

Bridging CRIS and National CV through Researcher-Controlled Synchronization

Tiago Figueiredo Santos
University Institute of Lisbon (ISCTE-
IUL)
Lisbon, Portugal
tiago.figueiredo.santos@iscte-iul.pt

João Carlos Ferreira
University Institute of Lisbon (ISCTE-
IUL), ISTAR
Lisbon, Portugal
joao.carlos.ferreira@iscte-iul.pt

António Luís Lopes
University Institute of Lisbon (ISCTE-
IUL), ISTAR
Lisbon, Portugal
antonio.luis@iscte-iul.pt

Abstract— Maintaining researcher profiles across both institutional CRIS systems and national CV platforms creates redundant data entry and metadata inconsistencies. This paper presents a researcher-centric, consent-based integration between Iscte – University Institute of Lisbon's institutional CRIS, Ciência-Iscte, and Portugal's national CV platform, CIÊNCIAVITAE. Unlike automated synchronization approaches, the mechanism places the researcher in full control: synchronization occurs only through explicit researcher action, following an identifier-based matching and validation process. Building on prior work that relied on ORCID as an intermediary data hub, the current approach establishes a direct connection between the two systems, thereby extending the scope of synchronization from articles to a significantly broader set of entity types, including proceedings, books and chapters, datasets, software, preprints, supervision, teaching, peer review, outreach, affiliations, projects, and funding. The paper describes the architecture, including the onboarding, reconciliation, and validation processes, and presents a cost-benefit analysis based on the institution's 203,756 records across 4,625 researchers. The simulation estimates a time saving of approximately two hours per researcher per synchronization cycle and a benefit-cost ratio of 27:1 relative to the development investment.

Keywords— CRIS, national CV, interoperability, persistent identifiers

I. INTRODUCTION

Numerous studies have demonstrated that the visibility of research outcomes is essential for maximizing their impact and fostering subsequent advancements [1][2]. Nevertheless, the existence of multiple information systems dedicated to the management of scientific outputs creates significant challenges for the efficient dissemination and sharing of the knowledge and innovations generated through research activities [3]. Universities and research institutions must aggregate researcher data from many platforms for assessment, funding, and reporting purposes. Maintaining CRIS systems for the institution and the national CV portal simultaneously creates a situation of duplication of efforts in filling in data and the occurrence of conflicting metadata. For example, in Portugal, if an institution has a local CRIS, there is a need to maintain the same metadata regarding individuals, affiliations, outputs, projects, and activities in both that local CRIS and also in the national CV portal (CIÊNCIAVITAE). To avoid having researchers fill in the information twice, CIÊNCIAVITAE offers an API that can be used to retrieve and send data to the platform, thus allowing every institution to adhere to the principle of “entering once, reusing multiple times”.

In this work, we describe the integration between the institutional CRIS of Iscte – University Institute of Lisbon and the national CV platform (CIÊNCIAVITAE) as a researcher-driven process and intentionally carried out by the researcher with no automated synchronization being done. This is done by connecting the researcher account to CIÊNCIA-ID – the national researcher identifier service, having a publicly accessible curriculum vitae on a national platform, and authorizing the institution to write on the national CV platform on behalf of the researcher. Once these steps are done, a reconciliation dashboard displays records that are paired or existing only on one system, and the researcher can select a direction of update (CRIS → CV, CV → CRIS) or create a corresponding counterpart. This study contributes to broadening the scope from a output-centric focus to a significantly broader set of entity types.

Our three contributions are: (1) a concise architecture for consent-driven, identifier-centric interoperability; (2) an expanded scope including outputs, activities, affiliations, projects, and funding; and (3) a cost-benefit analysis that supports such an approach.

II. BACKGROUND

Iscte has an institutional CRIS (Ciência-Iscte) for the purpose of profiling, reporting, and evaluating researchers. In Portugal, CIÊNCIAVITAE is the national CV system with an individual-centered perspective, and there is an overlap of information between the two systems concerning individuals, productions, activities, and projects.

Previously [4], the integration process at Iscte was based on a researcher-centric consent-based integration process, but one where the researcher was required to exchange the information through the ORCID platform first to eventually reach the national CV platform, and vice-versa. In other words, the ORCID platform acted as a data exchange hub between the two platforms. However, the coverage of metadata in ORCID limited the level of information exchange between the CRIS and CIÊNCIAVITAE. This meant that the synchronization was not comprehensive, mainly covering articles, which resulted in the exclusion or incomplete integration of several entities, including proceedings, books/chapters, datasets/software, preprints, supervision, teaching, peer review, outreach, and more complex affiliations, projects, and funding.

The synchronization mechanism was reimplemented, this time with a direct connection between the CRIS and CIÊNCIAVITAE, thus allowing it to be extended to cover the entirety of common structures between both systems, while keeping the same safety approach (researcher action only, consent, identifier first, matching, validation, and provenance). The dashboard provides the same three options: update data in CRIS → CV direction, update data in CV → CRIS direction, or simply creating the corresponding counterpart on the side that has the data missing.

Past CRIS/CV interoperability experiences also highlight the importance of persistent identifiers (ORCID, DOI/ISBN/ISSN), standards (e.g., CERIF), and harvesting/API exchange, often including automated synchronization [5][6]. The ISCTE model's approach is characterized by its strict researcher control (no unattended write allowed) and the broad range of entities synchronized, thereby illustrating that scope can be maximized without impacting consent, auditability, and data integrity.

III. SYSTEM OVERVIEW

The artifact consists of a Researcher Dashboard inside the CRIS, an Integration Service that handles API requests to the national platform, and the CIÊNCIA-ID/CIÊNCIAVITAE external services. The researcher dashboard has three requirements: association between accounts, publication of the researcher's CV, and write-authorization on behalf of the researcher. Finally, it shows a side-by-side reconciliation screen for each domain: identity, qualifications, outputs, activities, distinctions, affiliations, projects/funding. Every row shows delineation between institutional/national versions, a status badge, and three available options: adopt institutional → update CV, adopt CV → update institutional, generate missing counterpart.

All write operations are executed with appropriate OAuth scopes, are REST/JSON-based, and go through a validation pipeline that includes schema, business rule, identifier normalization, deduplication, authority matching, and provenance stamping. A local mapping database stores cross-system relationships between local ID and national ID. Conflicts are handled in the absence of background sync by exclusively relying on researcher input to prevent overwrites.

We have extended the scope of synchronization beyond articles to include proceedings, books/chapters, datasets, software, and preprints; supervision, teaching, peer review, and outreach; links to historical contexts; and projects/funding with funder/grant identifiers. The crosswalk also includes identifier-first match and vocabularies, such as article → article-journal, ISBN for books/chapters, and repository URLs/DOI for datasets/software.

To illustrate how the reconciliation process works in practice, we describe four representative scenarios that cover the main synchronization operations across different entity types:

1. **Adopt CRIS → update CV:** A researcher has a journal article in the CRIS with a DOI and a corrected title. The national CV contains an older version of the same record, matched by DOI. The researcher selects the CRIS version as authoritative; the system performs an idempotent write to the national CV and records provenance metadata linking both versions.
2. **Adopt CV → update CRIS:** A conference paper on the national CV includes proceedings volume and page numbers that are missing from the CRIS record. The records are matched by DOI, and the researcher chooses to adopt the CV version. The additional metadata is normalized into the CRIS schema and stored locally.
3. **Create counterpart:** A dataset with a DOI exists only on the national CV, with no corresponding entry in the CRIS. The researcher triggers the creation of a local counterpart; the system generates the CRIS record, maps it to the national identifier, and stores the cross-reference for future reconciliation.
4. **Activities and affiliations:** A supervision record shows a date range discrepancy between the two systems, and an affiliation entry has a missing end date on the CV side. Identifier fields are treated as immutable during reconciliation, while descriptive fields such as dates and role descriptions are updated only after explicit researcher confirmation.

IV. COST-BENEFIT ANALYSIS

To assess the practical value of the integration, we conducted a cost-benefit simulation based on the institution's current CRIS data. The analysis considers two dimensions: the one-time cost of developing the integration and the recurring time savings it provides to the research community.

The development of the integration between the institutional CRIS and CIÊNCIAVITAE reportedly required 513 hours of work, encompassing requirements analysis, API documentation review, software development, testing, and deployment. At the institution's internal development rate, this represents a cost of €8,421.

On the benefit side, we estimated the time savings by comparing the effort required for manual data entry with the time required using the automated synchronization mechanism. The institutional CRIS currently holds 203,756 records belonging to 4,625 researchers across five categories: outputs, activities, identification, qualifications, and distinctions. Based on observations of typical researcher behaviour, we estimate that manually entering a record into the national CV platform takes approximately 3 minutes on average, while the automated synchronization of that same record (including reconciliation, validation, and transfer) takes approximately 7 seconds. We used an average researcher hourly cost of €23.14 (estimated based on data from [7] and [7]) to convert time savings into monetary values.

Table I presents a breakdown of the time savings per researcher (for a total of 4625 researchers) by category. Activities and outputs represent the largest share of records and, consequently, the greatest time savings. On average, each researcher saves

approximately 127 minutes (over two hours) per synchronization cycle, compared to manually entering the same information on the national CV platform.

Table I — Time Savings per Researcher by Category

Category	Total Records	Records/Researcher	Manual Time (min.)	Sync Time (min.)	Time Saved (min.)
Outputs	77,385	16.73	50.20	1.95	48.24
Activities	105,849	22.89	68.66	2.67	65.99
Identification	13,423	2.90	8.71	0.34	8.37
Qualifications	6,886	1.49	4.47	0.17	4.29
Distinctions	213	0.05	0.14	0.01	0.13
Total	203,756	44.06	132.17	5.14	127.03

Table II summarizes the overall cost-benefit analysis at the institutional level.

Table II — Cost-Benefit Summary

Metric	Value
Total Researchers	4,625
Total Records in CRIS	203,756
Time saved per researcher	127.03 min. (~2.1 hours)
Monetary savings per researcher	€48.99
Total institutional time savings	9,792 hours
Total institutional monetary savings	€226,578
Development cost (500 hours)	€8,421
Benefit-cost ratio	27:1

The simulation shows that the development investment is recovered many times over when measured against the aggregate time savings across the institution's researcher base. The benefit-cost ratio of approximately 27:1 demonstrates a compelling return, even under conservative assumptions. It should be noted that these figures represent a single synchronization cycle (with the current existing data in the CRIS); the actual benefit grows with each subsequent use of the system, as researchers continue to produce new outputs and engage in new activities. Furthermore, the analysis does not account for indirect benefits such as improved data quality, reduced metadata inconsistencies between the institutional and national systems, and researcher satisfaction from reduced administrative burden.

V. CONCLUSIONS

This paper presented a consent-based, researcher-centric integration between an institutional CRIS and the Portuguese national CV platform CIÊNCIAVITAE, where synchronization occurs exclusively through explicit researcher action. By establishing a direct connection between the two systems (replacing the previous ORCID-mediated approach) the scope of synchronization was extended to cover outputs, activities, affiliations, qualifications, distinctions, and projects/funding, giving researchers the ability to consolidate their full academic profile from a single entry point.

The cost-benefit simulation, based on 203,756 records across 4,625 researchers, estimates a time saving of approximately two hours per researcher per synchronization cycle and a benefit-cost ratio of 27:1 relative to the development investment, demonstrating the practical viability of the approach even under conservative assumptions.

The researcher-controlled design ensures that no unattended writes occur, minimizing silent drift and preserving data integrity across systems. At the same time, the reliance on persistent identifiers (ORCID, DOI, ISSN/ISBN) and a structured validation pipeline reduces the risk of duplicates and metadata conflicts. The main limitations are legacy records without identifiers and incomplete ORCID adoption, both of which can be mitigated through conservative heuristics and curation queues. The approach is generalizable to other institutions with CRIS/CV interoperability needs and similar identifier infrastructures.

REFERENCES

- [1] Metag, J. (2021). Tension between visibility and invisibility: Science communication in new information environments. *Studies in Communication Sciences*, 21(1), 129-144.
- [2] Liu, L., & Liu, W. (2023). The engagement of academic libraries in open science: A systematic review. *The Journal of Academic Librarianship*, 49(3), 102711.
- [3] Bagstad, K. J., Balbi, S., Adamo, G., Athanasiadis, I. N., Affinito, F., Willcock, S., ... & Villa, F. (2025). Interoperability for ecosystem service assessments: Why, how, who, and for whom?. *Ecosystem Services*, 72, 101705.
- [4] Lopes, A. L. (2019). Integrating a local CRIS with the PTCRIS synchronization ecosystem. *Procedia computer science*, 146, 166-172.
- [5] Aghassibake, N., Castello, O. G., Gujilde, P., & Rabun, S. (2023). Visualizing institutional activity using persistent identifier metadata. *Information Services and Use*, 43(3-4), 335-342.
- [6] Schnieders, K., Mierz, S., Bocalini, S., Meyer zu Westerhausen, W., Hauschke, C., Hagemann-Wilholt, S., & Schulze, S. (2022). ORCID coverage in research institutions—Readiness for partially automated research reporting. *Frontiers in Research Metrics and Analytics*, 7, 1010504.
- [7] Teaching/Researching staff in Portugal in 2023/2024, https://estatisticas-educacao.dgeec.medu.pt/ef/2024/ensino_superior/docentes.asp, consulted 2026-02
- [8] Income for Teaching/Researching staff in Portugal. (2025). https://www.fenprof.pt/media/download/81EA277828A753E7FF6D3396B6861539/escaloes_e_indices_2025.pdf, consulted 2026-02