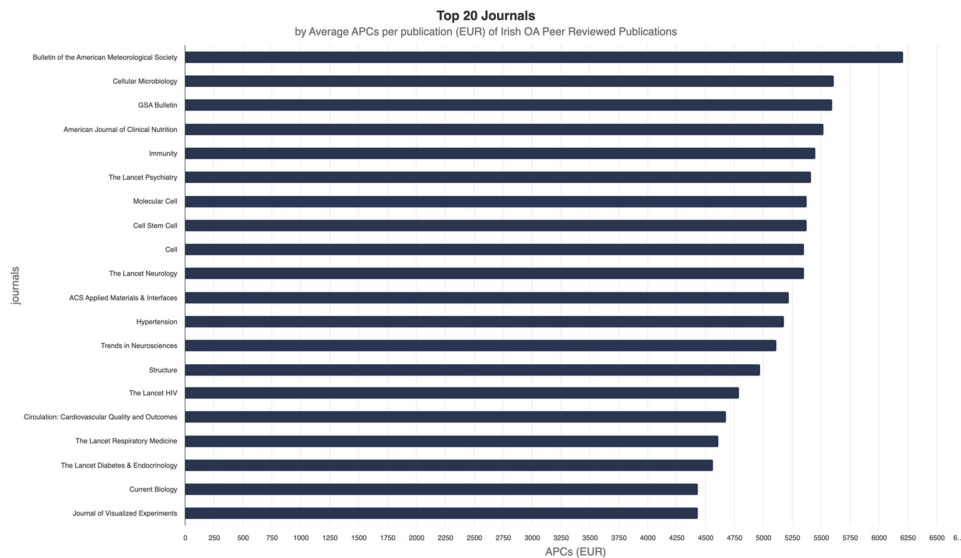


Figure 28: Top 20 Journals by Average APCs per publication (EUR) of Irish OA Peer Reviewed Publications

Commented [CF46]: For Advisory Group. Propose that this may not be a meaningful metric to include, as the Bulletin of American Meteorological Society could have come in first based on one APC paid. Do you agree?

To delve deeper into this trend, the following charts provide a detailed breakdown by Field of Science (FoS). This fine-grained perspective highlights Medical & Health Sciences domain driving the overall load of APCs. Further investigation will showcase more qualitative results especially with the increase of the FoS coverage until the delivery of the Monitor.

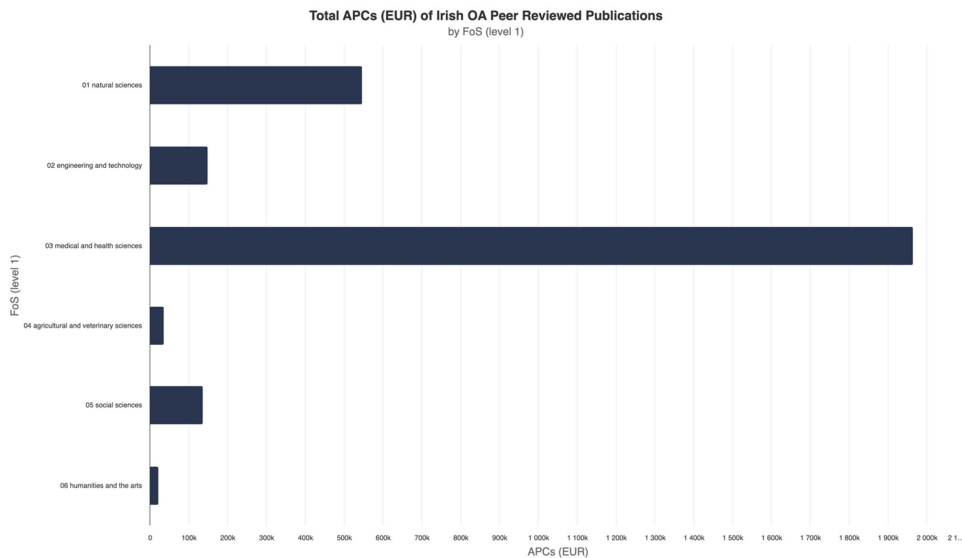


Figure 29: Total APCs (EUR) of Irish OA Peer Reviewed Publications by FoS (level 1)

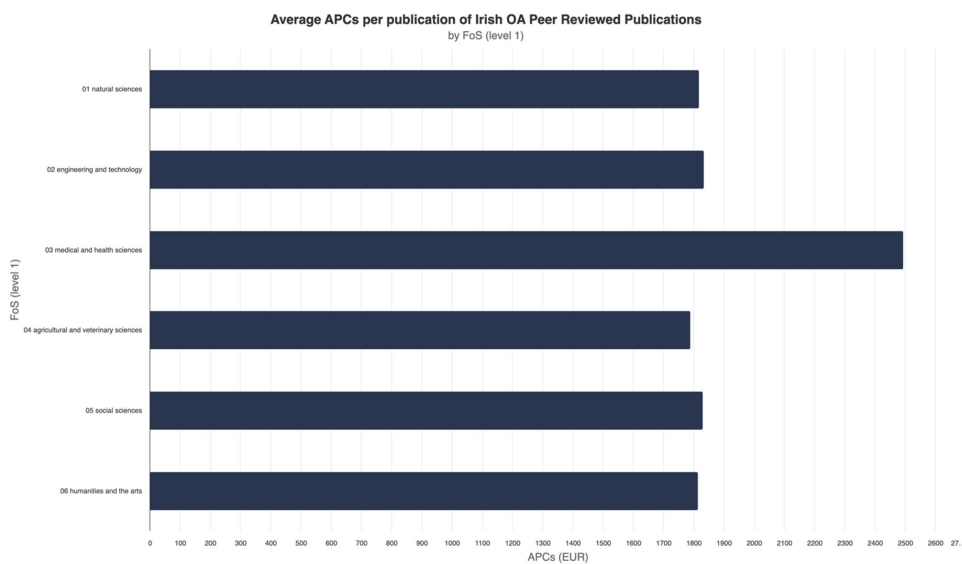


Figure 30: Average APCs per publication (EUR) of Irish OA Peer Reviewed Publications by FoS (level 1)

The following tables provide insights into the count and proportion of Irish peer-reviewed publications that have received grant support from funder members of the cOAlition S

consortium, including Science Foundation Ireland, Health Research Board, and Irish Research Council. Of the total Irish Open Access peer-reviewed publications, 16,742, comprising approximately 6.8%, have been funded by a Plan S Funder. Consequently, these publications adhere to Plan S requirements, ensuring their immediate availability as open access resources. An analysis of the distribution of grant-supported Irish peer-reviewed publications across various Open Access routes underscores Gold OA as the preferred publishing route, with Green, Hybrid, and Bronze options following. Furthermore, it's worth noting that approximately 40% of these publications are released under a Creative Commons (CC) license.

Commented [CF47]: For OpenAIRE: HRB and IRC aren't listed as endorsing organisations cOAlition S. <https://www.coalition-s.org/organisations/>. Please clarify their inclusion here.

Commented [le48R47]: Indeed they are not listed but here for example it is stated that they have endorsed Plan S. <https://atlantictu.libguides.com/openaccess/plans> We can remove them if they are not shown in the cOAlition S.

Table 12: Number and share of Grant supported Irish Peer Reviewed publications by a Plan S Funder

Plan S Funders	# Grant supported Irish peer-reviewed publications	Share of # of Grant supported Irish peer-reviewed publications
Science Foundation Ireland (SFI)	11,797	4,83%
Health Research Board (HRB)	2,125	0,87%
Irish Research Council (IRC))	2,840	1,16%

Commented [CF49]: For Advisory Group: as PlanS was launched in 2018, propose that this table either (a) only counts research outputs after 2018 in the context of PlanS or (b) splits pre-PlanS and post-PlanS. This table currently looks to feature from all time (or 2007-). Do you agree?

Commented [ed50R49]: Yes, agree, (b) to show pre- and post- Plan S would be useful. The Plan S principles took effect for most from 2021, so that might be a better start date?

Table 13: Number of Grant supported Irish Peer Reviewed publications by Open Access Route

Plan S Funders	Green OA	Gold OA	Hybrid OA	Bronze OA	With CC license
Science Foundation Ireland	2,090	2,892	1,514	5,60	4,417
Health Research Board	253	738	265	92	982
Irish Research Council	422	505	596	145	1,106

Commented [CF51]: A table that shows OA v. Closed would be meaningful here too.

Commented [le52R51]: OK, we will add it.

Table 14: Share of Grant supported Irish Peer Reviewed publications by Open Access Route

Share of Grant supported Irish Peer Reviewed publications by Open Access Route					
Plan S Funders	Green OA	Gold OA	Hybrid OA	Bronze OA	With CC license
Science Foundation Ireland	17,72%	24,51%	12,83%	0,05%	37,44%
Health Research Board	11,91%	34,73%	12,47%	4,33%	46,21%
Irish Research Council	14,86%	17,78%	20,99%	5,11%	38,94%

3 Methodology

This section provides an outline of the methodological steps taken to conduct the analysis of OA in Ireland presented in the previous section.⁹ This transparent approach ensures the integrity and reliability of our findings, guiding readers through the structured framework that underpins our assessment of the OA landscape.

3.1 OpenAIRE Graph: Foundation of the Monitor

The Monitor is built upon the OpenAIRE Graph (<https://graph.openaire.eu>). An open resource that aggregates a collection of research data properties (metadata, links) available for funders, organizations, researchers, research communities and publishers to interlink information by using a semantic graph database approach.

The Graph aggregates around 450 million metadata records from more than 120,000 trusted scholarly communication sources worldwide, including Crossref, Unpaywall, ORCID, institutional and thematic repositories (registered in OpenDOAR, re3data.org and FAIRSharing.org), Open Access journals, data archives, and the EOSC Service Catalogue. These metadata records are harvested and enriched with links between research results and projects, author affiliations, subject classifications, and links to domain-specific databases using dedicated inference algorithms. OpenAIRE's metadata records are cleaned, deduplicated, enriched, and transformed according to the OpenAIRE internal metadata model, generating the final OpenAIRE Graph. A new version of the OpenAIRE

⁹ Due to its relevance some of the text has been copied from the Inception Report.

Graph is available every month. The OpenAIRE Graph is available for download and reuse under a CC-BY license.

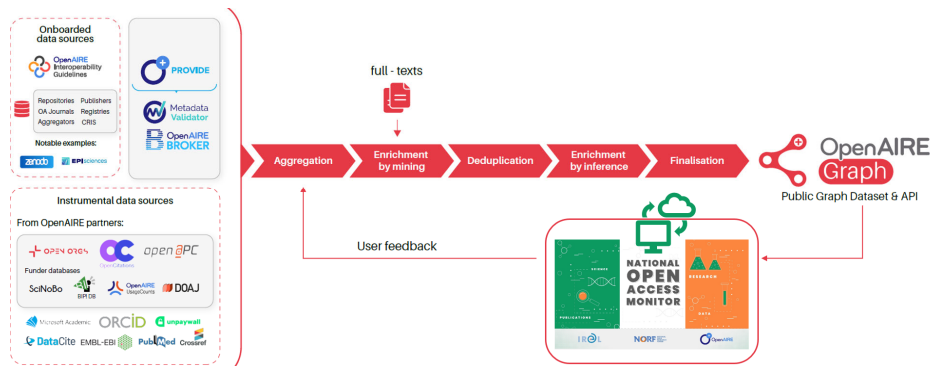


Figure 31: The Monitor and the OpenAIRE Graph Pipeline

3.2 The Publication Set of the Monitor

This section outlines the processes through which the Monitor compiles and utilizes data from Irish Research Performing Organizations (RPOs) and Research Funding Organizations (RFOs), ensuring comprehensive representation of peer-reviewed publications.

We take the following steps:

1. Identify Irish RPOs and their publications.
2. Identify Irish RFOs and their publications.
3. Exclude non-peer-reviewed publications
4. Make sure PIDs are used for Irish RPOs, RFOs and publications.

The final set for the Monitor currently includes **244.156** peer-reviewed publications.

Identification of Irish Research Performing Organisations' (RPOs) Publications

The Monitor leverages the comprehensive affiliation information already present in the OpenAIRE Graph to identify Irish RPO research output. The provenance of affiliation links in the Graph includes

1. Institutional data sources registered in OpenAIRE (repositories, CRIS, Open Access Journals)
2. Metadata from harvested data sources such as Crossref.
3. Inferred links via text mining.

- Links created via the claim and link functionalities in OpenAIRE EXPLORE¹⁰.

The table below provides some summary statistics for the RPOs in the Monitor¹¹.

Table 15: RPOs & their publications in the Monitor

# Irish RPOs	# publications of Irish RPOs
36	337,121

Irish institutional data sources registered in OpenAIRE

Institutional data sources (1. above) provide direct affiliation information to the OpenAIRE Graph, i.e. a publication from an institutional source is immediately given the corresponding affiliation. We present the coverage of Irish institutional data sources in the Monitor below.

Table 16: Irish Institutional Data Sources & Alignment with OpenAIRE

Institutional Data Source Type	Registered in OpenAIRE		Harvested by OpenAIRE (directly or via a compatible aggregator)	
	# sources	# publications	# sources	# publications
Repositories	17	98,151	0	0
CRIS Systems	0	0	0	0
OA Journals	0	0	18	2,628

Among the registered Irish repositories within OpenAIRE, a cumulative count of 98,151 publications is recorded. However, it's important to note that, among these repositories, 13 adhere to the BASIC and to version 2.0 of the OpenAIRE Guidelines. As a consequence, these repositories do not conform to the most recent IT and repository standards, which necessitate more contextually enriched content, including links and associations with various research outputs and entities. Furthermore, they do not accommodate diverse and enhanced vocabularies. These repositories also fall short of alignment with Open Science mandates and established standards, as they do not endorse well-established metadata schemas and namespace abbreviations for project identifications, including

¹⁰ <https://explore.openaire.eu/>

¹¹ The RPO list that will be provided by IReL, through the “National Open Access Monitor Survey: Organisational Identity” will be used to refine this set.

Commented [CF53]: For Advisory Group: for those of you that are involved in managing institutional repositories, is there content in your IRs that researchers might have carried out when not at your organisation (i.e. before they joined your organisation) - where a researcher will have given another organisation as their affiliation in the article itself? Is there a potential for research outputs to be counted as affiliated with your organisation that should be, due to this methodology?

For institutional sources that are CRIS, it's my expectation that a list of works associated with a researcher will include works published when they were not at the institution. Is this also your expectation?

EC funding programs, among others. The subsequent actions, in conjunction with the support of the “National Open Access Monitor Survey: Organisational Identity” and in collaboration with the Irish network of repositories, will involve the identification and registration or selective harvesting of repositories/CRIS that are currently not registered. Additionally, an ensuing phase will focus on enhancing compatibility with the OpenAIRE Guidelines, preferably upgrading to version 3.0 or, preferable, version 4.0.

OA Journals that are harvested through a compatible aggregator are not inherently linked to the institution or publisher within OpenAIRE. The subsequent actions involve institutions either registering their OA Journals in OpenAIRE or facilitating the registration and harvesting process. This is contingent upon the platforms supporting the OAI-PMH harvesting protocol. Alternatively, affiliations may be updated if these OA Journals have been associated with the institutions in the “National Open Access Monitor Survey: Organisational Identity”.

Commented [CF54]: This isn't an intended outcome of the survey. If a RPO tells us they host a journal, it's not appropriate to infer that all publications in that journal should be affiliated to that RPO.

Commented [le55R54]: We will change it.

Identification of Irish Research Funding Organisations' (RFOs) Publications

To guarantee a thorough representation of funded research outputs, the OpenAIRE Graph establishes links between publications and their associated funding data through a variety of methods:

- Harvesting links from repositories, OA Journals, CRIS systems.
- Merging information from Crossref's Open Funder Registry¹².
- Collecting links from users via the “link” functionality.
- Exchanging data with the EC's IT systems for EC/FP funding.
- Text mining of full text publications to identify the grants for 30+ funders that have joined OpenAIRE (see next paragraph). Science Foundation Ireland (SFI) is one of them.

Irish Funders in OpenAIRE

Irish funders are represented in the OpenAIRE graph through two primary avenues. The first is by directly joining OpenAIRE¹³, a process that entails providing a comprehensive list of research projects, the creation of a tailored text mining algorithm for data extraction, and meticulous curation of project-publication links to ensure accuracy. Science Foundation Ireland (SFI) has successfully undergone this process.

¹² <https://www.crossref.org/services/funder-registry/>

¹³ <https://www.openaire.eu/funders-how-to-join-guide>

Table 17: Irish Funders that have joined OpenAIRE

Irish RFOs that have joined OpenAIRE	Projects	Publications
Science Foundation Ireland	6,384	11,797

The second avenue for representation is through the Open Funder Registry (OFR)¹⁴ using the funders' fundref IDs. While this allows funders to be associated with publications via valid DOIs in the OpenAIRE Graph, it does not offer the granularity of the direct integration, notably the curated project-publication links.

Table 18: General Representation of Irish Funders in OpenAIRE (excluding SFI)

# Irish RFOs integrated via OFR	# Irish RFOs integrated via OFR with publications	# Irish RFOs' publications
142	79	11,351

For a funder to achieve the detailed representation observed with SFI, a direct integration with OpenAIRE is recommended. This not only ensures a comprehensive presence but also guarantees the precision of the data incorporated.

Peer-Reviewed Publications

We refine the set of Irish publications, as detailed in the preceding section, by focusing only on peer-reviewed articles, based on the following criteria:

- 1. Curated Peer-Review Assessment:** The OpenAIRE team has engaged in a curation process to determine the peer-review status of journals. This hand-curated assessment has been integrated into the Graph and is continuously under development.
- 2. Exclusion of Grey Literature:** We also filter out 'grey literature', which includes document types that typically bypass the peer review process, such as reports, theses, and white papers. Given that the OpenAIRE Graph aggregates data from various sources, resulting in merged records, we specifically exclude entries where all instances are identified as grey literature.
- 3. Presence of DOI from Crossref:** Since Crossref predominantly catalogues peer-reviewed content, its DOIs help maintain the scholarly credibility of our included publications.

¹⁴ Crossref, <https://www.crossref.org/services/funder-registry/>

The combination of these three criteria gives the following number of Irish peer-reviewed publications of the Monitor.

Table 19: Irish Peer Reviewed Publications

Irish peer-reviewed publications
244,156

Additional Criteria Under Examination

Beyond the core criteria, we are actively delving into additional parameters that might further refine our identification process.

- **Presence of DOI from DataCite:** We are ascertaining if such an inclusion can offer breadth to our dataset as we have been given examples of peer-reviewed data sources that get DOI from DataCite.

Table 20: Irish publications with DOI not assigned by Crossref and NOT included in current set

Irish publications with DOI not assigned by Crossref and NOT included in current set
5,202

- **Reference Count by Field of Study (FOS):** We are also exploring the possibility of adding a reference count, below which a publication can be considered non-peer-reviewed. Key in this process is examining the mean and standard deviation in the number of references per scientific field to ascertain whether this is a meaningful criterion.

Table 21: Mean and standard deviation in # of references by FoS (level 1)

# References in publications	Mean	Standard deviation
01 Natural Sciences	2.45	3.88
02 Engineering & Technology	1.92	2.02
03 Medical & Health Sciences	2.69	2.72

# References in publications	Mean	Standard deviation
04 Agricultural & Veterinary Sciences	3.13	5.71
05 Social Sciences	1.93	1.83
06 Humanities & The Arts	2.20	1.69

The final draft of our report will solidify our selection criteria. Nevertheless, the OpenAIRE infrastructure is designed for adaptability. Given that all processes and indicator workflows are fully integrated, we can easily make modifications in the future to meet evolving requirements or to refine our approach.

Use of Persistent Identifiers (PIDs)

To achieve accurate and comprehensive monitoring of Irish scholarly publications, we place emphasis on the use of Persistent Identifiers (PIDs). PIDs serve as essential building blocks, allowing us to uniquely identify these publications, facilitating the discoverability, accessibility, and reusability of research outputs.

The Monitor specifically defines an Irish scholarly publication as one that contains a persistent identifier (PID) associated with an Irish organization. These PIDs can be found in various places within the publication's metadata, PID metadata, or even within the publication content itself. We seamlessly integrate a range of PIDs for both research outputs and organizations. The process of deduplication ensures that metadata records from different data sources are effectively merged, accompanied by publicly displayed provenance information. This comprehensive approach guarantees not only the widest possible coverage but also maintains the integrity and consistency of our data.

The table below provides an overview of the PIDs used for publications, organizations, and authors within the Monitor, along with the corresponding number of publications associated with each type of PID.

Table 22: Irish peer-reviewed publications by PID type

PID type	# Irish peer-reviewed publications
Publication PIDs	
Digital Object Identifier (DOI)	244,156
Handle	19,354
PubMed Central ID	29,772
PubMed ID	87,020

PID type	# Irish peer-reviewed publications
arXiv	1
Organisation PIDs	
Participant Identification Code	224,434
ISNI	373,695
OrgRef	364,789
Open Funder Registry	353,241
Wikidata	377,701
GRID	379,940
RingGold	25,781
ROR	380,093
OrgReg	363,754
ORCID iDs	
ORCID iD	116,584

3.3 Data Disambiguation Techniques

Deduplication in OpenAIRE: The OpenAIRE Graph collects metadata records about scholarly works from different providers, which can carry different information. To provide accurate statistics, OpenAIRE merges duplicate records of the same scholarly work. The deduplication process is described in detail in the following link: <https://graph.openaire.eu/docs/graph-production-workflow/deduplication/>

Organizations: Organizations within OpenAIRE are aggregated from diverse registries and undergo a deduplication process via OpenOrgs. This tool merges automation with a "human in the loop" mechanism. It is designed to cluster records that are more likely to be analogous, employing both URL-based and title-based functions. Through the process of grouping duplicates, representative organizations not only inherit all attributes from the combined records but also maintain a record of their origin. On the Monitor, managers overseeing the national, RPO, and RFO dashboards will have access to OpenOrgs, empowering them to deduplicate Irish RPO records.

Journals, Publishers, and Licenses: To ensure precision and reliability in its data, the Monitor disambiguates journals using their ISSN numbers and publishers through the utilization of Crossref metadata, including ROR IDs and DOI prefixes, among other identifiers. This effort is bolstered by custom text similarity algorithms.

Additionally, OpenAIRE is systematically working on normalizing licenses. The table below provides an overview of the progress made in grouping and normalizing licenses. It is important to note that a merged record of a publication may include multiple licenses.

Table 23: Licence normalisation

Licence	# Irish OA peer-reviewed publications
CC-BY	30.561
CC-BY-NC-ND	13.679
CC-BY-NC	4.759
CC-BY-NC-SA	3.755
CC-0	664
CC-BY-SA	375
CC-BY-ND	97
CC-BY-ND-SA	22
Elsevier TDM	10.584
Elsevier OA	8.135
Wiley VOR	4.123
Springer TDM	5.618
Wiley TDM	4.611
arXiv non-exclusive distribution	3.055
Oxford Academic reuse	2.802
Sage TDM	2.595
CambridgeCORE	2.454
IEEE	2.377
APS	1.545
ASM TDM	1.042

Licence	# Irish OA peer-reviewed publications
% Irish OA peer-reviewed publications with non-normalised licenses	11,63%

As of now, approximately 89% of licenses have been successfully grouped and normalized. However, determining how these licenses should be accurately compared and categorized, particularly in the case of non-CC licenses, is a concern. For the construction of Hybrid OA publications (in a hybrid journal with a licence) we are currently only including CC licenses and are investigating which other ones should be included.

Authors: Researcher dashboards will be seamlessly integrated with their ORCID profiles. Originally, we had considered using text mining for author disambiguation. However, by anchoring the profiles to ORCID IDs, we have streamlined the process, allowing for easy identification of researchers *using all possible name variations* associated with their ORCID ID. This approach eliminates the need for additional text mining efforts.

3.4 Enrichment via Text Mining

In order to enrich metadata and enhance the comprehensiveness of scholarly records, OpenAIRE employs several effective text-mining methods. These methods include:

- **Affiliation Matching:** This process involves matching affiliations extracted from PDF and XML documents with organizations listed in the OpenAIRE organization database.
- **Funding Classifiers:** Utilizing a document classification algorithm, OpenAIRE analyses free text from abstracts of publications to categorize scientific text into one or more predefined content classes, such as funders and projects.
- **Extraction of Acknowledged Concepts:** OpenAIRE scans plaintexts of publications to identify acknowledged concepts. These may include grant identifiers (projects) from funders, accession numbers of bioentities, mentions of EPO (European Patent Office) patents, and custom concepts that link research objects to specific research communities and initiatives within OpenAIRE.
- **Metadata Extraction:** OpenAIRE employs the CERMINE project to extract plaintext and metadata from PDF documents. This extraction process covers various aspects, including titles, authors, affiliations, abstracts, keywords, journal names, volume and issue information, parsed bibliographic references, as well as the structure of document sections, section titles, and paragraphs.

For the Monitor, we have successfully completed the PDF aggregation exercise, which involved collecting accessible PDF documents of Irish RPOs.

Thus, we have meticulously reviewed 84K URLs extracted from the metadata of the 226K Irish OA publications. Among these URLs, we have successfully retrieved 61K PDFs. Our ongoing efforts involve inspecting the remaining 23K URLs to further enhance our coverage.

These documents will pass through the graph pipeline, enabling the implementation of the text-mining methods mentioned above. We are also mining them to classify them into FoS (FoS) and Sustainable Development Goals (SDGs) and identify the corresponding author affiliation. Methodology for the latter is addressed in Section 3.6

FoS (FoS) Classification System: To categorize into distinct levels FoS¹⁵, we have integrated an advanced classification system (Kotitsas, et al. 2023). This system utilizes Natural Language Processing (NLP) to analyse various components of the OpenAIRE Graph, including abstracts, citations, references, and venues. As a result, each publication is systematically classified into FoS classes down to level 3, adding precision to its scientific domain. This hierarchical categorization not only provides a structured framework but also bolsters our ability to pinpoint multidisciplinary overlaps within the research.

SDG Classification System: In order to contextualize the impact of research on addressing paramount global challenges, we have incorporated a classification mechanism aligned with the UN Sustainable Development Goals (SDGs). This schema is engineered to elucidate the alignment of research endeavours with critical issues, ranging from climate adaptation, biodiversity preservation, mitigation of environmental contaminants, to socioeconomic upliftment.

3.5 Indicators

The table below presents the construction methodology of indicators included in the Monitor, offering a detailed look at how each indicator in the Monitor is derived and calculated. The definitions of these indicators are given in the Glossary at this beginning of this report.

Table 24: Construction methodology of indicators

Attribute	Construction Methodology
Journal Business Models	
Open Access	Utilizing OpenAIRE's curated Gold ISSN list, Unpaywall metadata, and DOAJ journals.

¹⁵ The taxonomy is presented here: <https://explore.openaire.eu/fields-of-science>

Attribute	Construction Methodology
Subscription	Journals without any open access articles.
Hybrid	Journals with open access articles that are not OA journals.
Transformative	We identify Transformative Journals by ISSN matching with the publicly available Transformative Journals data ¹⁶ from Plan S initiative.
Journal APC Business Models	
Diamond OA	We obtain APC data from DOAJ using DOAJ's Public Data Dump ¹⁷ (an exportable version of the journal metadata). We cross-reference this with the OpenAPC ¹⁸ data integrated in the OpenAIRE Graph.
OA Types/Colours	
Green OA	An open access scientific publication deposited in a repository
Gold OA	A scientific publication published in an OA journal as defined above.
Hybrid OA	An open access scientific publication published in a hybrid journal with an open license. <i>At this point we consider only CC licenses "open"%. In principle, this means that we may be underestimating the number of hybrid OA articles and overestimating the number of bronze.</i>
Bronze OA	An open access scientific publication published in a hybrid journal without an open license.

Commented [CF56]: It is our expectation that at a minimum the list of licenses used by Unpaywall is used, see page 29/30 here <https://zenodo.org/records/7588788>

Commented [le57R56]: "

1. ACS Editors' Choice
2. APS License for Accepted Manuscripts
3. Open Access for APA Journals Authors
4. Many other less-common licenses, as long as they grant users sufficient rights to freely use and redistribute content.

"

1, 2, 3 OK. 4 is not well-defined.

¹⁶ <https://journalcheckertool.org/transformative-journals/>

¹⁷ <https://doaj.org/docs/public-data-dump/>

¹⁸ <https://openapc.net/>

¹⁹ See discussion in Section 3.3

Attribute	Construction Methodology
Accessibility – Interoperability	
Accessible	Accessible publications are sourced from OpenAIRE's full text collection, which holds PDFs of over twenty million OA publications. OpenAIRE's PDF aggregation system, recalibrated to prioritize Irish publications, examines the URL links in each publication's metadata to retrieve the corresponding PDF document. Given that multiple links can be associated with one publication, each is navigated. With ongoing automation, the system's coverage consistently expands, verifying if OA publications are accessible through thorough URL link inspection and PDF retrieval.
Interoperable	The construction of interoperable publications within the Monitor is intrinsically tied to their accessibility. Since we systematically fetch PDFs, any publication that is accessible through this process is also considered interoperable. In essence, the minimum threshold for interoperability is met when a publication's full text is accessible in a machine-readable format through the PDF aggregation system.

3.6 Additional Aspects

In this section, we outline the construction methodologies for specific metadata elements vital to the Monitor. While the subsequent Data Evaluation section delves into a comprehensive analysis of all key elements, here we focus solely on those requiring construction methodologies that have not been previously addressed.

Corresponding author affiliation: In identifying the corresponding author's affiliation in the Graph, we face limitations due to the lack of explicit tagging of this role in metadata from integrated data sources. While some sources are in development stages to provide such details, we devised two methodologies to bridge this gap:

1. **Contributor Rank Analysis:** We identify the corresponding author based on their position in the author list. Typically, *if the list is not alphabetical*, we consider the first author as the corresponding author. However, the sequencing of authors is not commonly represented in metadata, and we have managed to identify the first author in just 0.04% of Irish peer-reviewed publications using this method.
2. **Text Mining:** Exclusively for Irish publications, we are employing text mining on PDFs to discern the corresponding author's affiliation. It is worth noting that this approach is only possible for OA publications, as those PDFs are available.

Our final report will offer a comprehensive evaluation of the efficiency and limitations of these approaches.

Publicly-funded: To identify Irish scholarly publications that align with the definition of "publicly funded research as research undertaken in whole or in part via publicly funded resourcing or remuneration, e.g., salaries, grants, contracts, etc.," several steps are underway:

1. A comprehensive desk research is in progress to discern public RPOs and RFOs.
2. Survey currently conducted by IReL.
3. OFR's metadata offers pertinent information on funder type which is being integrated in the Graph. The OFR is donated by Elsevier, and is updated around every 4-6 weeks with new and updated funder records. Existing entries are also reviewed to make sure that they are accurate and up-to-date. The available funder types are "Government" and "Private". Private funding subtypes include: academic, corporate, foundation, international, other non-profit (private), professional associations and societies. Government funding subtypes include: federal (national government), government non-federal (state/provincial government).
4. Metadata from the European Commission's project participant RPOs includes their activity type and identifies the public ones.

Commented [CF58]: Clarification: as the definition above states "in whole or in part", an organisation only needs to be in receipt of public funds (and not wholly publicly funded) to be categorised as publicly-funded for this purpose.

Commented [le59R58]: But the resulting publications are not necessarily linked to public funding. Is there a database of government grants?

Methodological Concerns: We have identified the following issues with the above methodology. Research might be funded indirectly by public spending, e.g., a private RPO receiving public grants or subsidies and then funding a research project, in which case the publication metadata may not include an acknowledgement of public funding.

As of now, within OFR, Science Foundation Ireland is the only entity classified as "government." Consequently, we have a total of 11,797 Irish publications that are publicly funded. The assessment of data quality will be conducted in the final report and will be enhanced by applying the aforementioned methodology to identify additional publicly funded Irish scholarly publications.

Detailed results and further insights will be provided in the final report.

4 Data Evaluation

In the context of Open Access (OA) monitoring, the quality and completeness of data are of utmost importance. This section conducts a thorough evaluation, focusing on essential metadata elements crucial for displaying data, indicators and filters on the platform. The following table provides a detailed overview of these elements, including their coverage and data quality within the context of Open Access in Ireland. Additionally, we offer practical comments and steps that both OpenAIRE and stakeholders can take to improve data accuracy and depth, ultimately enhancing the effectiveness of OA monitoring.

Table 25: Metadata Analysis

Metadata element & Issue	Quality	OpenAIRE Actions	Stakeholder Actions
<p>Affiliations: Identification and association of research publications to Irish RPOs</p>	<p>The provenance of affiliation links in the Graph includes</p> <ul style="list-style-type: none"> • Institutional data sources registered in OpenAIRE (repositories, CRIS, Open Access Journals) • Metadata from harvested data sources such as Crossref. • Inferred links via text mining. • Links created via the claim and link 	<p>Besides including about 129K data sources, the OpenAIRE pipeline includes affiliation text mining algorithms (see Section 3.4). OpenAIRE is focusing on refining data harvesting from Irish institutional sources and enhancing text mining specifically for Irish affiliations. The survey that is currently being conducted by IReL is expected to enhance the quality of these actions.</p>	<p>Stakeholders can proactively register their institutional resources with OpenAIRE²⁰, ensuring that the most up-to-date and comprehensive data is accessible and integrated into OA monitoring.</p>

Commented [CF60]: More actionable information is needed in this column throughout.

Commented [le61R60]: We will try to improve.

Commented [CF62]: Are there other actions that stakeholders can take to assist with affiliations being identified correctly?

Commented [le63R62]: Register the institutional data sources. We will investigate it and add here other actions too.

²⁰ Via OpenAIRE PROVIDE (<https://provide.openaire.eu/home>)

Metadata element & Issue	Quality	OpenAIRE Actions	Stakeholder Actions
	<p>functionalities in OpenAIRE EXPLORE.</p> <p>Additionally, OpenORGs is utilized for name deduplication, ensuring consistency and precision in recognizing and distinguishing distinct RPOs and their outputs.</p>		
<p>Corresponding author affiliation: Determining the institution of the corresponding author.</p>	<p>0.04% publications with <i>first</i> author affiliation identified.</p>	<p>We currently have data on the first author of publications, and are text mining the collected PDFs to assess the possibility of identifying the corresponding author affiliation (see Section 3.6).</p>	<p>Publishers, can assist by consistently highlighting corresponding author affiliations in published articles, facilitating easier identification through automated processes.</p>
<p>Funded publications: Establishing comprehensive funder-publication links.</p>	<p>While SFI (Science Foundation Ireland) is curated due to having joined OpenAIRE, for other public funders, we rely on the harvested metadata. Its quality is difficult to assess externally.</p>	<p>OpenAIRE is still in the process of analysing data quality for integrated Irish Funders (those via OFR). The survey that is currently being conducted by IReL is expected to additionally enhance the quality.</p>	<p>Funders can join OpenAIRE by providing a project list, to allow us to build dedicated text mining algorithms that discover project-publication links and clean harvested metadata. Stakeholders can ensure clear labelling of funding sources in metadata and acknowledgements. Funding sources can have PIDs.</p>

Commented [CF60]: More actionable information is needed in this column throughout.

Commented [le61R60]: We will try to improve.

Commented [CF64]: The IReL OpenAPC dataset would also be useful here. The dataset includes the DOI + the institution, which by definition of the IReL OA publishing agreements is the organisation the corresponding author is staff or student of. Can this be utilised?
https://github.com/OpenAPC/openapc-de/tree/master/data/transformative_agreements/IReL

Commented [le65R64]: Thanks, we will investigate!

Commented [CF66]: What does this mean in practice? What actions are you specifically requesting publishers to take?

Commented [CF67]: What does this mean in practice? Which specific metadata elements, and in what way does acknowledgement text influence how OpenAIRE pulls this information in?

Metadata element & Issue	Quality	OpenAIRE Actions	Stakeholder Actions
Grant award ID: Ensuring availability & establishing project-publication links.	The quality assessment of this metadata element will be detailed in the final report. SFI has full coverage since they have joined OpenAIRE and provided a project list.	OpenAIRE is in the process of analysing data for Irish Funders (those in OFR) to examine the metadata coverage of this variable. For funders that have joined OpenAIRE (SFI for now) we have Grant award IDs and mine for project-publications links.	As above
Publicly-funded publications: Identification of all publications resulting from publicly funded research	Data Quality to be assessed in final report. As of now, within OFR, Science Foundation Ireland is the only entity classified as "government." Consequently, we have a total of 11,797 Irish publications that are publicly funded.	OpenAIRE is working on identifying the list of public RFOs and RPOs. The survey that is currently being conducted by IReL is expected to enhance this list. See Section 3.6 for a discussion.	As above
Publication type: Differentiating between Version of Record (VoR) and pre-print.	It is challenging to assess the quality of this field. About 3.3% of publications labelled as articles are also explicitly labelled as pre-prints, but checks have shown that there are more. Version of record is not a commonly shared metadata element and the sources that do expose it have low coverage.	Comparing date of deposition in a repository and date of publication of an article. If the article was deposited after it was published it is likely that it is the VoR that was deposited. We are currently examining this option.	Publishers and repositories can ensure clear metadata tags for Version of Record and pre-print instances.

Commented [CF60]: More actionable information is needed in this column throughout.

Commented [le61R60]: We will try to improve.

Commented [CF68]: In which metadata fields? In the DOI metadata.

Commented [le69R68]: In the DOI metadata for publishers and in the metadata records of the repositories, following the OpenAIRE Guidelines. eg. for the latest version: https://openaire-guidelines-for-literature-repository-managers.readthedocs.io/en/v4.0.0/field_publicationtype.html#aire-resourcetype

Metadata element & Issue	Quality	OpenAIRE Actions	Stakeholder Actions
<p>PIDs: Identifying and ensuring consistent coverage of Persistent Identifiers (PIDs) for publications, organisations, and authors.</p>	<p>As of now all Irish Peer-Reviewed publications have been identified as having a PID.</p> <p>For authors right now we have almost 50% coverage of peer-reviewed publications with authors with an ORCID iD.</p>	<p>For the Organisations the OpenOrgs platform is essential in order to disambiguate the organisation different names and PiDs.</p> <p>For authors, the ORCID iDs publications coverage can be further enriched by several methods including the organisations disambiguation in OpenOrgs, the signing in of the Irish researchers into the Monitor providing their ORCID iD, the claiming and linking of research products functionality.</p>	<p>Stakeholders, especially publishers and repositories, are encouraged to consistently apply and expose PIDs for publications, authors, and organisations.</p>
<p>Peer-reviewed publications: Identifying and verifying the peer-review status of Irish publications.</p>		<p>Refer to Section 3.2 for a discussion.</p>	<p>Publishers can label journals or proceedings as peer-reviewed.</p>
<p>Year of publication: Low(er) coverage for 2022</p>	<p>We have 3.2K peer-reviewed Irish publications in 2022, whereas in 2020 and 2021 they are 16.1K and 16.7K, respectively</p>	<p>The delay in the availability of 2022 Irish peer-reviewed publications, as of October 2023, can be attributed to the natural timeline associated with academic publishing. OpenAIRE is actively addressing this delay by focusing on two main strategies: refining data harvesting from Irish institutional sources, enhancing text mining for affiliations specifically for Ireland and re-examining the peer-reviewed criteria.</p>	<p>Self-deposition of peer-reviewed articles in repositories.</p>

Commented [CF60]: More actionable information is needed in this column throughout.

Commented [le61R60]: We will try to improve.

Commented [CF70]: Expose where/how?

Commented [CF71]: What does this mean in practice? In specific metadata fields? In Crossref when they mint the DOI?

Metadata element & Issue	Quality	OpenAIRE Actions	Stakeholder Actions
Publisher & Journal: Challenges in deduplication	Multiple instances of the same publisher or journal may be present due to variances in naming conventions, abbreviations, or metadata inconsistencies.	OpenAIRE has started the deduplication of publishers and journals (see Section 3.3).	Publishers can provide authoritative lists of names and identifiers.
Licence: Challenges in deduplications and interpretation of licence agreement	While about 89% of licenses have been grouped and normalized, there remains a challenge in categorizing non-CC licenses.	OpenAIRE has started the deduplication of licenses (see Section 3.3).	
FoS: Coverage and potential under-representation of specific disciplines	Currently the FoS coverage for the peer-reviewed publications is 22.71%.	OpenAIRE will classify the remaining Irish publications by running them through our FoS classification system to ensure comprehensive coverage (see Section 3.4).	We encourage researchers, institutions, and other stakeholders to provide feedback on any observed discrepancies in the classification.
SDGs: Coverage	Currently the SDGs coverage for the peer-reviewed publications is 1.81%.	The remaining Irish publications are scheduled to undergo SDG classification to ensure comprehensive coverage (see Section 3.4).	Feedback on perceived mismatches or overlooked research contributions towards SDGs is encouraged.
Access rights changing over time, such as an Embargoed publication becoming OA .	To the best of our knowledge, the <i>original</i> access rights of a publication deposited in a repository is not a metadata	After the Pilot Monitor's delivery on M6, OpenAIRE will begin uploading regular data dumps to Zenodo. Additionally, we will display the progression of	Repositories can expose the data of original access rights of an article.

Commented [CF60]: More actionable information is needed in this column throughout.

Commented [le61R60]: We will try to improve.

Commented [CF72]: Please specify the metadata field(s) that this data is expected to be provided in.

Commented [le73R72]: Eg. For the Embargoed access they can expose the Embargo Period Date according to the OpenAIRE Guidelines.
https://openaire-guidelines-for-literature-repository-managers.readthedocs.io/en/v4.0.0/field_embargoenddate.html

Metadata element & Issue	Quality	OpenAIRE Actions	Stakeholder Actions
	element that repositories expose.	indicators like % OA, % Green OA, etc., on the trendline graphs, utilizing monthly historical snapshots.	
URLs to PDF full texts: Coverage and validity of URLs linking to PDFs in publication metadata.	Of the 84K OA publications tested thus far, 61K had accessible PDF files linked directly from their metadata URLs.	OpenAIRE will continue testing Irish URLs as they show up in publication metadata.	Enhance metadata integrity by providing accurate, unambiguous links to PDFs

Commented [CF60]: More actionable information is needed in this column throughout.

Commented [le61R60]: We will try to improve.

Commented [CF74]: Who is this action directed to?

Commented [le75R74]: Publishers and repositories.

5 Conclusion

In this draft report, we've endeavoured to provide a transparent and detailed overview of our findings from the National Open Access Monitor. It serves as both an evaluative tool and a roadmap, pinpointing the accomplishments in the OA landscape and highlighting areas that warrant further examination and action.

To sum up, for the current baseline analysis, the Irish publications database encompasses a total of 345,316 publications. Among these:

- 244,156 (70.7%) have been identified as Peer Reviewed.
- 226,025 (65.45%) have been classified as Open Access.

The baseline analysis has provided us with valuable initial insights into the Open Access landscape in Ireland. It serves as a solid foundation for our understanding of the current state of Open Access publication trends in the country. However, it's important to note that this is just the beginning. We are actively engaged in a process of continuous improvement, working diligently to refine our data and analysis.

The ongoing enhancements we are currently implementing will significantly enrich the Monitor Data, allowing for a deeper and more nuanced understanding of Open Access within Ireland. These improvements are not merely technical; they represent our commitment to delivering high-quality, reliable data that can inform decisions, policies, and strategies related to Open Access.

As we move forward, the Monitor's data will become an even more valuable resource for researchers, institutions, and policymakers. It will serve as a powerful tool for tracking the progress of Open Access initiatives, identifying areas for improvement, and fostering a culture of transparency and openness in Irish research landscape. Our dedication to enhancing data quality and expanding coverage ensures that the Monitor will continue to play a vital role in advancing Open Access in Ireland and beyond.

The final iteration of the report will build upon the data and insights shared here, offering a holistic perspective of Ireland's OA environment, including challenges identified and recommended solutions.

References

Kotitsas, S., Pappas, D., Manola, N., & Papageorgiou, H. (2023). *SCINOBO: a novel system classifying scholarly communication in a dynamically constructed hierarchical Field-of-Science taxonomy*. *Frontiers in Research Metrics and Analytics*, 8, 1149834.

 OpenAIRE
SCIENCE.SET FREE.



National Open Access Report for Ireland

Draft Presentation



OpenAIRE Team Representatives



Natalia Manola

CEO



Ioanna Grypari

Technical Project Manager



Leonidas Pispiringas

Scholarly Communication
Technical Expert



Harry Dimitropoulos

Technical Manager & Text-mining
Expert

Agenda & Goals of Presentation

- **Introduction**
 - Scope and structure of the report
 - Differences from final report
- **Baseline Analysis Approach**
 - Strategy and methods for assessment
- **Methods**
 - Overview of applied techniques
- **Data Evaluation Approach**
 - Rationale behind data assessment
 - On-going data work
- **Discussion & Feedback**

Introduction

SCOPE OF NATIONAL OPEN ACCESS REPORT

- **State of OA in Ireland**
 - Holistic overview and domain-specific insights
- **Data Accuracy & Viability**
 - Examination of Irish OA publishing outputs
- **Challenges, Solutions and Recommendations**
 - Documentation of obstacles and proposed long-term monitoring solutions, including workflows for data validation and enrichment.
- **Goal**
 - Foster a transparent and effective OA ecosystem in Ireland.

STRUCTURE OF THE FINAL REPORT

- **Executive Summary**
 - Main findings & insights
 - **Introduction**
 - Background of the study
 - Objective and purpose of the report
 - General Methodological Approach (High-level overview)
 - **State of Open Access (OA) in Ireland (Baseline Analysis)**
 - Scholarly Production
 - Open & FAIR
 - Plan S & Transformative Agreements
 - **Methodology**
 - Methodological specifics (concise)
 - **Data Evaluation & Challenges**
 - Evaluation of current Irish OA publishing outputs
 - Challenges observed: both high-level and data-specific
 - **Recommendations & Solutions**
 - Proposals for addressing challenges
 - Suggested workflows for data validation and enrichment
 - Long-term monitoring strategies
- Conclusion**

THE DRAFT REPORT

- **Executive Summary**
 - Main findings & insights
 - **Introduction**
 - Background of the study
 - Objective and purpose of the report
 - General Methodological Approach (High-level overview)
 - **State of Open Access (OA) in Ireland (Baseline Analysis)**
 - Scholarly Production
 - Open & FAIR
 - Plan S & Transformative Agreements
 - **Methodology**
 - Methodological specifics (concise)
 - **Data Evaluation & Challenges**
 - Evaluation of current Irish OA publishing outputs
 - Challenges observed: both high-level and data-specific
 - **Recommendations & Solutions**
 - Proposals for addressing challenges
 - Suggested workflows for data validation and enrichment
 - Long-term monitoring strategies
- ## Conclusion

Visualizations will be enhanced for clarity in the final version.

Baseline Analysis Approach

BASELINE ANALYSIS

- Three Sections
 - **Scholarly production**
 - **Open & FAIR**
 - **Plan S & Transformative Agreements**
- Visualisations and tables breaking down these aspect by distinct areas of interest.
- Each graph followed by discussion on observed patterns and their potential interpretations/implications.

SCHOLARLY PRODUCTION

- **Publication Insights for IE**
- **Scope**
 - Foundational exploration of scholarly production in Ireland.
- **Significance – Setting the stage**
 - Grounds the subsequent analysis in the larger narrative of Ireland's research landscape.
 - Provides granular insights into distributions across variables.
 - Essential to gauge the potential and challenges of transitioning to OA.

SCHOLARLY PRODUCTION

IE peer-reviewed publications

- Number (#), share out of all IE publications (%) and by

☐ year - *evolution*

☐ FoS level 1 – *domain disparities*

- #, %

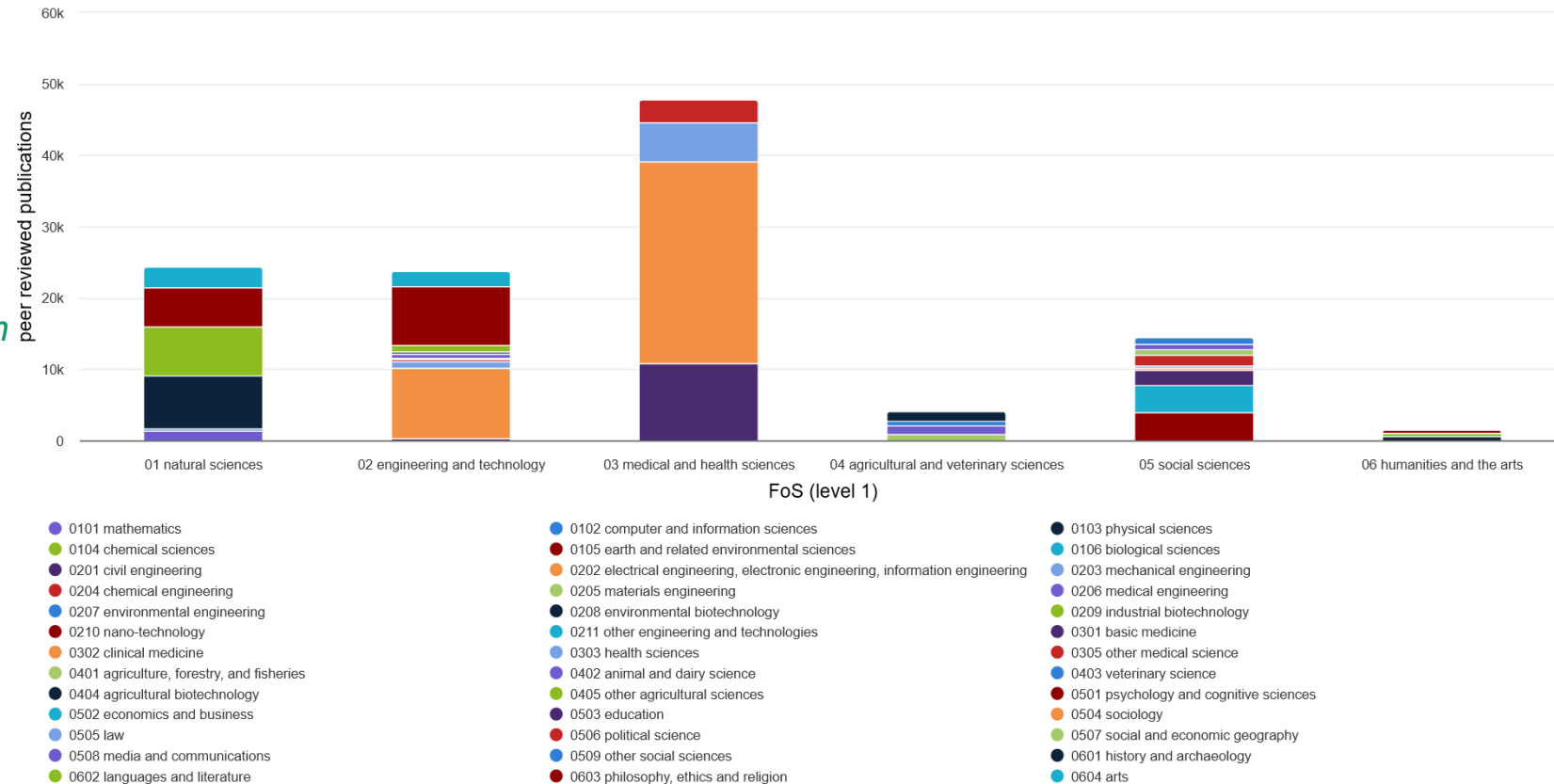
- FoS level 2 – *topic contribution*

☐ type of publication - *medium*

☐ data source type - *accessibility*

☐ SDG – *societal relevance*

Irish Peer Reviewed Publications
by FoS (levels 1 & 2)



OPEN & FAIR

- **Open & Fair Indicators**

- **Scope**

- Expansive analysis covering various dimensions of openness and fairness, including OA routes, licensing, PIDs, and more.

- **Significance – Setting the stage**

- Illuminate accessibility & transparency of publications in Ireland.
- Assess alignment with principles of fairness.
- Identify potential areas needing attention or intervention.

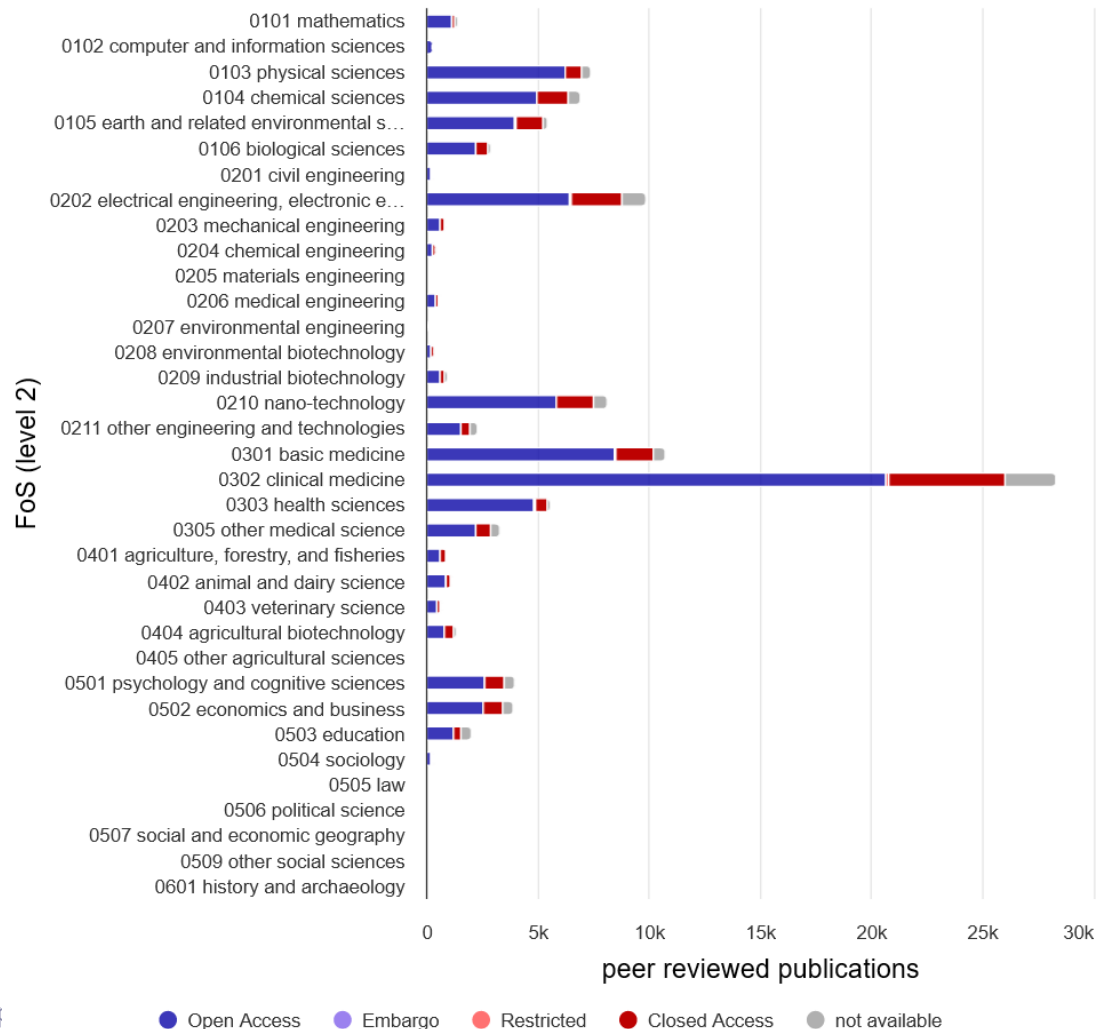
OPEN & FAIR

Access rights

- IE PR publications by access rights and #, %

- Over time - *evolution*
- FoS level 1 – *domain disparities*
- FoS level 2 – *topic contribution*
- type of publication – *medium*
- data source type - *accessibility*
- SDG – *societal relevance*

Irish Peer Reviewed Publications
by Access Rights by FoS (level 2)

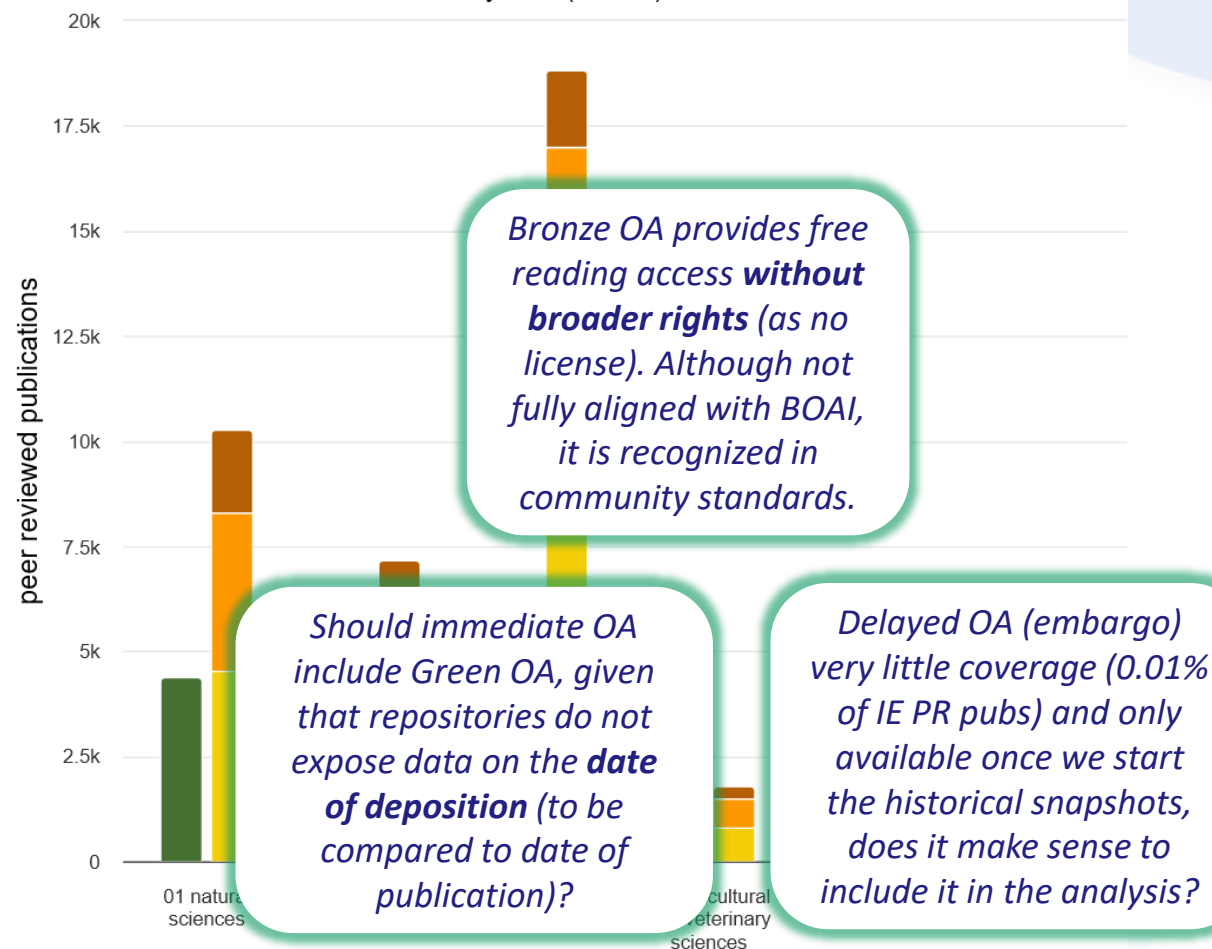


OPEN & FAIR

Open Access Types

- Repository mediated OA (Green) vs Publisher mediated (Gold, Hybrid, **Bronze**)
- Unrealized OA:
 1. non-OA & 2.
 2. Green but not Gold/Hybrid/Bronze
- Immediate OA (Gold, Hybrid, Bronze, **Green**) vs delayed OA (embargo)
 - Over time - **evolution**
 - FoS level 1 – **domain disparities**
 - FoS level 2 – **topic contribution**

Repository mediated (green) vs publisher mediated OA (gold, hybrid, bronze) by FoS (level 1)



OPEN & FAIR

What about non-CC licences?

FAIRness – Availability of

- CC licences, PIDs, ORCID iDs, abstract, funding reference in metadata, duplicates
- URL with access to PDF – **ACCESSIBILITY & INTEROPERABILITY**
 - Over time - *evolution*
 - FoS level 1 – *domain disparities*
 - FoS level 2 – *topic contribution*
 - type of publication – *medium*

Licence	# Irish OA peer-reviewed publications
CC-BY	30,561
CC-BY-NC-ND	13,679
CC-BY-NC	4,759
CC-BY-NC-SA	3,755
CC-0	664
CC-BY-SA	375
CC-BY-ND	97
CC-BY-ND-SA	22
Elsevier TDM	
Elsevier TDM	10,584
Elsevier OA	8,135
Wiley VOR	4,123
Springer TDM	5,618
Wiley TDM	4,611
arXiv non-exclusive distribution	3,055
Oxford Academic reuse	2,802
Sage TDM	2,595
CambridgeCORE	2,454
IEEE	2,377
APS	1,545
ASM TDM	1,042
% Irish OA peer-reviewed publications with non-normalised licenses	11.63%

PLAN S & TRANSFORMATIVE AGREEMENTS

Plan S

- Initiative by cOAlition S, SFI is a member
- Additional Irish Funders that have endorsed Plan S are:
 - Health Research Board
 - Irish Research Council
- Immediate and full open access to research outputs.
- Emphasizes transparency in publishing costs.
- Encourages publishing in fully open access journals, with allowances for transformative journals making the shift to full OA.
- Irish Plan S funders actively monitor compliance.

Transformative Agreements

- Agreements between (consortia of) RPOs and publishers, transitioning from subscription to open access.
- Negotiated by IReL
- Includes a range of journal types.
- Enables immediate open access publishing within existing subscription journals, bridging the gap between traditional and open access models.

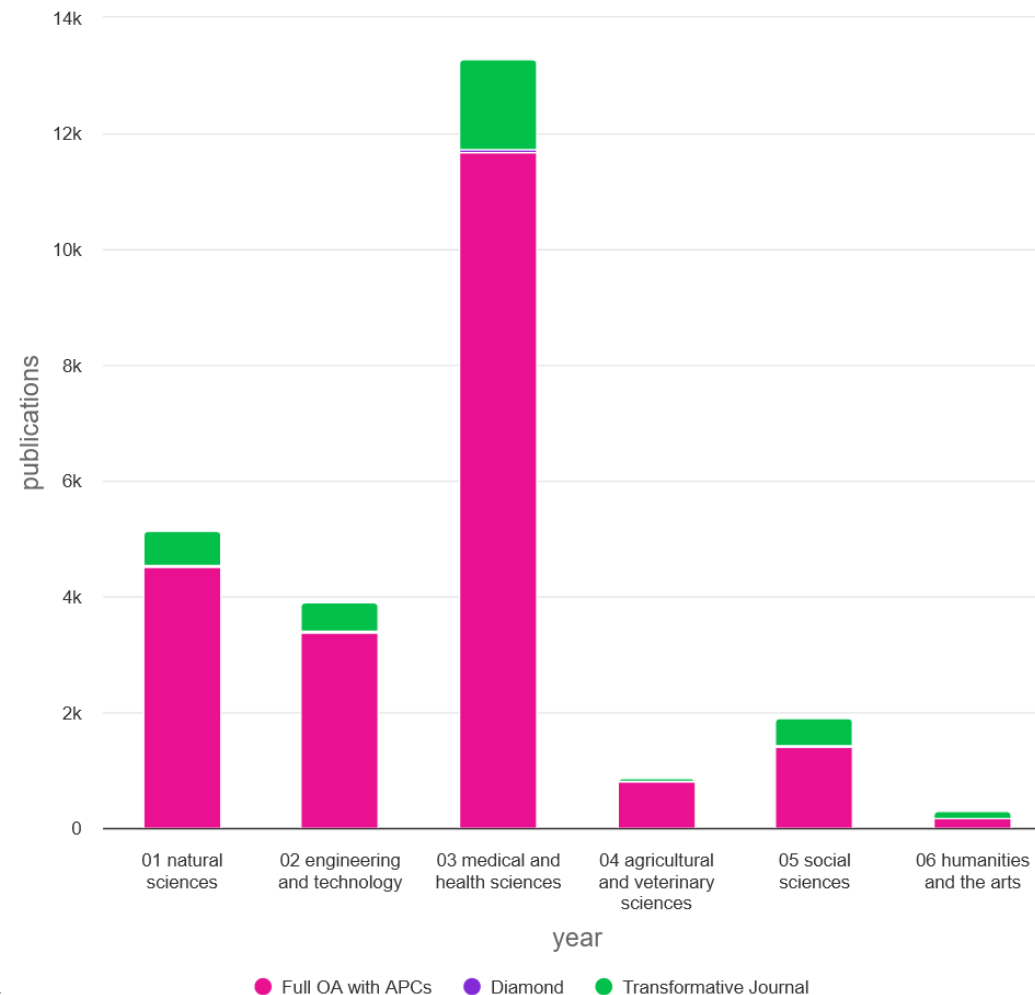
PLAN S & TRANSFORMATIVE AGREEMENTS

Plan S

- Grant-supported publications by Plan S funders
- IE PR publications in journals that are
 1. OA with APCs
 2. Diamond OA
 3. Transformative journals
 4. Journals that are part of transformative agreements
- APCs (total & average)
 - Over time - *evolution*
 - FoS level 1 – *domain disparities*
 - Journal – *expected impact*

Irish Open Access Peer Reviewed Publications

by journal business model, by FoS (level 1)



PLAN S & TRANSFORMATIVE AGREEMENTS

Transformative Agreements

- OA articles covered by transformative agreements using OpenAPC transformative agreement dataset.

❑ Over time - *evolution*

❑ FoS level 1 – *domain disparities*



IREL

WHAT IS A TRANSFORMATIVE PUBLISHING AGREEMENT?

OLD MODEL
(Paying to access)

- Traditionally universities would pay a subscription to publishers to get access to journal content
- That content was closed to anyone outside the university community

NEW MODEL
(Paying to publish)

- Open Access (OA) makes that quality research available to the world
- Funders want research to be Open Access
- Publishers require payment for this, called an APC - article processing charge
- Publishers often have a subscription fee and APCs

WHY ARE TRANSFORMATIVE AGREEMENTS IMPORTANT?

- Transformative agreements move us to a more equitable and sustainable model by accelerating the transition to OA and making it easier for authors

WHAT CAN YOU DO?

- **Researcher:** always look to make your research OA, even if the funding does not mandate it
- **Researcher:** your peer review is valuable - use it to encourage publishers to move to OA
- **Research Leaders:** encourage an OA culture
- **Library Staff:** advocate to your campus communities. Show how this can work

For more details see: <https://irel.ie/open-access>

Methods

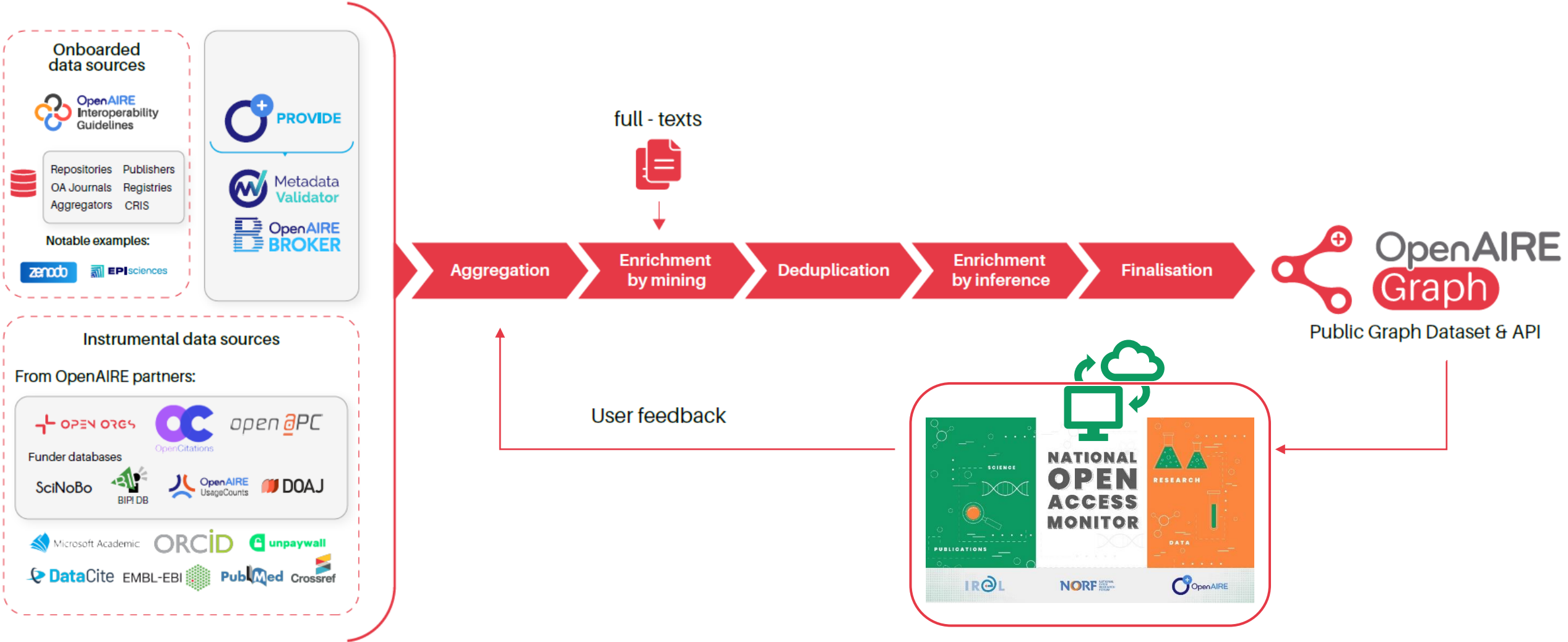
DATA BACKBONE: THE OPENAIRE GRAPH

- **A Scientific Knowledge Graph**
 - Timely and comprehensive coverage of research outputs
 - 129K data sources, 3M projects, and 235M research outputs (publications, software, datasets, other research products)
 - Near monthly updates
- **Precision, depth & processing**
 - Rigorous cleaning, deduplication, and enrichment for optimal accuracy
 - Enriched metadata links: Research results to projects, author affiliations, and classifications (FoS, SDG)
 - Techniques for duplicate identification in addition to OpenAIRE's curation tools
- **Robustness & openness**
 - Maintenance, load balancing, backups, overseen by the OpenAIRE technology centre
 - Open data & transparent methodologies



graph.openaire.eu

SEAMLESS INTEGRATION TO OPENAIRE AND EOSC



THE PUBLICATION SET OF THE MONITOR

We take the following steps

1. Identify Irish RPOs and their publications.
2. Identify Irish RFOs and their publications.
3. Exclude non-peer-reviewed publications
4. Make sure PIDs are used for Irish RPOs, RFOs and publications.

The final set for the Monitor currently includes 244.156 peer-reviewed publications.

1. IDENTIFY IRISH RPOS AND THEIR PUBLICATIONS

Publication affiliation in the OpenAIRE Graph

- Institutional data sources registered in OpenAIRE (repositories, CRIS, Open Access Journals)
- Metadata from harvested data sources such as Crossref.
- Inferred links via text mining.
- Links created via the claim and link functionalities in OpenAIRE EXPLORE

1. IDENTIFY IRISH RPOS AND THEIR PUBLICATIONS

Irish Repositories in OpenAIRE

- 98,151 publications recorded.
- 13 repositories use **outdated OpenAIRE Guidelines** → poor compliance with OS standards

Open Access (OA) Journals in OpenAIRE

- Not directly linked to institutions → for those only publishing institutional authors we may be **missing affiliations**

Institutional datasources urged to register with OpenAIRE PROVIDE and use the latest version of the OpenAIRE Guidelines

DRAFT: National Open Access Report for Ireland | 27 October 2023

2. IDENTIFY IRISH RFOS AND THEIR PUBLICATIONS

Funder – project – publication links in the OpenAIRE Graph

1. **Harvesting** links from repositories, OA Journals, CRIS systems.
2. Merging information from CrossRef's **Open Funder Registry** .
3. Collecting links from users via the “**link**” functionality.
4. Exchanging data with the **EC’s IT systems** for EC/FP funding.
5. **Text mining** of full text publications to identify the grants for 30+ funders that have joined OpenAIRE
 - **Science Foundation Ireland (SFI)** is one of them.
 - **We started working on Irish Research Council (IRC).**
 - **We are currently checking other Irish RFOs to see if they provide their project data openly on their site.**

3. EXCLUDE NON-PEER-REVIEWED PUBLICATIONS

Algorithm

1. DOI from Crossref
2. Publication types not only grey literature

Under consideration

- # of references above the minimum of the domain
 - Initial investigations show **too much variance** in numbers to meaningfully apply filter
- DOI from DataCite
 - PR journals that get DOIs from DataCite, however investigation shows this inclusion criterion **introduces noise**, considering alternatives.

Process automation in OpenAIRE graph → algorithm easy to adjust

4. USE OF PIDS

- Emphasis on **PIDs** for accurate monitoring of Irish publications.
- **Definition:** Irish scholarly publication = PID linked to an Irish organization.
- PIDs sourced from publication's metadata, content, and PID metadata.
- Comprehensive deduplication ensures consistent and wide coverage.

PID type	# Irish peer-reviewed publications
Publication PIDs	
Digital Object Identifier (DOI)	244,156
Handle	19,354
PubMed Central ID	29,772
PubMed ID	87,020
arXiv	1
Organisation PIDs	
Participant Identification Code	224,434
ISNI	373,695
OrgRef	364,789
Open Funder Registry	353,241
Wikidata	377,701
GRID	379,940
RingGold	25,781
ROR	380,093
OrgReg	363,754
ORCID IDs	
ORCID ID	116,584

DATA DISAMBIGUATION TECHNIQUES

- Deduplication in OpenAIRE
 - OpenAIRE Graph merges duplicate records of the same scholarly work. The deduplication process: <https://graph.openaire.eu/docs/graph-production-workflow/deduplication/>
- Organizations
 - Organizations in OpenAIRE are aggregated from different data sources
 - Deduplicated using the OpenOrgs platform
 - Designated RPO managers will have access
- Journals, Publishers and Licenses disambiguation
 - Disambiguate journals using the ISSN
 - Disambiguate publishers through Crossref metadata (ROR ids, DOI prefixes, et al)
 - Normalize licenses& custom text similarity algorithms

ENRICHMENT VIA TEXT MINING

- **Affiliation Matching:** Matches affiliations from documents with OpenAIRE organization database.
 - **Funding Classifiers:** Categorizes text from abstracts into predefined classes (e.g., funders, projects).
 - **Acknowledged Concepts Extraction:** Identifies grant identifiers, patents, bioentities, and custom concepts in publications.
 - **Metadata Extraction:** Uses CERMINE to extract plaintext, metadata (titles, authors, affiliations, etc.) from PDFs.
 - **FoS Classification:** NLP-based system classifying publications into FoS classes up to level 3.
 - **SDG Classification:** Aligns research with UN Sustainable Development Goals, pinpointing contributions to global challenges.
-
- PDF aggregation for IE complete (for current Graph)

74K (66%) of IE PR OA publications are accessible

INDICATORS

Attribute	Construction Methodology
Journal Business Models	
Open Access	Utilizing OpenAIRE's curated Gold ISSN list, Unpaywall metadata, and DOAJ journals.
Subscription	Journals without any open access articles.
Hybrid	Journals with open access articles that are not OA journals.
Transformative	We identify Transformative Journals by ISSN matching with the publicly available Transformative Journals data from Plan S initiative.
Journal APC Business Models	
Diamond OA	We obtain APC data from DOAJ using DOAJ's Public Data Dump (an exportable version of the journal metadata). We cross-reference this with the OpenAPC data integrated in the OpenAIRE Graph.
OA Types/Colours	
Green OA	An open access publication deposited in a repository
Gold OA	A publication published in an OA journal as defined above.
Hybrid OA	An open access publication published in a hybrid journal with an open license. At this point we consider only CC licenses "open". In principle, this means that we may be underestimating the number of hybrid OA articles and overestimating the number of bronze.

Attribute	Construction Methodology
Bronze OA	An open access publication published in a hybrid journal without an open license.
Accessibility – Interoperability	
Accessible	Accessible publications are sourced from OpenAIRE's full text collection, which holds PDFs of over twenty million OA publications. OpenAIRE's PDF aggregation system, recalibrated to prioritize Irish publications, examines the URL links in each publication's metadata to retrieve the corresponding PDF document. Given that multiple links can be associated with one publication, each is navigated. With ongoing automation, the system's coverage consistently expands, verifying if OA publications are accessible through thorough URL link inspection and PDF retrieval.
Interoperable	The construction of interoperable publications within the Monitor is intrinsically tied to their accessibility. Since we systematically fetch PDFs, any publication that is accessible through this process is also considered interoperable. In essence, the minimum threshold for interoperability is met when a publication's full text is accessible in a machine-readable format through the PDF aggregation system.

CORRESPONDING AUTHOR AFFILIATION

Identifying Corresponding Author Affiliation

- **Challenge:** Not a shared metadata element

Devised Methodologies

- **Contributor Rank Analysis using metadata**
 - Identify corresponding author based on position in the author list (Crossref metadata element)
 - Consider first author as corresponding if list isn't alphabetical.
 - Achieved identification in only 0.04% of Irish peer-reviewed publications using this approach.
- **Text Mining**
 - Employed exclusively for Irish publications.
 - Mine PDFs to discern the corresponding author's affiliation.
 - Only feasible for OA publications with available PDFs.

Comprehensive evaluation on final report.

PUBLICLY-FUNDED PUBLICATIONS

Definition

- "Publicly funded research" is research undertaken wholly or partly via public funding or remuneration (e.g., salaries, grants, contracts).

Steps Underway

1. Desk Research to discern public RPOs and RFOs.
2. IReL Survey: Currently being conducted.
3. OFR Metadata Integration, Funder types include: "Government" and "Private".
 - Private subtypes: academic, corporate, foundation, etc.
 - Government subtypes: federal, non-federal.
4. European Commission's Data: Metadata from project participant RPOs identifies public ones.

Methodological Concerns

- Indirect public funding may not be acknowledged in publication metadata (e.g., Private research institution gets public grants ([database?](#)), then finances a specific research project independently.)



GRANT ID's

Funders in OpenAIRE (SFI)

- Grant ID's available

Irish RFOs in OFR (Crossref)

- No projects registered with Crossref, we have funder-publications links
- Text mining module cannot be built without list of projects

We are working on creating funder-publication links for IRC, but with unidentified projects.

Data Evaluation Approach

DATA EVALUATION APPROACH

- For each metadata element pertinent to a high quality monitoring platform we fill out the following table.

Metadata element & Issue	Quality	OpenAIRE Actions	Stakeholder Actions

DATA WORK

- **Publications in 2022:** Enhance number of IE PR publications in 2022 (text-mining boost and additional data sources) – *doubled the number since draft report*
- **FoS/SDG:** Complete classification of all publications to FoS levels 1,2 and SDGs
- **Corresponding author affiliation**
- **Publicly funded publications**
- **Text mining for Irish funders**

- **On demand:** integration of institutional data sources & funder data

Discussion & Feedback

QUESTIONS

- Bronze OA provides free reading access **without broader rights** (as no license). Although not fully aligned with BOAI, it is recognized in community standards.
- Should immediate OA include Green OA, given that repositories do not expose data on the date of deposition (to be compared to date of publication)?
- Delayed OA (embargo) very little coverage (0.01% of IE PR pubs) and only available once we start the historical snapshots, does it make sense to include it in the analysis?
- What about non-CC licences?

Thank you!