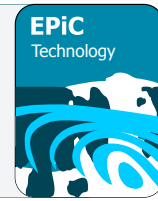




EPiC Series in Technology

Volume 5, 2026, Pages 280–295

Proceedings of 17th International Conference on Current Research Information Systems



Building FAIR Metadata Mapping Model for Scientific Instruments in the Ukrainian Research Information System (URIS): Pathways to Recovery of Research Infrastructures

Sabina Auhunas

TIB – Leibniz Information Centre for Science and Technology, Hannover, Germany
sabina.auhunas@tib.eu

Abstract

Scientific instruments remain insufficiently represented as interoperable metadata entities within current CRIS implementations. This gap is particularly critical in conflict-affected research environments, where structured metadata is required not only for infrastructure visibility but also for damage assessment, recovery planning, and preservation of scientific capacity. Since the full-scale Russian invasion of February 2022, approximately 35% of Ukrainian research infrastructure has been damaged or destroyed, including 251 units of scientific equipment. This paper presents results from an ongoing project investigating metadata requirements for scientific instruments within the Ukrainian research information ecosystem. Drawing on policy analysis, stakeholder requirements, and comparative assessment of CERIF, MERIL, PIDINST, OpenAIRE, and B2INST, the study identifies two systematic gaps: the absence of recovery-oriented metadata support across all examined systems, and insufficient instrument-level CRIS integration within URIS. In response, the project outlines a FAIR-aligned, CERIF-compatible conceptual framework representing scientific instruments as first-class CRIS entities and proposes a recovery-oriented metadata layer comprising six structured dimensions, operational status, damage classification, relocation history, software dependencies, recovery priority, and knowledge preservation. The framework will be empirically validated through a structured pilot implementation involving 60–100 instrument records from Ukrainian research institutions, planned for 2027.

Keywords: CRIS, scientific instruments, digital research infrastructure, FAIR, metadata mapping, URIS, CERIF, MERIL, PIDs, post-war recovery, conflict-affected environments, Ukraine

1 Introduction

Research infrastructures (RIs) constitute strategic assets underpinning scientific excellence, technological competitiveness, and knowledge production across all disciplines [1, 2]. Beyond physical facilities and laboratories, their effective governance increasingly depends on structured, interoperable, and machine-readable metadata describing the resources that enable research and innovation [3, 4]. Current Research Information Systems (CRIS) have evolved into central platforms for managing and integrating such information across institutional, national, and international environments, supporting interoperability between projects, publications, organizations, datasets, and researchers [5]. Within this context, scientific instruments (SI) and RI components are increasingly recognized as integral elements of national CRIS ecosystems requiring standardized, interoperable metadata representation [6].

Despite this recognition, SIs remain insufficiently represented as interoperable metadata entities within many CRIS implementations [7]. Instruments constitute the operational core of research infrastructures, enabling experimentation, measurement, validation, and scientific data production across virtually all domains [8]. Yet metadata describing them frequently remains fragmented across institutional inventories, spreadsheets, procurement records, laboratory documentation, and disconnected local databases lacking persistent identifiers, controlled vocabularies, and interoperability mechanisms linking instruments to projects, publications, datasets, and organizations. Initiatives such as CERIF-based infrastructure models, MERIL, PIDINST and national CRIS implementations, including Finland's Research.fi, demonstrate growing interest in machine-readable instrument metadata and infrastructure discoverability [9, 10]. Nevertheless, the integration of SIs as first-class CRIS entities linked within the broader research graph remains underdeveloped. The requirements of conflict-affected research ecosystems, where metadata may additionally support damage assessment, recovery planning, and preservation of scientific capacity, remain largely unexplored in the CRIS literature.

Ukraine represents a particularly critical case. Since the full-scale Russian invasion of 24 February 2022, Ukrainian research institutions and higher education institutions (HEIs) have experienced extensive destruction and disruption of RIs. Approximately 35% of Ukrainian research infrastructure has been damaged or destroyed [11]. At the instrument level, the scale of documented losses is considerable: a UNESCO assessment reports that 29.4% of all research institutions have been affected, with 1,443 buildings and 188 engineering structures damaged. Among movable assets, 677 units sustained damage, including 251 units of scientific equipment, of which 423 were destroyed and 77 damaged in practice, rendering all affected equipment unfit for further use [12]. figures represent only documented cases; the actual scope of instrument-level losses is likely considerably larger, given the absence of systematic, machine-readable documentation of scientific equipment at national level before and during the conflict. Beyond physical destruction, the war has exposed substantial governance and metadata-related weaknesses: damage documentation has been conducted unsystematically by different organizations using heterogeneous methodologies, and existing infrastructure records remain distributed across spreadsheets, PDF reports, local databases, and paper documentation [13]. Current national mechanisms do not address SIs as interoperable metadata entities that require standardized, machine-readable representations.

Ukraine has nonetheless been developing national policies aimed at integrating its research infrastructure ecosystem into the European Research Area. The Ukrainian Research Information System (URIS), established in 2020 and governed by regulations approved in 2022, aims to support aggregation, interoperability, and analytical processing of research information, integrating data on researchers, organizations, projects, publications, and elements of research infrastructure [14, 15]. However, FAIR-oriented metadata representation of SIs within URIS remains limited: current infrastructure modules primarily focus on institutional visibility and registry functions rather than machine-readable instrument-level metadata, operational condition, persistent identifiers, and linkages

to projects, datasets, and publications. The development of FAIR-aligned metadata frameworks for SIs is therefore important not only for interoperability and Open Science implementation but also for post-war recovery governance and integration into the European Research Area. Without structured and interoperable instrument metadata, it is difficult to assess infrastructure losses, coordinate shared access, support evidence-based recovery planning, or integrate Ukrainian research infrastructures into European ecosystems.

This paper presents results from the project "Identifying challenges of metadata mapping for scientific instruments in the Ukrainian Research Information System: Pathways to recovery of research infrastructures in conflict-affected societies", conducted at the Open Science Lab at TIB – Leibniz Information Centre for Science and Technology and University Library and funded through the Philipp Schwartz Initiative of the Alexander von Humboldt Foundation (<https://vivo.tib.eu/fis/display/gr59000044>).

The objective of this study is to investigate the metadata requirements for SIs within the Ukrainian research infrastructure ecosystem and to develop a FAIR-aligned metadata framework that supports their representation as interoperable CRIS entities within URIS. The study further examines how recovery-oriented metadata dimensions can be incorporated into CRIS environments to support infrastructure reconstruction and long-term sustainability.

Building upon policy analysis, stakeholder requirements, and comparative examination of CERIF, MERIL, PIDINST, OpenAIRE [16], B2INST [17] and URIS infrastructure representations, this paper makes four contributions: (1) identifying metadata requirements and outlining a conceptual direction for a FAIR-aligned metadata framework for SIs as interoperable CRIS entities; (2) integrating recovery-oriented metadata dimensions into CRIS interoperability discussions; (3) extending infrastructure-aware CRIS approaches toward instrument-level metadata governance; and (4) supporting future integration of Ukrainian RI into broader European Open Science ecosystems.

2 Ukrainian Research Infrastructure Landscape

2.1 Research infrastructure landscape in Ukraine

The Ukrainian research and development ecosystem comprises higher education institutions, research organizations, national academies, and specialized research infrastructures. As of 2025, Ukraine counts approximately 625 higher education institutions and 160 National Academy of Sciences institutes [18]. Before the full-scale invasion, approximately 80,000 research and teaching staff were engaged in scientific activities [19]; however, subsequent displacement, emigration, and wartime disruption have significantly reduced active research capacity, with estimates suggesting a loss of approximately 18% of scientists since February 2022 [12, 20].

Research infrastructures are legally defined as sets of tools, scientific equipment, resources, and related services used by the scientific community to conduct research at the highest level, encompassing centralized, distributed, virtual, and network-based environments [21]. Public funding supports research institutions and scientific programmes, while the National Research Foundation of Ukraine (NRFU) provides competitive support for equipment acquisition, laboratory modernization, and infrastructure development. Despite these mechanisms, Ukrainian research infrastructure has historically experienced uneven modernization, institutional fragmentation, and chronic underfunding conditions that became critically visible under wartime disruption [22].

In 2021, the Cabinet of Ministers of Ukraine approved the Concept of the State Target Programme for the Development of RIs of Ukraine until 2026, aimed at modernization of the national research infrastructure ecosystem and its alignment with the European Research Area [23]. The Horizon Europe project RIFF (Research Infrastructures for the Future of Ukraine) additionally supports

infrastructure road mapping, policy coordination, and integration of Ukrainian RIs into European governance processes, including alignment with ESFRI roadmaps and participation in European Research Infrastructure Consortia (ERIC) [24].

The national research infrastructure ecosystem comprises five principal categories, each characterized by distinct governance mechanisms and metadata challenges summarized in Table 1.

RI type	Main function	Governance	Metadata challenges
CSUSE Centres for Shared Use of Scientific Equipment	Shared access to scientific instruments and laboratory facilities; interdisciplinary resource sharing	MESU · NASU universities 100+ centres	Heterogeneous descriptions; no PIDs; no standardised schema; limited output linkage
National Scientific Centres	Strategic national research coordination and priority domain development	State-level governance Cabinet of Ministers of Ukraine	Fragmented interoperability; siloed systems; limited cross-system linkage
State Key Laboratories	Advanced research in priority scientific domains; specialised instrumentation	MESU	Limited machine-readable metadata; no interoperability APIs; sparse PID coverage
Scientific Facilities — National Heritage	Preservation of unique scientific facilities of national significance	Cabinet of Ministers of Ukraine · MESU 138 registered items	Limited accessibility metadata; no lifecycle records; paper-based reporting
Specialised and Unique Structures	Domain-specific infrastructure services; unique experimental facilities and platforms	Institutional governance varies by domain	Non-standardised metadata; disconnected inventories; no shared vocabulary
■ key priority for pilot ■ national coordination ■ heritage / preservation ■ domain-specific			

Table 1: Principal categories of research infrastructures within the Ukrainian research ecosystem and associated metadata challenges [19, 25].

Among these categories, Centres for Shared Use of Scientific Equipment (CSUSE) play a particularly important role for instrument-based research. CSUSEs provide researchers with shared access to specialized SIs, analytical equipment, and laboratory facilities unavailable at the level of individual laboratories or institutions, supporting collaborative use of expensive equipment and facilitating interdisciplinary cooperation. Before the war, Ukraine maintained more than one hundred CSUSEs distributed across universities and research institutions under the Ministry of Education and Science and NASU [25]. At the same time, existing national registries, including the State Register of Scientific Objects Constituting National Heritage and the State Register of Scientific Institutions Supported by the State, primarily support administrative reporting and institutional accounting rather than FAIR-oriented interoperability or machine-readable metadata exchange [19].

The fragmentation of infrastructure metadata became especially critical following the full-scale Russian invasion. Information regarding damaged and destroyed RIs has frequently been collected using inconsistent methodologies and disconnected reporting mechanisms [13]. Metadata describing scientific equipment, operational conditions, damage status, relocation needs, and infrastructure accessibility remains dispersed across spreadsheets, local institutional databases, PDF reports, and non-standardized documentation systems. Consequently, SIs remain weakly connected with publications, projects, datasets, organizations, and researchers within the broader research graph.

These challenges directly motivated the development of an initial infrastructure mapping initiative within URIS in 2024 through the Analysis Module of Research Infrastructure (AMRI) [26]. The module focused primarily on collection and aggregation of infrastructure data from CSUSEs under the Ministry of Education and Science and NASU, including metadata describing participating organizations, available scientific equipment, infrastructure needs and operational requirements. While the initiative demonstrated the feasibility of national-level infrastructure mapping, it also revealed substantial metadata-related challenges, including inconsistent equipment descriptions, absence of standardized schemas, lack of persistent identifiers, and limited interoperability between institutional records.

2.2 Digital research infrastructures and the Ukrainian CRIS ecosystem

Ukraine has progressively developed digital RIs aimed at improving accessibility, interoperability, and visibility of scientific information. A key milestone was the establishment of the URIS in 2020, with governing regulations approved by the Cabinet of Ministers of Ukraine in 2022 [14]. URIS represents the first national CRIS-oriented infrastructure in Ukraine, designed to aggregate, verify, manage, and interconnect information related to scientific and scientific-technical activities. The system supports collection, storage, interoperability, and analytical processing of research information covering researchers, organizations, projects, publications, scientific services, and elements of research infrastructure. The development of URIS reflects FAIR-oriented approaches to research information management, emphasising interoperability, machine-readability, and reuse of research information across distributed infrastructures [27]. Before URIS, research information in Ukraine was severely fragmented across institutional repositories, local databases, spreadsheets, and disconnected administrative registries without interoperable metadata exchange mechanisms [19].

2.2.1. Current state of instrument representation in URIS

The URIS infrastructure platform currently provides searchable infrastructure categories, including products, services, scientific equipment, infrastructure areas, and scientific resources. At the time of analysis, the platform contained metadata records for more than 1,200 products, 583 services, 428 scientific equipment entities, 121 infrastructure areas, and 117 scientific resources. The platform enables filtering of scientific equipment by year of manufacture, institutional affiliation, access conditions, and infrastructure type. Figure 1 illustrates examples of scientific equipment currently represented within the URIS infrastructure platform, including mass spectrometry and chromatography equipment available through Ukrainian research institutions.

Existing metadata records include basic descriptive information such as instrument title, infrastructure type, institutional ownership, manufacturer, year of manufacture, scientific domains, and access conditions. However, as illustrated in Figure 2, the current metadata structure remains primarily descriptive and registry-oriented rather than fully interoperable and machine-readable. Existing records provide limited support for persistent identifiers, standardized vocabularies, API-based interoperability, and semantic linking with publications, datasets, projects, and researchers. Lifecycle and operational-status metadata are largely absent.

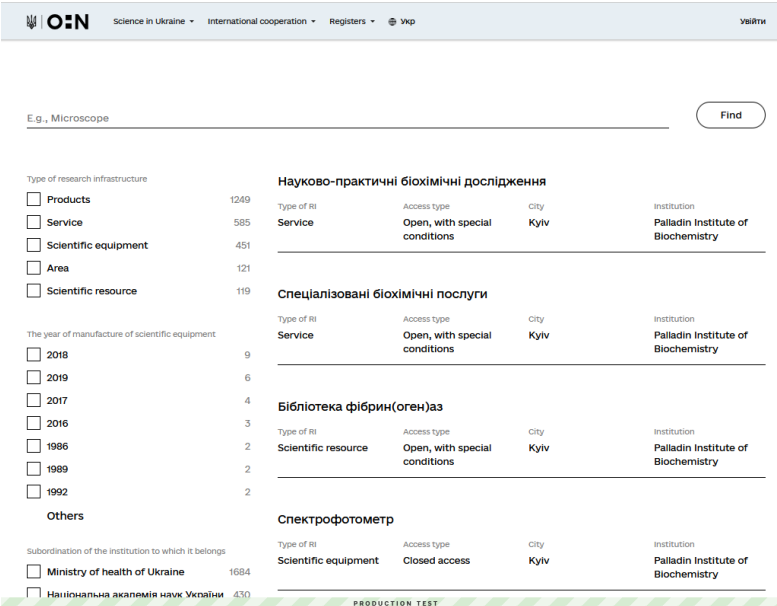


Figure 1: Examples of scientific equipment records in URIS infrastructure platform (<https://nauka.gov.ua/en/infrastructure/>). The platform currently supports basic filtering by equipment type, institutional affiliation, year of manufacture, and access conditions

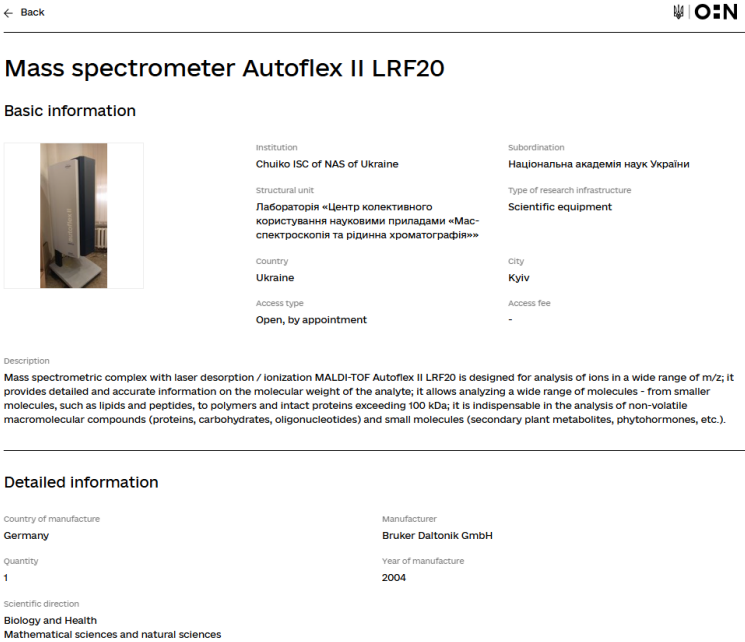


Figure 2: Current metadata representation for a scientific instrument in URIS infrastructure platform. Existing records include basic descriptive fields but lack persistent identifiers, semantic relationships, and interoperability mechanisms required for CERIF-compatible CRIS integration (<https://nauka.gov.ua/en/infrastructure/e.KaRLo49s/>)

Compared to CERIF-oriented infrastructure models and European registries such as MERIL, URIS infrastructure module provides limited interoperability support at instrument level, machine-readable semantic relationships between instruments, projects, datasets, publications, and researchers remain underdeveloped [28, 29]. Recent development in URIS additionally includes the implementation of the Analysis Module of Research Infrastructure (AMRI) and initial preparation of the Science Map of Ukraine [26]. These initiatives support collection and aggregation of infrastructure-related information, including metadata describing research organizations, scientific equipment, infrastructure needs, and operational conditions. While these developments demonstrate that URIS already provides an important institutional and technical foundation for future infrastructure interoperability, they simultaneously confirm the need for harmonized, FAIR-oriented metadata models for scientific instruments – the gap this study directly addresses.

2.3 Metadata Fragmentation and Challenges for Scientific Instruments

The representation of scientific instruments within CRIS environments is considerably more complex than representation of traditional research entities such as publications, projects, organizations, or researchers. Scientific instruments possess highly heterogeneous technical characteristics, operational conditions, ownership models, accessibility regimes, and lifecycle states, which complicate the development of standardized metadata structures across institutional and national infrastructures [7]. Within many CRIS implementations, research infrastructure information is commonly represented at the organizational or facility level rather than at the level of individual instruments, with existing registries prioritizing institutional visibility and inventory management over machine-readable, interoperable instrument-level metadata.

The Ukrainian case illustrates these challenges particularly clearly. Existing infrastructure-related information within institutional systems is distributed across spreadsheets, procurement records, laboratory inventories, websites, PDF reports, and disconnected local databases. Metadata structures differ substantially between institutions, scientific domains, and infrastructure types. Even basic instrument descriptions frequently remain inconsistent due to variations in terminology, naming conventions, technical specifications, and local documentation practices [19]. The fragmented state of instrument metadata across the Ukrainian research ecosystem is illustrated in Figure 3.

A further critical limitation concerns the absence of persistent identifier infrastructures for SIs. Existing infrastructure records rarely support globally interoperable PID systems capable of uniquely identifying instruments across institutional and national environments. Consequently, SIs remain difficult to integrate into broader research information graphs and Open Science ecosystems, and linkages between instruments, publications, projects, datasets, and researchers remain largely absent [30, 31]. Emerging platforms such as LabsArena.com, which supports search, sharing, and sale of scientific equipment across European and Ukrainian research institutions, additionally demonstrate growing demand for structured equipment databases and play a notable role in cataloguing Ukrainian research infrastructure under wartime conditions [32]. However, such marketplace-oriented platforms operate independently from national CRIS environments, providing no support for CERIF-compatible semantic relationships, persistent identifiers, or recovery-oriented metadata.

Scientific instruments additionally represent dynamic research entities that continuously participate in scientific workflows. Instruments generate experimental data, support scientific services, contribute to publications and projects, and may simultaneously belong to distributed infrastructures shared between institutions and research communities. This dynamic character requires metadata relationships considerably more complex than those typically supported by current CRIS implementations, including usage history, access modalities, calibration records, and service linkages [7].

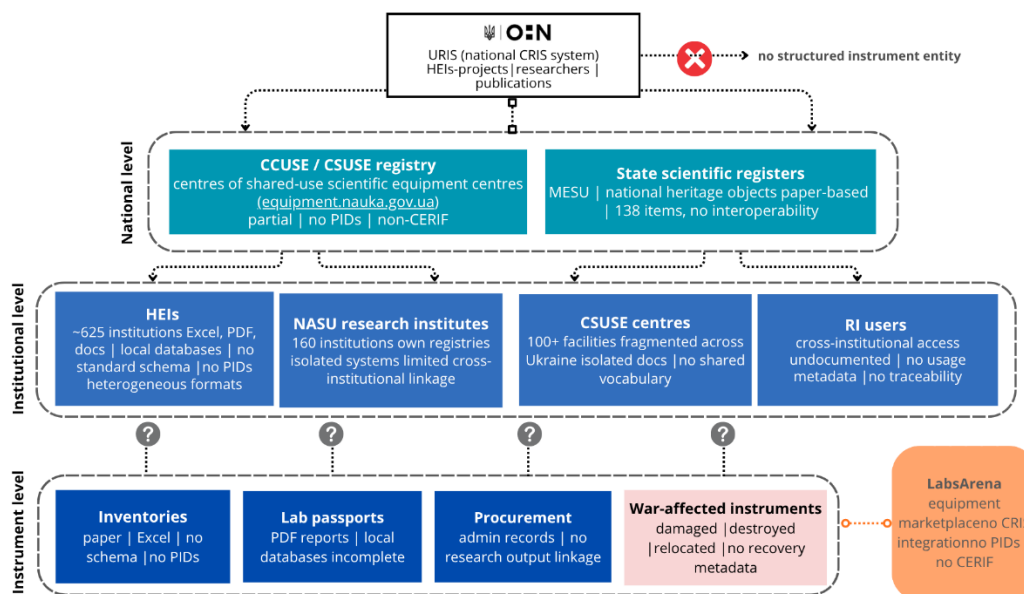


Figure 3: Fragmentation of scientific instrument metadata across institutional layers of the Ukrainian research infrastructure ecosystem. Dashed arrows indicate weak or informal linkages; structured CERIF-compatible integration with URIS is currently absent at instrument level

The wartime context has further exposed the limitations of existing infrastructure metadata practices. Under conditions of military aggression, SIs may become damaged, relocated, inaccessible, partially operational, or destroyed. As documented by UNESCO, 251 units of scientific equipment have been damaged or destroyed since February 2022 [12], with the actual scope likely considerably larger given the absence of systematic, machine-readable documentation at national level. Existing infrastructure records rarely include metadata capable of describing infrastructure condition, damage classification, relocation history, operational risks, preservation status, or recovery priorities. This significantly complicates evidence-based recovery planning and coordination of scientific infrastructure resources – a gap that existing CRIS schemas, including CERIF, MERIL, and PIDINST, do not currently address [7].

These challenges collectively demonstrate the need for a FAIR-aligned metadata framework capable of representing SIs as interoperable entities within the broader CRIS and research infrastructure ecosystem, supporting not only inventory description and discoverability but also lifecycle management, semantic linkage with research outputs, and recovery-oriented documentation in conflict-affected environments. The comparative analysis and proposed metadata architecture addressing these requirements are presented in the following section.

3 Implementation phases of project

The study applies a five-phase mixed-methods research design aimed at developing a FAIR-aligned, CERIF-compatible metadata framework for SIs and supporting their integration into URIS as interoperable CRIS entities. The design combines policy analysis, stakeholder survey, expert interviews, comparative metadata analysis, and pilot implementation. Each phase produces evidence that feeds directly into the framework development.

Phase 1: Policy and document analysis (completed). The objective of Phase 1 was to identify metadata requirements, documentation gaps, and governance challenges relevant to scientific instrument representation in conflict-affected research environments. Policy documents were retrieved via the Overton platform and official national and European databases. A structured coding framework of 109 codes across 13 thematic blocks and five analytical layers was applied. Key outputs include a structured gap analysis of existing documentation practices and identification of recovery-oriented metadata requirements absent from current CRIS implementations.

Phase 2: Stakeholder survey (in progress). The objective of Phase 2 is to collect empirical evidence on current metadata practices, instrument documentation conditions, interoperability requirements, and recovery needs across Ukrainian research institutions. The survey targets 250-350 responses from Ukrainian research institutions, higher education institutions (HEIs), National Register of Centres for Shared Use of Scientific Equipment (CCUSE/CSUSE), National Academy of Ukraine and research infrastructure users, covering instrument condition, documentation completeness, digital system usage, and perspectives on a future national instrument registry. Expected outputs include a quantitative dataset of institutional metadata practices and a prioritized list of metadata requirements to be incorporated into the proposed framework.

Phase 3: Expert interviews (planned). The objective of Phase 3 is to validate survey findings and identify implementation priorities, governance barriers, and technical requirements for a national scientific instrument registry. The phase will conduct 12-20 semi-structured interviews with researchers, laboratory managers, CRIS specialists, and metadata experts until thematic saturation. Expected outputs include validated metadata requirements, identified governance gaps, and stakeholder input for the metadata model development in Phase 4.

Phase 4: Comparative metadata analysis (preliminary results presented). The objective of Phase 4 is to benchmark URIS against existing scientific instrument metadata schemas and interoperability frameworks MERIL, PIDINST, OpenAIRE and B2INST, across six criteria: metadata element coverage, FAIR alignment, PID support, governance model, access mechanisms, and CRIS interoperability. Expected outputs include a consolidated metadata element set, a metadata crosswalk across schemas, a draft JSON metadata schema, a JSON-LD representation, and a gap analysis forming the technical foundation.

Phase 5: Pilot implementation (planned, 2027). The objective of Phase 5 is to validate the proposed metadata framework using 60–100 real scientific instrument records from approximately five Ukrainian institutions. Instruments will be stratified into three metadata maturity groups: complete, partial, and inventory-only documentation to assess framework robustness across varying conditions. The four-stage validation procedure covers record collection and gap analysis, metadata mapping to CERIF and MERIL, harmonization of terminology and identifiers, and quality evaluation against completeness, FAIR alignment, and URIS compatibility criteria. Expected outputs include a validated metadata schema, a metadata crosswalk, a pilot dataset, URIS implementation recommendations, and a scientific instrument registry prototype.

The five phases collectively feed a framework synthesis integrating governance-layer findings (Phases 1-2), stakeholder evidence (Phase 3), and technical outcomes (Phases 4-5) into a unified, empirically grounded metadata architecture for SIs in URIS.

3.1 Comparative metadata analysis of scientific instrument registries

Phase 4 of the study conducts a systematic comparative analysis of existing metadata schemas, infrastructure registries, and interoperability frameworks relevant to the representation of SIs within CRIS environments. The primary objective is to identify metadata elements, governance mechanisms, interoperability gaps, and PID implementation practices relevant to integrating SIs into URIS as CERIF-compatible entities. Particular attention is given to semantic interoperability models, metadata granularity, infrastructure classification schemes, persistent identifier support, FAIR implementation

mechanisms, and interoperability relationships between instrument entities and broader research information ecosystems [28, 33].

The analysis includes the CERIF 1.6 semantic model, the MERIL infrastructure categorization framework, the PIDINST metadata schema, OpenAIRE Explore Research Resources, B2INST, URIS infrastructure module, and recent European initiatives addressing the integration of research instruments and facilities into CRIS environments. Additional attention is given to the RIFF initiative, which seeks to strengthen interoperability, infrastructure visibility, and strategic coordination of Ukrainian RIs within ESFRI- and EOSC-related environments [24]. The comparison is structured around six criteria applied consistently across all systems: metadata element coverage, FAIR alignment, PID support, governance model, access mechanisms, and CRIS interoperability. Results are summarized in Table 2.

	URIS	MERIL	PIDINST	OpenAIRE	B2INST	Proposed framework
Core entity and relationships						
Instrument entity	1	2	2	1	2	3
Instrument–project link	1	2	1	0	0	3
Instrument–publication link	0	2	2	2	2	3
Persistent identifiers (PID)	1	1	2	2	2	3
Interoperability						
Machine-readable metadata	1	2	2	2	2	3
FAIR alignment	1	2	2	2	1	3
Recovery-oriented metadata (identified gap across all existing systems)						
Operational status	0	1	0	0	0	3
Damage classification	0	0	0	0	0	3
Relocation history	0	0	0	0	0	3
Software and calibration dependencies	0	0	1	0	0	3
Recovery priority	0	0	0	0	0	3
Knowledge preservation	0	0	0	0	0	3

0 not supported 1 partial 2 supported 3 included in proposed framework

Table 2: Preliminary comparative assessment of scientific instrument metadata capabilities. Score: 0 = not supported; 1 = partially supported; 2 = supported; 3 = included in the proposed framework. Scores for PIDINST, OpenAIRE, and B2INST represent preliminary estimates pending completion of Phase 4 systematic analysis. Recovery-oriented dimensions are absent from all existing systems. Scores represent a qualitative assessment derived from documentation review and should be interpreted as indicative rather than absolute. Formal benchmarking will be completed during Phase 4 of the project.

Recent developments within the CRIS community increasingly recognize SIs and facilities as essential components of the research graph. Rather than being treated solely as inventory objects, SIs are increasingly conceptualized as research entities linked with projects, organizations, researchers, publications, datasets, and services [34].

Within the CERIF 1.6 data model, scientific equipment is representable through the *cfEquip* entity, linked to other research information entities through standardized semantic relationships. CERIF supports semantic interoperability by separating core entities from classification schemes and relationship entities, enabling flexible representation of RI within distributed information environments [28, 35]. This architecture provides a foundation for machine-actionable infrastructure

metadata and facilitates interoperability between institutional, national, and international CRIS implementations [36, 37, 38]. MERIL supports harmonized categorization, visibility, and discoverability of RIs through standardized descriptions, scientific domain classifications, and access-related metadata aligned with the European research infrastructure landscape [9, 29]. PIDINST provides a structured schema for instrument-level persistent identification, establishing minimum metadata requirements and typed relationship structures including *isComponentOf*, *isDescribedBy*, and *hasComponent*, enabling instrument traceability across research graphs [30, 31, 39]. B2INST operationalizes PIDINST workflows and demonstrates practical PID assignment at scale.

The increasing integration of RIs into Open Science environments requires alignment with FAIR principles and EOSC interoperability recommendations. Recent EOSC initiatives emphasize machine-actionable metadata and semantic interoperability across distributed European research environments [33]. The PIDINST initiative and RDA recommendations [40] additionally highlight the growing importance of persistent identifiers for SIs as a mechanism for improving traceability, attribution, and reuse within evolving research graphs [7, 30].

Despite the maturity of these frameworks in stable cataloguing contexts, the comparative analysis identifies two categories of systematic gaps.

The first and primary gap concerns recovery-oriented metadata. Existing schemas prioritize inventory visibility, institutional reporting, and access facilitation under normal operating conditions. None of the examined systems provides structured metadata support for: operational status under disruption; damage classification by nature, extent, and cause; relocation history including cross-border instrument transfers; software and calibration dependencies required for recommissioning; recovery priority assignment; or preservation of tacit operational knowledge at risk of loss through personnel displacement. These dimensions are absent not only from URIS but from MERIL, PIDINST, OpenAIRE, and B2INST, confirming that the gap is systemic rather than context-specific, consistent with recent observations on instrument-level metadata governance [7].

The second gap concerns depth of CRIS integration. The URIS infrastructure records represent SIs primarily as administrative inventory objects with limited linkage to projects, researchers, publications, or datasets. MERIL similarly operates at facility description level rather than instrument-entity level. Neither system currently implements persistent identifiers at instrument granularity nor supports machine-actionable metadata harvesting for individual instruments, capabilities explicitly required for EOSC interoperability [33].

The comparative analysis produces four technical artefacts forming the input basis for the metadata architecture: (i) a consolidated metadata element set derived from crosswalk analysis across CERIF, MERIL, PIDINST, and URIS; (ii) a metadata crosswalk table mapping equivalent and non-equivalent elements across schemas; (iii) a draft JSON metadata schema encoding core instrument metadata in a machine-readable structure; and (iv) a JSON-LD representation supporting linked data integration and EOSC interoperability. These artefacts, together with the gap analysis in Table 2, provide the empirical and technical foundation for the recovery-oriented metadata extension.

Based on the gaps identified through the comparative analysis, the project outlines several conceptual directions for the proposed metadata architecture. Scientific instruments are conceptualized as first-class CRIS entities representable through the CERIF *cfEquip* entity and linked to organizations, projects, researchers, publications, datasets, services, and RIs through typed semantic relationships. The architecture will adopt an extension rather than replacement approach, augmenting existing CERIF and MERIL models with metadata elements not currently accommodated by those frameworks. A recovery-oriented metadata layer is envisaged to address six dimensions identified through the Phase 1 policy analysis: operational status, damage classification, relocation history, software and calibration dependencies, recovery priority, and preservation of operational knowledge. The detailed metadata element definitions, relationship mappings, and schema formalization will be developed in subsequent project phases informed by stakeholder survey findings (Phase 2), expert interviews (Phase 3), and validated through the pilot implementation (Phase 5).

3.2 Pilot implementation and validation

Phase 5 of the study will validate the proposed metadata framework through a structured pilot implementation using real scientific instrument records from Ukrainian research institutions. The pilot is planned for 2027 and constitutes the final phase of the project. Its primary objective is to assess the applicability, technical robustness, and interoperability potential of the proposed CERIF-aligned architecture before its integration into URIS as a CRIS-compliant component.

The pilot dataset will comprise 60-100 scientific instrument records. This sample size provides adequate coverage of scientific domains, MERIL infrastructure categories, and metadata complexity levels, while remaining feasible for detailed mapping and iterative validation. Only instruments meeting two baseline conditions are eligible for inclusion: the scientific instrument must be deployed for research purposes, and it must be supported by at least minimal institutional documentation. Instruments for which metadata can be linked to other CERIF entities, projects, publications, datasets, researchers, services, or organizational units are prioritized, as this linkability criterion directly operationalizes the interoperability objectives of the proposed architecture.

Instruments are included if they: (i) are used for research purposes; (ii) are owned or operated by a Ukrainian research institution; (iii) are supported by at least one form of institutional documentation; and (iv) can be attributed to a specific organizational unit. Instruments are excluded if they: (i) are used exclusively for teaching or administrative purposes; (ii) have no available documentation in any form; or (iii) belong to institutions that declined to participate in the validation activities.

The pilot will involve approximately five institutions selected to ensure diversity across geographic location, scientific discipline, institutional type, level of war-related infrastructure impact, and instrument availability. Candidate groups include national research universities, technical universities, medical universities, and institutes of the National Academy of Sciences of Ukraine. Institutions that participated in the Phase 2 survey or Phase 3 expert interviews are prioritized to ensure methodological continuity. The deliberate inclusion of institutions representing heterogeneous infrastructure conditions, fully operational, partially disrupted, relocated, and war-damaged, distinguishes this pilot from analogous validation exercises conducted in stable research environments.

The pilot covers five scientific domains characterized by intensive reliance on physical equipment and laboratory infrastructure: physical sciences, engineering and technology, life sciences, medical sciences, and environmental sciences. This domain diversity constitutes a deliberate source of methodological variation, enabling assessment of the framework's generalizability across heterogeneous CRIS contexts.

To evaluate framework robustness across varying documentation conditions, instruments are stratified into three metadata maturity groups. Group 1 (complete documentation) comprises instruments supported by inventory records, laboratory passports, technical specifications, calibration histories, and documented linkages to research outputs. Group 2 (partial documentation) includes instruments for which only a subset of metadata elements is available. Group 3 (inventory-only) consists of instruments documented exclusively through administrative registry entries, representing the minimum metadata conditions commonly observed in Ukrainian institutional repositories. This stratification enables assessment of how effectively the proposed schema accommodates incomplete metadata and provides a basis for defining minimum viable metadata profiles for resource-constrained environments.

Metadata will be collected from the following institutional sources: equipment inventories, laboratory passports, procurement documentation, maintenance and calibration records, spreadsheet-based registries, PDF technical reports, local databases, and existing URIS infrastructure records. The inclusion of heterogeneous, non-standardized source formats reflects realistic documentation conditions in Ukrainian research organizations.

The validation is structured across four sequential stages, illustrated in Figure 2. Stage 1 covers record collection and gap analysis: existing institutional records are collected and analyzed to identify available metadata elements and document gaps relative to the proposed schema. Stage 2 covers metadata mapping: identified elements are mapped to the proposed architecture and aligned with *cfEquip* and related CERIF entities, MERIL classification categories, and EOSC interoperability constructs. Stage 3 covers harmonization: terminology, classification assignments, persistent identifier schemes, and organizational references are normalized. Stage 4 covers quality evaluation: transformed records are assessed against five dimensions: metadata completeness, semantic consistency, interoperability, FAIR alignment, and URIS compatibility.

The pilot examines targeted validation scenarios corresponding to core CERIF relationship types: semantic links between instrument records and associated projects, publications, datasets, researchers, organizational units, services, and research infrastructures. Particular attention is directed toward recovery-oriented metadata elements – operational status, damage classification, relocation history, software and calibration dependencies, and recovery priority and their capacity to be represented within a CERIF-compatible, machine-readable structure.

Successful validation requires the ability to: (i) transform heterogeneous institutional records into a harmonized, schema-conformant metadata structure; (ii) establish typed semantic links between instrument records and other CRIS entities in accordance with the CERIF data model; (iii) assign appropriate CERIF classification terms and MERIL category codes; and (iv) represent all recovery-oriented metadata dimensions in a machine-readable, interoperable format. The pilot outputs, validated schema, metadata crosswalk, pilot dataset, URIS implementation recommendations, and registry prototype will provide the evidentiary basis for finalizing the proposed architecture and defining a roadmap for instrument integration into URIS.

4 Conclusion

This project investigated metadata requirements for scientific instruments within the Ukrainian research infrastructure ecosystem and outlined a FAIR-aligned, CERIF-compatible framework supporting their integration into URIS as interoperable CRIS entities. The project was conducted in the context of wartime infrastructure disruption, where the absence of structured, machine-readable instrument metadata has significantly complicated infrastructure governance, damage assessment, and evidence-based recovery planning.

The policy analysis and comparative metadata assessment confirmed that existing documentation mechanisms across URIS, MERIL, PIDINST, OpenAIRE, and B2INST provide insufficient support for instrument-level metadata governance. While established frameworks offer mature support for infrastructure visibility and semantic interoperability, none provides structured metadata support for operational status under disruption, damage classification, relocation history, software and calibration dependencies, recovery priority, or preservation of tacit operational knowledge. This gap is systemic rather than context-specific and has received limited attention in the current CRIS literature.

In response, the project identified the metadata requirements and conceptual foundations for a metadata architecture in which SIs are represented as first-class CRIS entities connected to organizations, projects, researchers, publications, datasets, services, and RIs through typed CERIF-compatible semantic relationships. The architecture extends rather than replaces existing interoperability models, supports incremental adoption through tiered metadata profiles, and adheres to FAIR research information principles. A recovery-oriented metadata layer comprising six structured dimensions was introduced, designed to remain machine-readable and interoperable with the CERIF core model rather than functioning as free-text annotations.

The proposed framework is currently at the conceptual stage, and several limitations should be acknowledged. The comparative assessment of PIDINST, OpenAIRE, and B2INST presented in Table 2 represents preliminary estimates pending completion of the full Phase 4 systematic analysis. The framework has not yet been validated empirically. Validation will be conducted through a structured pilot implementation planned for 2027, involving approximately five Ukrainian institutions and 60-100 instrument records stratified across three metadata maturity groups. The pilot will produce a validated metadata schema, a metadata crosswalk and URIS implementation recommendations. Survey and interview findings from Phases 2 and 3 will be reported separately.

For URIS, this project contributes to its ambition to represent SIs not merely as inventory objects but as semantically connected, interoperable CRIS entities within the national research graph. Achieving this will strengthen infrastructure discoverability, simplify damage reporting and recovery coordination across Ukrainian research institutions, and facilitate long-term integration of Ukrainian RIs into the European Research Area and EOSC ecosystem.

Acknowledgments

The author thanks Professor Dr. Sören Auer for his academic guidance during the project. The author thanks Christian Hauschke in particular for his mentoring and support in shaping the direction of this research. The author also thanks Dr. Alla Zharinova, Director of the State Scientific and Technical Library of Ukraine, for her support and cooperation throughout the project. This project would not have been possible without the financial support of the Philipp Schwartz Initiative of the Alexander von Humboldt Foundation, funded by the Federal Foreign Office of Germany. The author thanks TIB – Leibniz Information Centre for Science and Technology and the Open Science Lab for providing the institutional environment for this project.

References

- [1] ESFRI. (2020). Supporting the Transformative Impact of Research Infrastructures on European Research. European Commission, Brussels. Retrieved from <https://op.europa.eu/en/publication-detail/-/publication/310d694d-3f53-11eb-b27b-01aa75ed71a1>
- [2] Anderson, S. (2013). What are Research Infrastructures? *International Journal of Humanities and Arts Computing*, 7(1-2), 4–23. <https://doi.org/10.3366/ijhac.2013.0078>
- [3] Hallonsten, O. (2020). Research Infrastructures in Europe: The Hype and the Field. *European Review*, 28(4), 617–635. <https://doi.org/10.1017/S1062798720000095>
- [4] OECD. (2014). International Distributed Research Infrastructures. Retrieved from <https://www.oecd.org/sti/inno/international-distributed-research-infrastructures.pdf>
- [5] Jeffery, K.G., & Asserson, A. (2010). Institutional Repositories and Current Research Information Systems. *New Review of Information Networking*, 15(1), 71–83. <https://doi.org/10.1080/13614576.2010.518851>
- [6] Zendulková, D., Gavurníková, G., & Turňa, J. (2022). Research Infrastructure Map as a part of Slovak Current Research Information System. *Procedia Computer Science*, 211, 47–56. <https://doi.org/10.1016/j.procs.2022.10.175>
- [7] de Castro, P. (2025). Expanding the research graph underpinning the implementation of Open Science: research instruments and facilities. *Scottish Journal of Open Research*, 10(1). <https://doi.org/10.36399/dme15z31>
- [8] Latour, B., & Woolgar, S. (1979). *Laboratory Life: The Social Construction of Scientific Facts*. Beverly Hills: Sage Publications. 279 p.

- [9] Martin, P. et al. (2019). Mapping heterogeneous research infrastructure metadata into a unified catalogue for use in a generic virtual research environment. *Future Generation Computer Systems*, 101, 1–13. <https://doi.org/10.1016/j.future.2019.05.076>
- [10] Nikkanen J (2024). Handling research infrastructures within Research.fi. Autumn 2024 euroCRIS Strategic Membership Meeting (26-28 Nov, INRAE, Paris). <http://hdl.handle.net/11366/2653>
- [11] ERAC. (2023). ERAC Report on the role of Research and Innovation in Ukraine's Recovery. Retrieved from <https://data.consilium.europa.eu/doc/document/ST-10034-2023-INIT/en/pdf>
- [12] UNESCO. (2024). Analysis of war damage to the Ukrainian science sector and its consequences. Paris: UNESCO. Retrieved from <https://ukrainet.eu/wp-content/uploads/2024/03/388803eng.pdf>
- [13] Science at Risk. (2023). Forming a Database of Damaged and Destroyed Research Infrastructure in Ukraine. White Paper. Retrieved from <https://scienceatrisk.org/uk/whitepaper/bazy-danykh-poshkodzhenoi-ta-zruinovanoi-doslidnytskoi-infrastruktury/>
- [14] Cabinet of Ministers of Ukraine. (2022a). Resolution No. 1067 On Approval of the Provision on the National Electronic Scientific and Information System of 27 September 2022. Retrieved from <https://zakon.rada.gov.ua/laws/show/1067-2022-%D0%BF#Text>
- [15] Cabinet of Ministers of Ukraine. (2022b). National Plan for Open Science. Order No. 892-p of 8 October 2022. Retrieved from <https://zakon.rada.gov.ua/laws/show/892-2022-%D1%80#Text>
- [16] OpenAIRE. (2024). OpenAIRE Explore – Research Resources. Retrieved from <https://explore.openaire.eu/>
- [17] B2INST. (2024). B2INST – Instrument PID Registry Service. EUDAT / GWDG. Retrieved from <https://b2inst.gwdg.de/>
- [18] NAQA. (2025). Review on the quality of higher education in Ukraine 2024. National Agency for Quality Assurance in Higher Education. Retrieved from <https://en.naqa.gov.ua/wp-content/uploads/2025/06/REVIEW-ON-THE-QUALITY-OF-HIGHER-EDUCATION-IN-UKRAINE-2024.pdf>
- [19] Kaliuzhna, N., & Auhunas, S. (2022). Research information infrastructure in Ukraine: first steps towards building a national CRIS. *Procedia Computer Science*, 211, 230–237. <https://doi.org/10.1016/j.procs.2022.10.196>
- [20] de Rassenfosse, G. et al. (2023). The impact of the war on Ukrainian scientists. *Humanities and Social Sciences Communications*. <https://doi.org/10.1057/s41599-023-01651-3>
- [21] Law of Ukraine. (2016). On Scientific and Scientific-Technical Activity (No. 848-VIII). Verkhovna Rada of Ukraine. Retrieved from <https://zakon.rada.gov.ua/laws/show/848-19>
- [22] Drach, I., Borodiyenko, O., Petroye, O., Reheilo, I. Y., Bazeliuk, N., Slobodianiuk, O., & Kuzminska, O. H. (2023). Assessing the state of research e-infrastructures for open science in Ukrainian higher education institutions. In CEUR Workshop Proceedings (Vol. 3406, pp. 234–254). CEUR-WS.
- [23] Cabinet of Ministers of Ukraine. (2021). Concept of the State Target Programme for the Development of Research Infrastructures of Ukraine until 2026. Kyiv. Retrieved from <https://zakon.rada.gov.ua/laws/show/956-2021-%D1%80#Text>
- [24] Research Infrastructures for the Future of Ukraine (RIFF). (2025). Research infrastructure roadmap and strategic framework for Ukraine. RIFF Initiative. Retrieved from <https://riff.muni.cz/>
- [25] MESU. (2023). Research infrastructures. Ministry of Education and Science of Ukraine. Retrieved from <https://mon.gov.ua/ua/nauka/nauka/derzhavni-klyuchovi-laboratoriyi>
- [26] Zharinov, S., Zendulková, D., Tsybenko, I., & Zhrebchuk, S. (2024). The importance for Ukrainian R&D system RI's mapping and the Slovak case study. *Procedia Computer Science*, 249, 40–50. <https://doi.org/10.1016/j.procs.2024.11.048>

- [27] Hauschke, C., Nazarovets, S., Altemeier, F., & Kaliuzhna, N. (2021). Roadmap to FAIR research information in open infrastructures. *Journal of Library Metadata*, 21(1–2), 45–61. <https://doi.org/10.1080/19386389.2021.1999156>
- [28] euroCRIS. (2024). CERIF 1.6 Specification. Retrieved from <https://www.eurocris.org/cerif/main-features-cerif>
- [29] MERIL. (2022). Mapping of the European Research Infrastructure Landscape. Retrieved from <https://portal.meril.eu>
- [30] Stocker, M., Darroch, L., Krah, R., Habermann, T., Devaraju, A., Schwarzm, U., Huber, R., & Hggstrm, I. (2020). Persistent Identification of Instruments. *Data Science Journal*, 19(1), 1–12. <https://doi.org/10.5334/dsj-2020-018>
- [31] Research Data Alliance (2017). Persistent Identification of Instruments (PIDINST) WG Case Statement. <https://www.rd-alliance.org/groups/persistent-identification-instruments-wg/work-statement/>
- [32] Basnukaieva, R., & Postolnyi, B. (2024). LabsArena.com: Research Facilities. In A. Zharinova (Ed.), *Open Science and Innovation in Ukraine 2023: IInd International Conference* (pp. 14–15). Bentham Science Publishers.
- [33] EOSC Association. (2021). Strategic Research and Innovation Agenda (SRIA). Retrieved from <https://eosc.eu>
- [34] Auer, S., Kovtun, V., Prinz, M., Kasprzik, A., & Stocker, M. (2018). Towards a knowledge graph for science. *Proceedings of the 8th International Conference on Web Intelligence, Mining and Semantics (WIMS '18)*, 1–6. <https://doi.org/10.1145/3227609.3227689>
- [35] Dvořák, J., Chudlarský, T., & Spacek, J. (2019). Practical CRIS Interoperability. *Procedia Computer Science*, 146, 66–75. <https://doi.org/10.1016/j.procs.2019.01.068>
- [36] Engelman, A., Enkvist, C., & Pettersson, K. (2019). A FAIR archive based on the CERIF model. *Procedia Computer Science*, 146, 190–200. <https://doi.org/10.1016/j.procs.2019.01.076>
- [37] Jeffery, K., Houssos, N., Jörg, B., & Asserson, A. (2014). Research information management: The CERIF approach. *International Journal of Metadata, Semantics and Ontologies*, 9(1), 5–14. <https://doi.org/10.1504/IJMSO.2014.059142>
- [38] Jörg, B. (2010). ‘CERIF: The Common European Research Information Format Model’, *Data Science Journal*, 9(0), p. CRIS24-CRIS31. <https://doi.org/10.2481/dsj.CRIS4>
- [39] De-Castro P, Herb U, Rothfritz L, Schöpfel J (2023). Persistent identifiers for research instruments and facilities: an emerging PID domain in need of coordination. Knowledge Exchange case study. <https://doi.org/10.5281/zenodo.7330372>
- [40] Krah, R., Darroch, L., Huber, R., Devaraju, A., Klump, J., Habermann, T., Stocker, M., & The RDA PIDINST WG. (2022). Metadata Schema for the Persistent Identification of Instruments (1.0). Research Data Alliance. <https://doi.org/10.15497/RDA00070>