

The Role of the Iscte Library in the Scientific Information Curation Process in the CRIS, *Ciência_Iscte*

Ricardo Reiçadas,^{1*} Teresa Segurado¹ and Vera Batalha¹

¹Iscte – University Institute of Lisbon, Portugal

ricardo.reicadas@iscte-iul.pt, teresa.segurado@iscte-iul.pt,

vera.batalha@iscte-iul.pt

Abstract

For the academic and scientific community whose institutions operate *Current Research Information Systems* (CRIS), it is essential that the scientific output deposited in CRIS is described responsibly and with high quality, in accordance with best practices and the principles of Open Science. In the academic context, science is not realized solely through research or publication phases, but also through the ways in which information is organized, described, evaluated, made available, reused, and preserved over time – tasks in which libraries play a central role.

Quality and responsibility in metadata description – ensuring the integration of scientific content into reliable, coherent, transparent, and interoperable information systems, in line with the FAIR Principles – are essential to academic and research assessment. These processes directly affect degree attribution, career progression, access to funding, performance evaluation, and personal and professional recognition by peers.

At Iscte - University Institute of Lisbon, the Library plays a key role in curating metadata related to researchers' scientific output in the institutional CRIS, *Ciência_Iscte*. By enhancing and enriching metadata and enabling integration with the Institutional Repository, the Library ensures accessibility and visibility of Iscte's scientific production and promotes the dissemination of knowledge to a broader audience.

This article aims to present and describe the scientific publication curation work carried out by the Iscte library within *Ciência_Iscte* and its integration with the Institutional Repository, supported by quantitative data that illustrate the work developed in this context during the 2020–2025 period.

* Corresponding author. Tel.: +351 210 464 254
E-mail address: ricardo.reicadas@iscte-iul.pt

Keywords: Metadata curation; *Ciência_Iscte*; Current Research Information Systems (CRIS); Institutional Repository; Academic libraries; Open Science

1 Literature review

1.1 CRIS

Current Research Information Systems (CRIS) have established themselves as essential infrastructures within the academic research ecosystem by enabling the integrated, structured, and standardized management of scientific information produced by researchers at higher education institutions. In a period marked by the exponential growth of scientific output and increasing demands for accountability, ethics, transparency, and evaluation, these systems support the organization of scientific production, including data on publications, projects, funding, and collaborations. In addition, they contribute to the monitoring of scientific performance, institutional evaluation, and the definition of more informed research strategies, while also strengthening personal and institutional recognition.

CRIS are defined by Sivertsen (2019) as “databases or other information systems used within and among research organizations to store, manage, and exchange data for documentation, communication, and administration of research activities” (p. 667). Clements and Proven (2015) also describe CRIS as “a single portal bringing corporate and academic research activity together, reducing duplicate data entry, increasing data quality, identifying authority sources of information and recording complex relationships between researchers, projects, outputs and impact” (para. 4).

One of the most relevant aspects of CRIS is their capacity to interoperate with other information systems, particularly Institutional Repositories (IR). This interoperability, often based on shared standards and conceptual models such as CERIF, ensures the efficient circulation of information between systems, avoids redundancies, promotes consistency of records, and facilitates the dissemination of scientific output in accordance with Open Science principles. Integration between CRIS and IR thus enables the optimization of workflows, improves the visibility of scientific production, and enhances the quality, integrity, and reliability of metadata and research data, thereby increasing the impact, reuse, and long-term preservation of research outputs. Although originally developed with distinct objectives, technological advances and Open Science requirements have fostered a progressive convergence between these two systems, evolving towards increasingly deep and complementary interoperability (Schöpfel & Azeroual, 2021).

As a result, interoperability between CRIS and IR has led to deep changes in the functioning of repositories themselves. As self-archiving has gradually been replaced by Open Access or embargoed deposits made directly from CRIS, in compliance with national and international Open Access policies, workflows between the two systems have become smoother and more efficient (Biesenbender et al., 2019). Following the curation and validation of scientific publication metadata in the CRIS, records are transferred directly to the IR, which assumes responsibility for assigning handles and ensuring discovery, accessibility, dissemination, and long-term preservation not only of scientific publications but also of research data. This model highlights how several functions traditionally attributed to IR have increasingly been assumed by CRIS, while the repository remains the deposit infrastructure and the primary point of access and dissemination for institutional scientific output (de Castro et al., 2014; de Castro & Puuska, 2023).

The volume and diversity of metadata collected by CRIS demonstrate their importance in organizing scientific production, as they encompass titles, abstracts, authorship and affiliation data, publication details, research software, data and samples, instruments, funding, and information on participating organizations and entities (Barcelona Declaration on Open Research Information, 2024;

Schöpfel & Azeroual, 2024). This integrative capacity reinforces the role of CRIS as central pillars of metadata management by ensuring coherence, quality, and visibility of scientific output, supporting institutional strategies and public policies, and sustaining an integrated and sustainable approach to research management.

Within this framework, academic libraries play a central role in the curation of scientific publications and research data within CRIS. Beyond their traditional functions, libraries – owing to their expertise and experience in bibliographic record management – assume responsibility for the validation, standardization, and enrichment of metadata, ensuring data quality, reliability, interoperability, and reuse (Bryant et al., 2017).

This role is intrinsically linked to compliance with the FAIR principles of Open Science, which promote free and Open Access to knowledge, transparency of scientific processes, and the reuse of research data. Through their involvement in CRIS, libraries play a decisive role in the curation of scientific information by ensuring the validation, standardization, and enrichment of metadata, as well as compliance with ethical standards related to privacy and the protection of personal data. By applying standards, controlled vocabulary, and best descriptive practices, libraries contribute to ensuring that scientific publications and research data are easily findable, interoperable, accessible, and reusable. In this way, academic libraries assume an essential role in ensuring the quality, consistency, and interoperability of metadata, contributing to more robust and reliable CRIS (Schöpfel & Azeroual, 2021, 2023, 2024).

The quality of metadata description is one of the core elements of the effective functioning of CRIS. Only complete, accurate, and consistent information can underpin reliable processes of evaluation, analysis, and decision-making in the institutional context (Azeroual & Lewoniewski, 2020; Sivertsen, 2019). Enhancing metadata quality depends largely on curation practices capable of integrating information from multiple sources, such as IR, funding platforms, bibliographic databases, and persistent identifiers. This process also entails responsibility regarding ethical and legal issues related to privacy, confidentiality, and consent management, alongside the rigorous application of the FAIR principles.

In this sense, the *FAIRification* of CRIS emerges as a strategic necessity for the sustainable, transparent, and reliable management of institutional scientific output. As infrastructures that collect, store, process, and interlink information from multiple sources – often including sensitive and strategically relevant data – CRIS must ensure adequate protection mechanisms and, where necessary, data anonymization. The increasing convergence of these systems with other research infrastructures makes it even more essential to ensure not only the quality and reliability of metadata, but also the transparency, interoperability, and robustness of the systems themselves (Azeroual et al., 2022).

1.2 The role of libraries in metadata curation

Metadata curation undertaken by academic libraries – covering both scientific publications and research data – is a central element in ensuring the quality, visibility, and impact of academic research. The production of knowledge is no longer sufficient in itself; it is essential that research outputs are accurately described, organized, preserved, and shared. Well-structured metadata – such as information on authorship and institutional affiliation, titles, keywords, scientific fields, persistent identifiers (ORCID, DOI), publication versions, and funding data – are crucial to ensuring that scientific results are easily discoverable, accessible, and correctly integrated into other systems and platforms. For researchers, this is reflected in greater recognition of their scientific output, increased visibility and impact, and more effective integration into academic and scientific collaboration networks.

In the domain of research data, curation assumes an even more critical role. Ensuring that data is adequately described, standardized, and preserved guarantees their reliability, interoperability, and

potential for reuse over time. This process not only supports the development of new research projects but also helps to avoid duplication of effort and the unnecessary creation of new datasets.

Johnston et al. (2018) defines data curation as “the encompassing work and actions taken by curators of a data repository in order to provide meaningful and enduring access to data” (p. 5). Yoon and Donaldson (2019) reinforce this perspective, understanding data curation as “the active and ongoing management of data through its lifecycle to best serve the needs of current and future users” (p. 812).

This concept emerged “in the context of growing pressure to increase access to scientific research output and to improve the overall efficiency of the technological and human resources necessary to facilitate data creation, preservation, and reuse” (Tammaro et al., 2019, p. 89).

Within this context, academic libraries and librarians assume a strategic and increasingly specialized role in the scientific research ecosystem. Their activities clearly go beyond the traditional management of collections and information access services, shifting towards qualified mediation throughout the entire lifecycle of scientific information. According to Tammaro et al. (2019), the responsibilities of academic libraries include:

- The standardization and quality control of metadata, including correction, enrichment, and validation;
- The application of standard metadata schemes and controlled vocabularies (e.g. Dublin Core, CERIF, MeSH);
- Ensuring the deposit and preservation of scientific publications and research data in the IR;
- The assignment of persistent identifiers;
- Ensuring the anonymization and security of sensitive data, in compliance with applicable legislation;
- Ensuring compliance with the FAIR principles, so that data are easily findable, accessible, interoperable, and reusable in the future.

However, academic libraries face significant challenges in carrying out curation tasks, arising both from the increasing complexity of the scientific information ecosystem and from the technical, financial, organizational, and ethical demands associated with it. According to Zareef and Jabeen (in press), the main constraints include:

- A shortage of staff qualified in metadata schemes, digital preservation, copyright, standardized formats, and the protection of sensitive data;
- Lack of funding, which limits the acquisition and updating of technological infrastructures, storage capacity, and specialized software;
- Technological obsolescence requires continuous maintenance of platforms, formats, and digital systems.

Another relevant limitation concerns the availability and level of specialization of human resources in academic libraries. Although professionals with expertise in data curation and scientific information management do exist, many libraries lack sufficiently sized teams or staff who can devote themselves exclusively to these tasks or deepen the necessary specialization in specific scientific

domains. The accumulation of functions, combined with a lack of continuous training, often compromises the consistency and quality of curation activities.

In summary, despite the challenges and constraints identified, academic libraries play a fundamental role in the scientific research ecosystem. Through their curation practices, they ensure high-quality metadata, promote interoperability between systems, contribute to the long-term preservation of information, and support compliance with the FAIR principles. In doing so, they make a decisive contribution to more visible, transparent, and reusable science, strengthening the value of scientific production infrastructures and sustainable management of research.

2 Evolution of Open Science policies and movements (2020-2025)

In the European context, Open Science has consolidated its position as a strategic pillar of science policy, promoting profound transformations in the models of research production, communication, and evaluation. Grounded in principles of transparency, accessibility, reproducibility, and reuse, this approach has driven the adoption of Open Access, research data management practices, and the development of open infrastructures throughout the scientific lifecycle.

Between 2020 and 2025, the strengthening of public policies, the consolidation of the FAIR principles, and the evolution of Open Access models decisively shaped both the European and national contexts. The European Commission reinforced requirements for the immediate availability of publications, data, software, algorithms, protocols, and other research outputs, emphasizing the importance of verification, reproducibility, and the responsible management of scientific results in accordance with the FAIR principles, thereby promoting openness across the entire research lifecycle (European Commission, n.d.).

A major milestone during this period was the implementation of Plan S, which established that “from 2021, all scholarly publications resulting from research funded by public funds must be published in Open Access journals or platforms that comply with the established requirements or be made immediately available through Open Access repositories without embargo” (cOAlition S, n.d.).

In 2023, cOAlition S ended its financial support for Transformative Agreements, considering that these arrangements perpetuated the hybrid Open Access model. This decision sought to reduce reliance on subscription-based publishing and to accelerate the transition to immediate Open Access. Despite the withdrawal of financial support, Transformative Agreements continue to meet the requirements of Plan S, with decisions regarding their continuation delegated to individual countries. In Portugal, Open Access publishing conditions under these agreements were negotiated for the 2025–2027 period.

Between 2020 and 2023, Europe witnessed a strengthening of national and institutional Open Science policies, supported by strategic plans that emphasized FAIR data, repository infrastructures, and research evaluation. In Portugal, the Fundação para a Ciência e a Tecnologia (FCT) adopted a new Open Access Policy in February 2025, mandating Open Access availability for all scientific publications resulting from research funded by the organization.

In 2023, Iscte also approved its *Política de gestão e partilha de dados de investigação*, expressing its commitment to scientific integrity, the promotion of innovation, and the generation of social impact, while contributing to international discussions on responsible and efficient research data management practices.

In 2024, the Barcelona Declaration on Open Research Information was published, advocating transparency, openness, and a reduction in dependence on proprietary infrastructures that restrict access to, reuse of, and auditing of scientific information.

In Portugal, the *Re.Data* consortium was established in 2025 with the mission of setting up and operationalizing a national support network for Research Data Management. Its objective is to promote a culture of data management and openness aligned with the FAIR principles and the European Open Science ecosystem, particularly the European Open Science Cloud.

Overall, during the 2020–2025 period, the most significant transformations in Open Science and research data management materialized in the strengthening of public policies and institutional frameworks oriented towards openness and transparency in the scientific process. This period was also marked by the progressive elimination of hybrid access models for scientific output, the consolidation of FAIR practices, the growing adoption of artificial intelligence in research processes, and the consolidation of open scientific infrastructures. At the same time, national policies were reviewed and updated with a view to fully integrating scientific openness as a structural axis of research systems, promoting greater coherence between practices, infrastructures, and research assessment mechanisms.

3 Methodology

This research uses a mixed methodology, combining a qualitative analysis of processes with a quantitative study of institutional data. The work began with a literature review using international reference databases like *Web of Science* and *Scopus*. It focused on the importance of Current Research Information Systems (CRIS) and the role of libraries in curating metadata for scientific publications. The study focuses on *Ciência_Iscte*. We have analyzed the library's curation model by creating a flowchart to explain internal validation and how it integrates with the Institutional Repository (IR).

The quantitative part of the study is based on the library's own records of curation work done within the *Ciência_Iscte* system. To this analysis, we reviewed a total of 5,398 records between 2020 and 2025. We specifically selected the "Scientific Journal Article" category because it has the highest volume of records and requires the most detailed curation. This choice was also based on the fact that journal articles are the most valued document type in the Iscte internal Regulation that governs the evaluation of the work of researchers employed by the University (2020), and that directly impacts the researchers' career progression and institutional assessments.

The data shows the significant scope of the CRIS system, which managed 86.3% of all documents deposited in the repository during this period. By identifying versions compatible with Open Access, the curation team was able to move a high volume of research into IR. We have used Microsoft Excel for statistical analysis and to create charts. The analysis focused on the following indicators: Curation Outcomes: The number of records reviewed, archived, or rejected. Document versions: A comparison between the *Version of Record* (72.1%) and *Author Accepted Manuscript/Post-print* (27.9%) version. Access Types: The distribution between Open Access (80.6%) and Embargoed Access (19.4%). This data allows us to measure how the library's intervention improves the quality of scientific information and helps the university meet Open Science requirements.

4 CRIS: *Ciência_Iscte*

Ciência_Iscte, the *Current Research Information System* of Iscte – University Institute of Lisbon, developed and maintained by the institution's Information Systems Development Services, has, since its inception, established itself as a strategic instrument for the organization, management, and enhancement of research produced at the institution.

This system enables each researcher to manage their own scientific activity by submitting scientific outputs, centralizing and structuring this information within a single system. In this way,

duplication of records is avoided, the consistency of scientific information is improved, and the time spent by both support services and researchers themselves is significantly reduced, allowing researchers to focus more effectively on their research activities.

The information added to *Ciência_Iscte* is subsequently automatically disseminated across multiple systems with which the platform is integrated, ensuring the interoperability and consistency of scientific information. This process operates in a transparent and integrated manner, requiring no additional intervention from researchers, thereby contributing to the optimization of workflows, the reduction of redundancies, and increased efficiency in the management of scientific information (Lopes, 2019).

Regarding datasets, the transition implemented within *Ciência_Iscte* emerged in response to the requirements of funding bodies such as Fundação para a Ciência e a Tecnologia (FCT) and the European Commission. The Iscte *Policy for the management and sharing of research data* (2023) establishes, as a requirement the deposit, within the Iscte community of the Zenodo repository, of the data essential for the validation of the scientific results produced.

This model allows the institution to benefit from a robust infrastructure while maintaining control over research outputs through an integrated workflow. As noted by Inácio et al. (2023), the information circuit is based on the export of metadata from Zenodo to *Ciência_Iscte*, ensuring that data are recorded only once and remain visible both in researchers' profiles and in the IR. However, this automation does not eliminate the need for human verification by the library, which ensures compliance with quality requirements and institutional standards, both at the level of metadata and of the deposited files.

This logic of simplification provided by CRIS is reinforced by Joint (2008), who states:

In the dual approach, it is essential to share data between the two local systems, to avoid exasperating researchers by asking them to deposit in two places at the same time, and to avoid the workload of checking and editing identical self-archived submissions on two systems (p. 573).

By providing structured, standardized, and continuously updated information on the scientific activity produced at Iscte, *Ciência_Iscte* establishes itself as a central instrument supporting the strategic management of research. The system enables the institution to obtain an integrated and reliable overview of its scientific output, facilitating internal planning, monitoring, and evaluation processes, both at the institutional level and within the context of its Research Units and projects.

At the same time, *Ciência_Iscte* contributes decisively to strengthening the national and international visibility of research carried out at Iscte by ensuring the organization, coherence, and dissemination of scientific information. Through its integration with IR, making this information accessible and reusable, it supports the recognition and enhancement of scientific output and researchers alike, promotes their integration into academic and scientific networks, and strengthens Iscte's institutional position within the scientific and higher education system.

5 Library curation of scientific publication metadata at Iscte

This article highlights the importance of the metadata curation and validation work carried out on scientific journal articles authored by researchers at Iscte – University Institute of Lisbon, developed by the institution's Library. This work is essential to ensure the quality, reliability and consistency of metadata in *Ciência_Iscte* and in other internal and external scientific information systems, ensuring interoperability between systems. By ensuring metadata standardization and interoperability, it contributes to the visibility of institutional scientific output and alignment with the FAIR principles,

thereby reinforcing Iscte's commitment to Open Science and the sustainable management of scientific information.

Over recent years, the library has consolidated an internal curation model that simultaneously responds to the requirements of *Ciência_Iscte*, to the legal responsibilities associated with the deposit of scientific publications in the IR, and to researchers' needs in terms of the recognition and impact of their research. This model is based on a structured workflow – illustrated in Figure 1 – which is documented and supported by defined quality criteria.

Within this framework, the library assumes responsibility for verifying and validating bibliographic information, with particular attention to author affiliation and metadata accuracy, using reliable sources such as journal websites, publication PDF files, as well as reference databases and services such as *Scopus*, *Web of Science* and *Crossref*. Among these, *Crossref* plays a central role by providing standardized and persistent DOI-based metadata, which support the normalization and enrichment of bibliographic records, ensure the reliable identification of publications, and facilitate interoperability and citation tracking within *Ciência_Iscte*.

As a result of this curatorial intervention, descriptive metadata are systematically completed and enriched, ensuring that the information is properly validated, consistent, and harmonized in accordance with widely recognized international standards and models (such as Dublin Core), controlled vocabularies (*EuroVoc*, *ERIC*, *MeSH*, *Library of Congress Subject Headings*), the *Fields of Science and Technology (FOS)* classification, and authority lists.

This standardization is essential to ensure data interoperability, particularly through the adoption of the OAI-PMH protocol and compliance with the *OpenAIRE* guidelines. These frameworks enable the harvesting, aggregation, circulation, and reuse of metadata by national and international repositories and research infrastructures, in accordance with Open Science guidelines.

These descriptive metadata include, among other elements, the name(s) of the author(s) and the publication title described according to APA style, the publisher, abstract, keywords, bibliographic information, persistent identifiers (ORCID, ISBN, ISSN, DOI, *WoS/Scopus*), funding data, and the identification of different versions of the document.

The qualitative and structured description of metadata constitutes an essential basis for improving the discoverability, accessibility, and interoperability of research output. Complete, consistent, and semantically enriched metadata allows for precise identification of resources, facilitate their retrieval across different information systems, and ensure effective linking between publications, authors, institutions, and funding sources. In this sense, metadata curation plays a central role in the effective implementation of the FAIR principles (Azeroual et al., 2022), promoting the visibility, reuse, and impact of scientific knowledge within the Open Science ecosystem.

One of the most important aspects of exporting records from *Ciência_Iscte* to the IR concerns the correct identification and selection of the document version to be deposited – namely the *Author Accepted Manuscript/Post-print* or the *Version of Record* – to ensure compliance with publisher policies and Open Science requirements. In addition, it is essential to clearly define access conditions, particularly regarding immediate Open Access availability or embargo periods. *Creative Commons* licenses play a decisive role in this context, as they clearly specify permitted reuse conditions, thereby supporting copyright protection and the dissemination of scientific output.

All stages of the scientific output curation process developed by the Iscte library are illustrated in the corresponding workflow diagram.

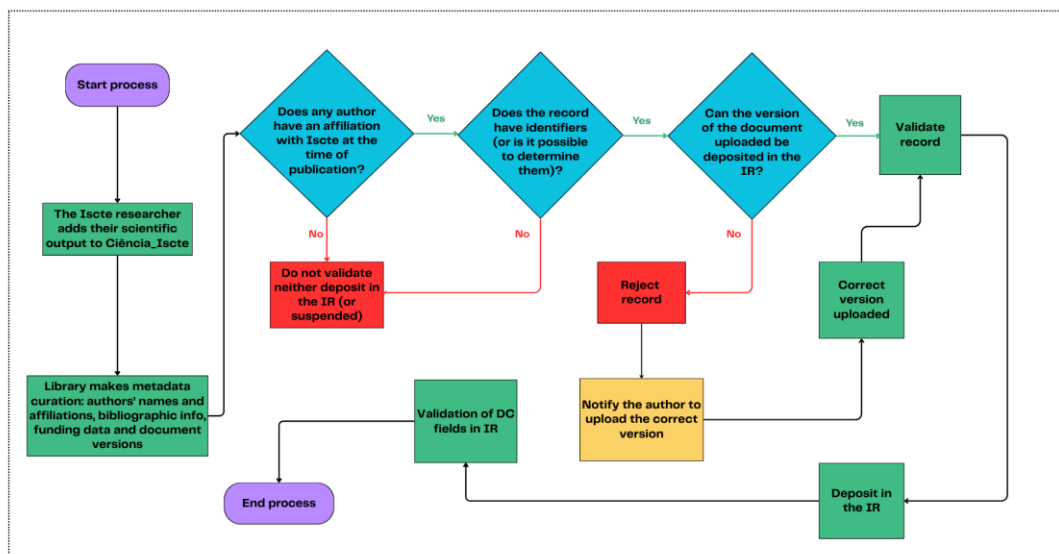


Figure 1: Diagram of the scientific publication's curation process at *Ciência_Iscte*; Source: Own authors

Interoperability between *Ciência_Iscte* and the IR is achieved through processes of metadata mapping, export, and synchronization to the corresponding Dublin Core fields within the IR, which is based on the DSpace platform. At the same time, this mapping ensures the correct association of records with the appropriate author's community and document typology. In addition to metadata, the full-text file of each publication is also exported (Amante et al., 2015).

This process operationalizes the principle of “one input, many outputs”, whereby metadata are added once into *Ciência_Iscte* and subsequently disseminated automatically and synchronously to multiple systems and platforms, including the Iscte Institutional Repository, researchers' profile pages, and those of their respective research centers, as well as national and international scientific infrastructures and systems, like *Sherpa Romeo*, *Crossref*, *Scopus*, *Web of Science* and *OpenAIRE*. This ensures consistency, interoperability, and visibility of institutional scientific output.

Within the IR, the scientific output curation process is completed through the validation of Dublin Core fields related to funding data, using the *OpenAIRE* API to establish links between publications, projects, and funders, thereby ensuring system interoperability. This process also includes the validation of journal titles, *FOS* classification, and the *Sustainable Development Goals (SDGs)*, based on authority lists. In addition, bibliographic references are included and standardized in accordance with APA style (7th edition), and PDF files are normalized according to document type.

The automation provided by *Ciência_Iscte* has contributed to a substantial simplification of the IR deposit workflow, reducing the need for direct contact with authors and researchers. Nevertheless, the library seeks to ensure that its intervention does not replace, but rather complement, the role of researchers in managing their scientific output. Researchers remain responsible for adding their scientific outputs to *Ciência_Iscte* and are notified whenever their intervention is required in a given record, particularly in cases where the version of the document submitted in *Ciência_Iscte* for deposit in the IR is rejected for not complying with publishers' archiving policies. In such cases, researchers are notified to upload the correct and eligible deposit version, ensuring compliance with editorial policies and Open Access availability whenever possible.

The library also provides regular support to researchers on various issues related to the validation and rejection of records in *Ciência_Iscte*. Among the most frequent queries are those concerning the

distinction between different document versions – particularly between the *Author Accepted Manuscript/Post-print* and the *Version of Record* – an essential aspect for ensuring compliance with publisher policies and applicable Open Access requirements. Questions regarding publication dates are also common, especially when it is necessary to distinguish between online publication dates and final publication dates, ensuring accurate description and record compliance. This support work is articulated with the training provision promoted by the library, which develops training sessions and guidance materials within the scope of Open Science, aimed at strengthening scientific literacy and clarifying the legal, practical, and ethical implications associated with the deposit, dissemination, and reuse of scientific output.

At the same time, the growing availability of Open Access publishing routes – whether through Transformative Agreements or the diamond route – has contributed to a significant increase in the number of publications added to *Ciência_Iscte* and deposited in the IR, resulting in greater visibility, accessibility, and impact of Iscte's scientific output. In this context, the integration between *Ciência_Iscte* and the IR plays a decisive role in the efficiency of this process: by facilitating content transfer and validation, it has helped to increase the number of deposits of scientific and peer-reviewed journal articles, thereby contributing to the democratization of access to knowledge.

6 Quantitative analysis

The implementation of the *Ciência_Iscte* (CRIS) system in 2013 was a major turning point for the Iscte's Repository. Before 2013, the number of documents added each year was relatively low and inconsistent. However, once the new system and professional curation began, there was an immediate and significant increase in the records archived.

The effect of the new management strategy became clear in 2014. The number of deposits more than doubled, jumping from 272 in 2013 to 558 in 2014. This growth was driven by a specific focus on scientific journal articles. By identifying versions of documents compatible with Open Access, the curation team was able to move a much higher volume of research into IR.

Between 2014 and 2025, the Repository maintained a high and stable level of activity. During this period, the annual count never dropped below 460 documents, reaching a peak of 1,217 deposits in 2017. These figures show that the integration between the CRIS system and the IR successfully created a consistent workflow for preserving and sharing the university's scientific output (Table 1).

Year	Frequency (n)
2006	21
2007	133
2008	343
2009	145
2010	32
2011	162
2012	230
2013	272
2014	558
2015	757
2016	649
2017	1,217
2018	552
2019	466
2020	650
2021	798
2022	694
2023	800
2024	897
2025	718
Total	10,094

Table 1: Annual distribution of documents deposited in the Iscte Repository (2006–2025); Source: IR

Table 2 shows that the *Ciência_Iscte* (CRIS) provides most of the repository's records. From 2020 to 2025, 86.3% of all archived documents were managed through this system.

The percentage of records from this system remained above 69.7% of the yearly totals throughout the period. In 2023, the CRIS system accounted for 95.9% of all deposits. This indicates that the documents added that year resulted from the curation and export process of the CRIS platform.

In the last six years, out of 4,557 total records, 3,933 came from the *Ciência_Iscte* system. This data indicates the system's role in repository performance and the archiving of university research.

Year	Total Records (n)	CRIS-Sourced (n)	% of Total
2020	650	454	69.8
2021	798	735	92.1
2022	694	612	88.2
2023	800	767	95.9
2024	897	727	81.0
2025	718	638	88.9
Total	4,557	3,933	86.3

Table 2: Proportion of records managed via *Ciência_Iscte* (2020–2025); Source: Curation database record

From 2020 to 2025, a total of 5,398 records were managed. As Figure 2 shows, most of these records were successfully achieved, while a much smaller number was rejected. This high success rate is tied to the *Ciência_Iscte* (CRIS), which provided 86.3% of all the documents in this period. The system was especially active in 2023, when it accounted for nearly 96% of all deposits. These figures show that the current workflow is effective at processing and preserving the university's research output.

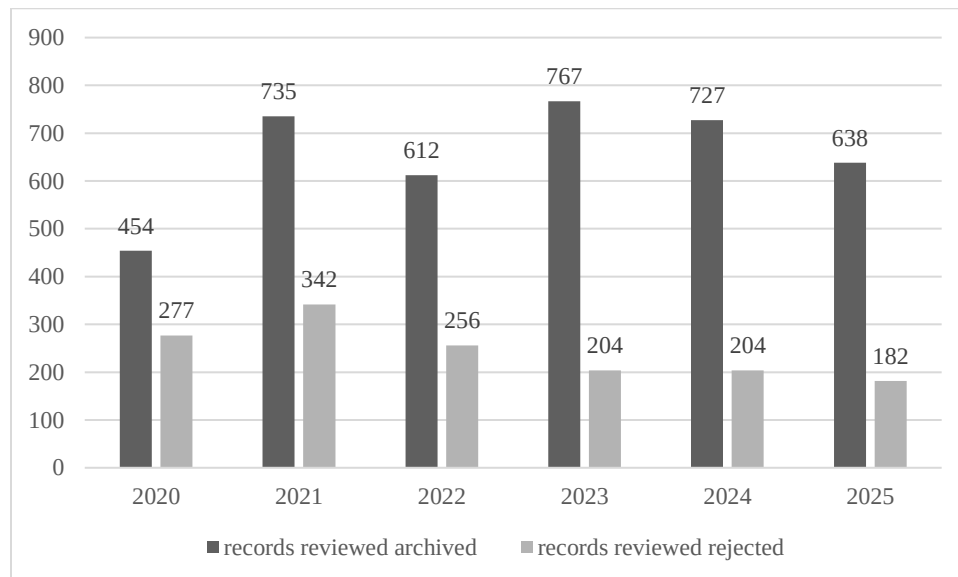


Figure 2: Annual distribution of managed records by curation outcome (2020-2025) (n=5,398); Source: Curation database record

Regarding the access options, the dataset is divided into two distinct categories based on availability. Open Access accounts for 3,170 records, representing 80.6% of the total. This high percentage indicates that most of the university's research is immediately available to the public without barriers. Embargoed Access comprises 763 records (19.4%), which applies to documents currently under a temporary restriction, typically imposed by publisher policies, that will automatically transition to Open Access once the designated period expires. Together, these two categories account for the total of 3,933 records managed through the *Ciência_Iscte* system.

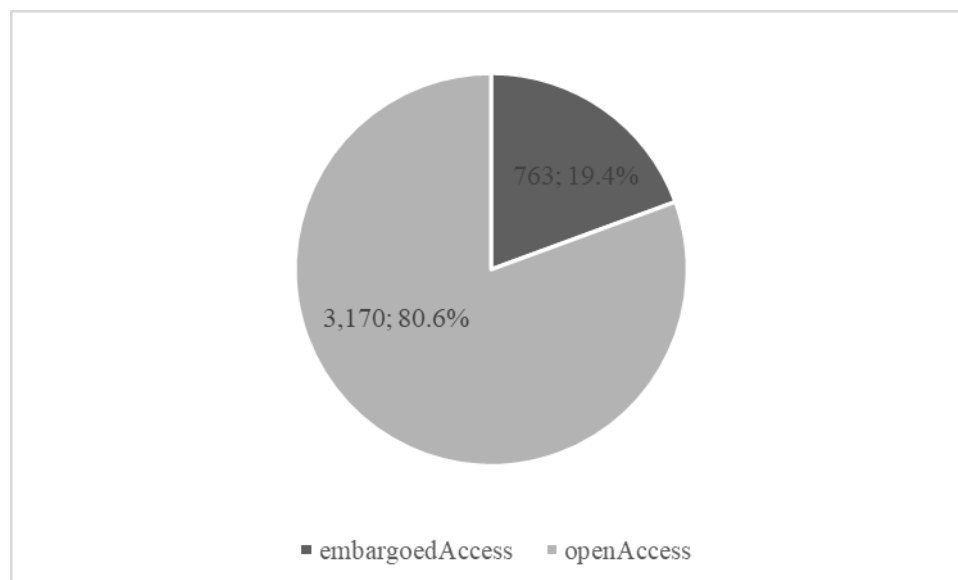


Figure 3: Distribution of records by Access Type (n=3,933); Source: Curation database record

The archived records are categorized by the technical version of the document deposited in the repository. The *Version of Record* (VoR), the final publisher-formatted version of the article, accounts for 2,837 records, representing 72.1% of the total. The prevalence of this version indicates a high frequency of deposits allowed through *Gold Open Access* models or institutional agreements with publishers.

The remaining 1,096 records (27.9%) consist of *Author Accepted Manuscript/Post-print* versions. These documents contain peer-reviewed content but lack the final publisher typesetting and layout. These versions are typically used to fulfill *Green Open Access* requirements when the publisher restricts the archiving of the *Version of Record*. Together, these two categories comprise a total of 3,933 records analyzed in this period.

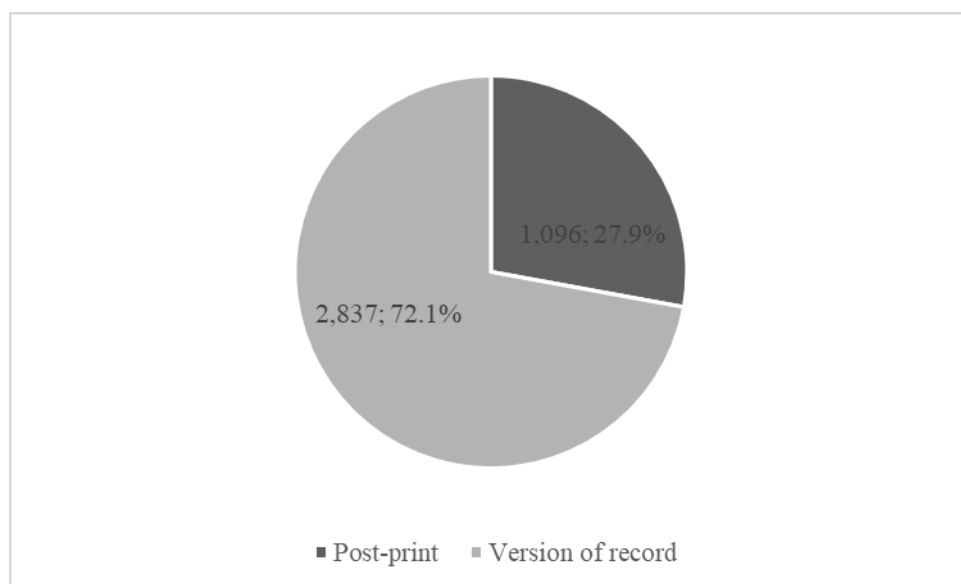


Figure 4: Distribution of records by Version Type (n=3,933); Source: Curation database record

7 Conclusion

Current Research Information Systems (CRIS) have become a key component of the contemporary research information ecosystem, particularly when integrated with IR. This integration is strategically significant for ensuring the quality, visibility and reuse of research outputs, while promoting interoperability and alignment with the FAIR principles and Open Science practices.

At Iscte – University Institute of Lisbon, the *Ciência_Iscte* system plays a central role in institutional research management. By centralizing and standardizing scientific research outputs, the CRIS supports transparent, rigorous and equitable evaluation processes, while providing a reliable and authoritative source of scientific information. Its integration with the IR enhances the visibility of institution's scientific output and facilitates its dissemination within international research networks, in accordance with Open Science best practices.

The reliability of CRIS-based workflows is critically dependent on data quality. In this context, metadata curation performed by the library within *Ciência_Iscte* is a key operational and strategic activity. The adopted curation model, based in international metadata standards and reliable information sources, ensures the quality, consistency, and reliability of the metadata. This approach

not only supports the dissemination and long-term preservation of research outputs but also ensures compliance with Open Access policies and the FAIR principles.

Between 2020 and 2025, the library team analyzed 5,398 records, with 86.3% of IR deposits managed through the CRIS. This high level of integration enables a consistent and controlled workflow, contributing to sustained high standards of scientific information quality. The curation process directly addresses Open Science requirements through the identification of appropriate document versions for deposit. Of the analyzed records, 72.1% correspond to the *Version of Record* and 27.9% to the *Author Accepted Manuscript/Post-print*, ensuring compliance with Open Access mandates while maximizing the visibility and impact of institutional research outputs.

Beyond its technical dimension, metadata curation has a significant institutional impact. Validated CRIS data underpin researcher evaluation processes and directly influence career progression, in accordance with the *Regulamento de avaliação da atividade dos investigadores contratados do Iscte - Instituto Universitário de Lisboa* (2020). The high success rate of validated and deposited records, compared to rejections, demonstrates the effectiveness of the current model and positions it as a transferable good practice for institutions seeking to articulate CRIS platforms with IR.

In a research landscape where CRIS data increasingly supports national and international evaluation and funding systems, data quality assumes critical importance. Inaccurate or inconsistent metadata can negatively affect researcher careers, institutional assessments, and access to funding. The findings presented reinforce the need for robust quality control mechanisms, transparency and accountability in research information management, as essential conditions for fairly and contextualized evaluation aligned with the principles of Open Science.

CRIS systems are now essential tools for managing scientific outputs. They help integrate data, improve quality, and share knowledge. *Ciência_Iscte* is a key part of Iscte's research policy. It helps the university evaluate research fairly and connect with international research networks. High-quality data is vital for these systems to work reliably. Libraries are the most specialized organizations for this task within universities. Using international metadata standards and reliable sources is important to avoid errors and ensure information is accurate. Connecting the IR and the CRIS system makes data quality even more important. If the metadata is poor, it can negatively affect researchers, institutions, and projects. In performance reviews, the expression “garbage in, garbage out” is a serious ethical issue. Wrong or incomplete metadata can unfairly impact career progress and funding. This makes high standards for checking and validating information essential.

Experience shows that the library's specialized curation is necessary for *Ciência_Iscte* to work effectively and to share the university's research. This model may be used by other institutions wishing to integrate their CRIS systems and IR in order to improve research management.

References

- Amante, M. J., Lopes, S., Marçal, B., & Segurado, T. (2015). Partilhar ciência: O repositório do ISCTE-IUL e a interoperabilidade com um sistema CRIS. In Associação Portuguesa de Bibliotecários, Arquivistas e Documentalistas (Ed.), *12º Congresso da BAD: Ligar. Transformar. Criar Valor.* Évora, Portugal.
<https://publicacoes.bad.pt/revistas/index.php/congressosbad/article/view/1269>
- Azeroual, O., & Lewoniewski, W. (2020). How to inspect and measure data quality about scientific publications: Use case of Wikipedia and CRIS databases. *Algorithms*, 13(5), 107.
<https://doi.org/10.3390/a13050107>
- Azeroual, O., Schöpfel, J., Pölönen, J., & Nikiforova, A. (2022). FAIRification of CRIS: A review. In F. Coenen, A. Fred, D. Aveiro, J. Dietz, J. Bernardino, E. Masciari, & J. Filipe (Eds.),

- Knowledge discovery, knowledge engineering and knowledge management. 14th International Joint Conference, IC3K 2022, Valletta, Malta* (Vol. 3, pp. 63–71). Springer. <https://doi.org/10.5220/0011548700003335>
- Barcelona Declaration on Open Research Information. (2024). Zenodo. <https://doi.org/https://doi.org/10.5281/zenodo.10958522>
- Biesenbender, S., Petersohn, S., & Thiedig, C. (2019). Using Current Research Information Systems (CRIS) to showcase national and institutional research (potential): Research information systems in the context of Open Science. *Procedia Computer Science*, 146, 142–155. <https://doi.org/10.1016/j.procs.2019.01.089>
- Bryant, R., Clements, A., Groenewegen, D., Huggard, S., Mercer, H., Missingham, R., Oxnam, M., Rauh, A., & Wright, J. (2017). *Research Information Management: Defining RIM and the library's role*. OCLC Research. <https://doi.org/10.25333/C3NK88>
- Clements, A., & Proven, J. (2015). The emerging role of institutional CRIS in facilitating open scholarship. In Paul Ayris (Chair), *44th LIBER Annual Conference 2015: Towards open science*. euroCRIS. <http://hdl.handle.net/11366/393>
- cOAlition S. (n.d.). Plan S. Retrieved March 26, 2026, from <https://www.coalition-s.org/about/>
- cOAlition S. (2023, January 26). cOAlition S confirms the end of its financial support for Open Access publishing under transformative arrangements after 2024. <https://www.coalition-s.org/coalition-s-confirms-the-end-of-its-financial-support-for-open-access-publishing-under-transformative-arrangements-after-2024/>
- de Castro, P., & Puuska, H.-M. (2023, August 9). Research information management systems: Covering the whole research lifecycle. In Jean-François Desnos & Martín López Nores (Eds.), *Proceedings of European University Information Systems Congress 2023, Vigo, Spain. EPiC Series in Computing* (Vol. 95, pp. 257–265). easychair. <https://doi.org/10.29007/z7vk>
- de Castro, P., Shearer, K., & Summann, F. (2014). The gradual merging of repository and CRIS solutions to meet institutional research information management requirements. *Procedia Computer Science*, 33, 39–46. <https://doi.org/10.1016/j.procs.2014.06.007>
- European Commission. (n.d.). *The EU's open science policy*. Open Science. Retrieved March 26, 2026, from https://research-and-innovation.ec.europa.eu/strategy/strategy-research-and-innovation/our-digital-future/open-science_en
- Fundação para a Ciência e a Tecnologia. (2025, February 11). *Nova Política sobre Acesso Aberto a Publicações Científicas da FCT promove a visibilidade da produção científica*. <https://www.fct.pt/nova-politica-sobre-acesso-aberto-a-publicacoes-cientificas-da-fct-promove-a-visibilidade-da-producao-cientifica/>
- Inácio, A., Lopes, A., Cunha, C., Boavida, C. P., & Segurado, T. (2023, September). Fluxo do depósito de dados de investigação no Iscte: Estratégia, infraestrutura e formação [Poster]. *14ª Conferência Lusófona de Ciência Aberta (ConfOA)*. Natal, Brasil. Zenodo. <https://doi.org/10.5281/zenodo.8402150>
- Iscte – Instituto Universitário de Lisboa. (2023). *Política de gestão e partilha de dados de investigação do Iscte*. Zenodo. <https://doi.org/10.5281/zenodo.11202971>
- Johnston, L. R., Carlson, J., Hudson-Vitale, C., Imker, H., Kozłowski, W., Olendorf, R., & Stewart, C. (2018). How important are data curation activities to researchers? Gaps and opportunities for academic libraries. *Journal of Librarianship and Scholarly Communication*, 6(1), eP2198. <https://doi.org/10.7710/2162-3309.2198>
- Joint, N. (2008). Current Research Information Systems, open access repositories and libraries. *Library Review*, 57(8), 570–575. <https://doi.org/10.1108/00242530810899559>
- Lopes, A. L. (2019). Integrating a local CRIS with the PTCRIS synchronization ecosystem. *Procedia Computer Science*, 146, 166–172. <https://doi.org/10.1016/j.procs.2019.01.091>

- Rede para a Gestão de Dados de Investigação. (2025). *Re.Data: Conectar dados, potenciar ciência*. <https://redata.pt/>
- Regulamento nº 369/2020: Regulamento de avaliação da atividade dos investigadores contratados do Iscte - Instituto Universitário de Lisboa. (2020). *Diário da República, 2.ª Série, Parte E, nº 72 (2020-04-13)*, 156-171. <https://diariodarepublica.pt/dr/detalhe/regulamento/369-2020-131348780>
- Schöpfel, J., & Azeroual, O. (2021). Current Research Information Systems and institutional repositories: From data ingestion to convergence and merger. In D. Baker & L. Ellis (Eds.), *Future directions in digital information* (pp. 19–37). Chandos. <https://doi.org/10.1016/B978-0-12-822144-0.00002-1>
- Schöpfel, J., & Azeroual, O. (2023). Ethical dimensions of research information management: A new challenge for information professionals. In M. C. Fombad (Ed.), *Information services for a sustainable society* (pp. 150–165). De Gruyter. <https://doi.org/10.1515/9783110772753-012>
- Schöpfel, J., & Azeroual, O. (2024). Responsible research assessment and research information management systems. *Encyclopedia*, 4(2), 915–922. <https://doi.org/10.3390/encyclopedia4020059>
- Schöpfel, J., Azeroual, O., Chaudiron, S., Jacquemin, B., Kergosien, E., Prost, H., & Thiault, F. (2024). From open repositories to CRIS: A case study. *Procedia Computer Science*, 249(C), 173–178. <https://doi.org/10.1016/j.procs.2024.11.061>
- Sivertsen, G. (2019). Developing Current Research Information Systems (CRIS) as data sources for studies of research. In W. Glänzel, H. F. Moed, U. Schmoch, & M. Thelwall (Eds.), *Springer handbook of science and technology indicators* (pp. 667–683). Springer. https://doi.org/10.1007/978-3-030-02511-3_25
- Tammaro, A. M., Matusiak, K. K., Sposito, F. A., & Casarosa, V. (2019). Data curator's roles and responsibilities: An international perspective. *Libri*, 69(2), 89–104. <https://doi.org/10.1515/libri-2018-0090>
- Zareef, M., & Jabeen, M. (in press). Systematic literature review of digital curation services in academic libraries (2001–2023): A global perspective. *Journal of Information Science*. <https://doi.org/10.1177/01655515241305348>