DCPAPERS

Bridging FAIR and CARE in ETD Metadata: An LLM-based Cross-Repository Evaluation Framework

Somesh Rai[1,2,†], Rajani Mishra[2,*,†]

[1] *Central University of Punjab, Bathinda 151401, India*
[2] *Banaras Hindu University, Varanasi 221005, India*

Abstract

Electronic Theses and dissertations that are submitted electronically are known as electronic theses and dissertations (ETDs), and they are distributed through a variety of institutional and aggregated repositories. Even though the majority of ETD platforms adhere to the FAIR principles, which specify that data must be Findable, Accessible, Interoperable, and Reusable, these platforms frequently fail to take into account the ethical and community-centered aspects that are encapsulated in the CARE principles, which are as follows: Collective Benefit, Authority to Control, Responsibility, and Ethics. The purpose of this research is to present a novel cross-repository evaluation framework that utilises an LLM-assisted technique to bridge the gap between the FAIR and CARE principles. A rubric-based review was combined with the reasoning skills of three big language models—ChatGPT, Grok, and DeepSeek R1—in order to conduct an evaluation of nine of the important open-access electronic text databases (ETD) repositories. A standardised rubric served as a guide for each model, and it was prompted to conduct an analysis of the quality of the metadata as well as ethical constraints. Despite the fact that the data demonstrate that FAIR compliance is resilient across repositories, they also highlight systemic weaknesses in CARE alignment, particularly with regard to cultural context, ethical reuse, and authorial control. Moreover, the comparative analysis among three agents suggests that it should be used for evaluating FAIR compliance. CARE compliance evaluation may need more sophisticated 'Human in the Loop' setup. This framework offers a scalable and transparent approach to analysing metadata governance. Additionally, it gives schema-agnostic recommendations for encouraging inclusivity and ethical stewardship in digital academic infrastructure. These characteristics are achieved through the triangulation of assessments given by artificial intelligence.

Keywords

Repository Evaluation, FAIR, CARE, LLM, ChatGPT, Grok, DeepSeek

1. Introduction

ETDs represent most of the early-career research, often exploring emerging, interdisciplinary, or culturally rooted novel topics. Therefore, its dissemination becomes increasingly important in digital and global landscape. It can be ensured only when their metadata is both technically interoperable and ethically grounded. While the FAIR principles [1] have guided improvements in data interoperability and reuse, the CARE principles [2] emphasize human agency, collective rights, and ethical governance—dimensions often absent in digital library metadata practices.

* Corresponding author.
† These authors contributed equally.
somesh.rai15@bhu.ac.in (S. Rai); rajanimishra5@gmail.com (R. Mishra)
0000-0003-0322-3915 (S. Rai); 0000-0002-1507-4665 (R. Mishra)

https://doi.org/10.23106/dcmi.952576294

This paper assesses how ETD repositories operationalize both principles and proposes a path toward more inclusive metadata design.

2. Background

The FAIR principles are designed to ensure that metadata enables efficient discovery and reuse, particularly by machines. They have gained significant traction in research data management, shaping global standards and influencing repository policies [1]. In contrast, the CARE principles, introduced by the Global Indigenous Data Alliance (GIDA), emerged from Indigenous data sovereignty movements and emphasize ethical engagement, power dynamics, and benefit-sharing in data practices [2].

Electronic Theses and Dissertations (ETD) repositories are notably diverse in terms of architecture and metadata schemas. Institutional repositories based on platforms such as DSpace and Vireo typically utilize metadata schemas like Dublin Core or MODS [3, 4]. Meanwhile, aggregators such as PQDT Global and OATD often rely on minimal or standardized metadata profiles, which may limit semantic richness and interoperability. Despite the critical importance of metadata in discovery and reuse, studies and evaluations have predominantly focused on FAIR principles, often overlooking CARE-oriented assessments. There remains a pronounced gap in the literature examining how repository architectures and schema choices align with both FAIR and CARE principles. While FAIR evaluations emphasize machine-readability and technical structure, CARE critiques draw attention to representational equity, cultural context, and ethical reuse [5]. Furthermore, the potential of AI-assisted automated evaluation of repositories in this dual-framework context has been significantly under-explored, despite its promise for scalable, repeatable assessment methodologies [6].

3. Methodology

We developed a scoring rubric with 16 FAIR points and 10 CARE points, broken into 4 subdimensions each. Repositories were evaluated based on public (accessible to AI agents) metadata schemas, documentation, and sample records. Following are the truncated codes used for various subdimensions throughout.

Sl. No.	Subdimension Codes	Explanation
1	*F_PI: Findable*	*Persistent Identifiers (e.g., DOIs, Handles)*
2	*F_MR: Findable*	*Metadata Richness (e.g., titles, authors, abstracts)*
3	*A_OA: Accessible*	*Open Access Policies (e.g., open access with minimal restrictions)*
4	*A_ML: Accessible*	*Metadata Longevity (e.g., retrievable via stable protocols like OAI-PMH)*
5	*I_US: Interoperable*	*Use of Standards (e.g., Dublin Core, MODS)*
6	*I_VR: Interoperable*	*Vocabulary Reuse (e.g., LCSH, ORCID)*
7	*R_LM: Reusable*	*Licensing Metadata (e.g., CC-BY)*
8	*R_MP: Reusable*	*R_MP: Reusable - Metadata Provenance (e.g., documentation of creation/modification)*

https://doi.org/10.23106/dcmi.952576294

9	*CB_CV: Collective Benefit*	*Community Visibility (e.g., enhances visibility for underrepresented groups)*
10	*AC_AR: Authority to Control*	*Author Rights & Restrictions (e.g., authors can specify access/usage rights)*
11	*AC_IC: Authority to Control*	*Indigenous/Community Tags (e.g., provisions for cultural affiliations)*
12	*R_EU: Responsibility*	*Ethical Usage Metadata (e.g., ethical considerations documented)*
13	*E_PR: Ethics*	*Provenance and Representation (e.g., system to track provenance and address misrepresentation)*

3.1. Repositories Evaluated:

- ❖ Institutional:
  - MIT DSpace
  - TDL
  - Theses Canada
  - Shodhganga
- ❖ Aggregators:
  - OATD
  - PQDT Global
  - EBSCO Open
  - DART-Europe
  - EThOS

3.2. Schema Grouping

- ❖ Dublin Core (DC): Used by 7 repositories
- ❖ Non-DC: MODS, ETD-MS, proprietary (e.g., PQDT)

3.3. Scoring

- **0** – Does **not** meet the indicator (i.e., there is no evidence of that feature in the metadata or policy).
- **1** – Partially meets the indicator (some evidence, but incomplete or inconsistent).
- **2** – Fully meets the indicator (clear, consistent, and complete support).

4. Results

4.1. Iteration 1: ChatGPT (o4-mini)

4.1.1. FAIR alignment

Most repositories scored highly on FAIR (12–14/16), but CARE alignment was low (1–4/10). The heatmaps (see Figures 1–2) reveal consistent implementation of FAIR principles (persistent identifiers, access protocols, standardized formats) and widespread neglect of CARE dimensions like ethical use and cultural context.

Figure 1 presents the FAIR subdimension evaluation across nine digital repositories, capturing their relative performance in alignment with the FAIR principles. Each repository

https://doi.org/10.23106/dcmi.952576294

was scored on eight subdimensions, ranging from persistent identifier usage to metadata provenance documentation.

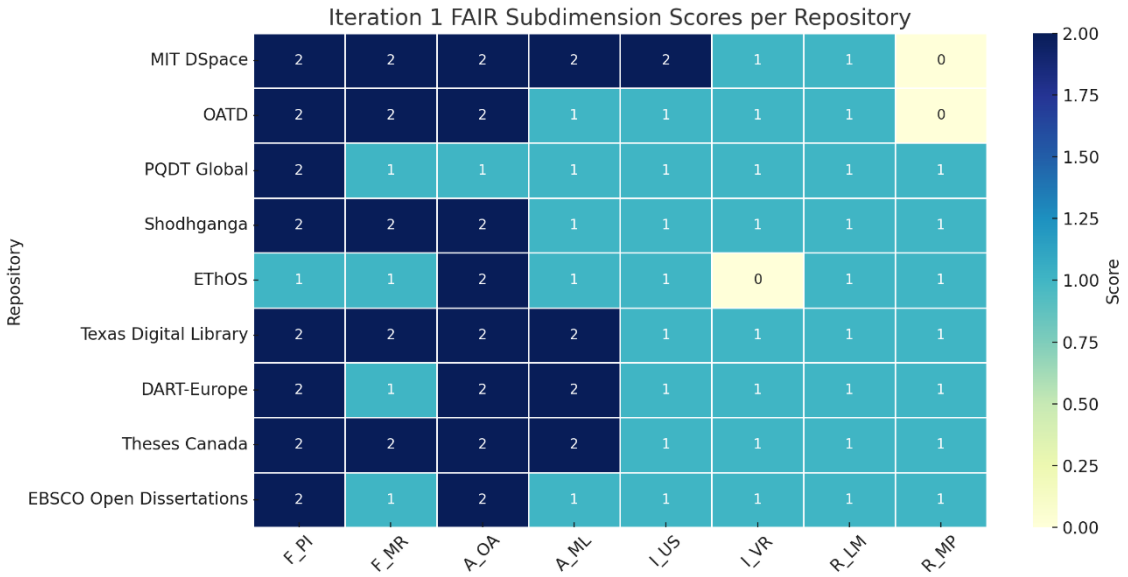

Figure 1: FAIR Scores for evaluated ETD Repositories in Iteration 1

MIT DSpace, Shodhganga, Texas Digital Library, and Theses Canada emerged as strong performers, scoring 2 (maximum) across most subdimensions. These repositories consistently utilized persistent identifiers (F_PI), provided rich metadata (F_MR), supported standardized access protocols (A_ML), and included clear licensing information (R_LM). Their adherence to recognized metadata standards and controlled vocabularies also contributed to higher interoperability scores (I_US, I_VR).

In contrast, repositories like EThOS and EBSCO Open Dissertations displayed moderate to low FAIR alignment. While many supported open access (A_OA), they fell short in areas such as vocabulary reuse (I_VR) and metadata provenance (R_MP), with several scoring 0 in the latter—indicating a lack of transparency in metadata creation and revision history.

Aggregators such as OATD and DART-Europe showed mixed performance, often reflecting variability inherited from their source repositories. These platforms generally performed well in accessibility but scored inconsistently on findability and interoperability due to non-uniform metadata quality.

The heatmap illustrates that while some institutional repositories have well-structured metadata infrastructures in place, others—particularly aggregators and those with legacy systems—require targeted improvements in metadata richness, interoperability, and provenance tracking to align fully with FAIR best practices.

https://doi.org/10.23106/dcmi.952576294

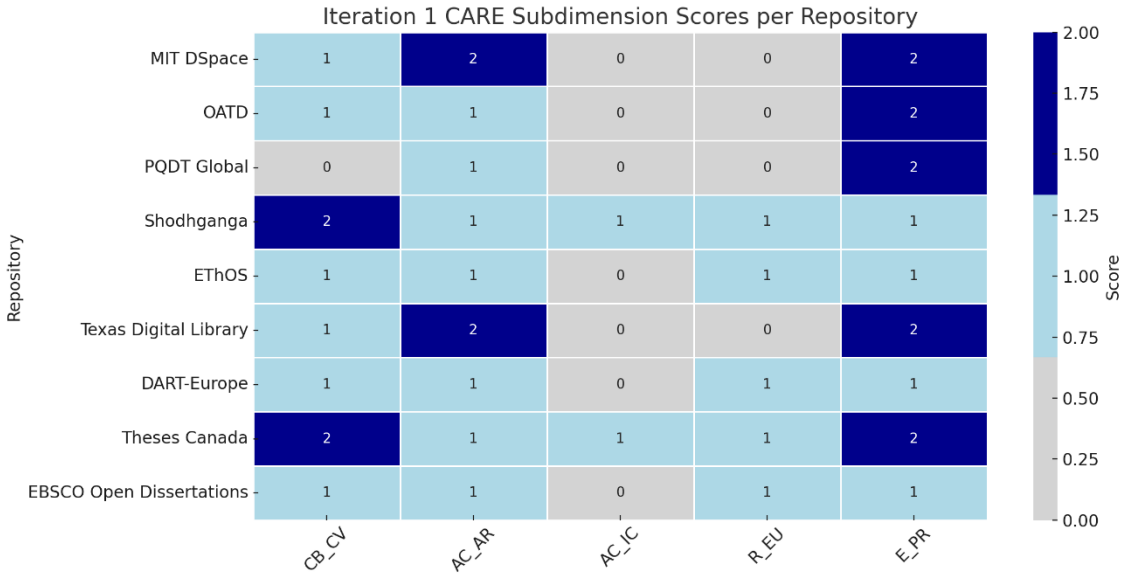

Figure 2: CARE Scores for evaluated ETD Repositories in Iteration 1

Figure 2 presents the Iteration 1 CARE subdimension scores for nine digital repositories, evaluating their alignment with the principles of Collective Benefit, Authority to Control, Responsibility, and Ethics as adapted from the CARE Principles for Indigenous Data Governance.

Across the board, CARE compliance was markedly lower and more variable than FAIR compliance. Most repositories scored 1 out of 2 in subdimensions such as Community Visibility (CB_CV) and Ethical Provenance (E_PR), reflecting baseline acknowledgment of academic value and limited efforts to track data lineage. However, no repository scored a perfect 2 across all CARE dimensions, and several consistently scored 0 in key ethical and cultural domains.

A particularly notable pattern is the universal absence of Indigenous or community-specific metadata tagging (AC_IC). All repositories except Shodhganga and Theses Canada scored 0 in this category, indicating a widespread lack of support for cultural identifiers, heritage metadata, or community-level attribution frameworks. This deficiency significantly limits the Authority to Control for underrepresented communities and undermines the CARE principle of Representation.

MIT DSpace, PQDT Global, and Texas Digital Library demonstrated relatively better compliance in Author Rights (AC_AR) and Ethical Provenance (E_PR), likely due to established licensing workflows and metadata versioning. However, these strengths are largely procedural rather than relational and do not substitute for culturally aware governance models.

The repositories EThOS, OATD, and DART-Europe showed limited CARE alignment, reflecting systemic issues in aggregator models that inherit inconsistencies from upstream sources. In these cases, the absence of structured mechanisms for community governance, ethical metadata tracking, and inclusive metadata schemas is particularly apparent.

While some repositories show preliminary structures supporting individual rights (e.g., licensing), the broader values of collective governance and ethical stewardship are underdeveloped. This analysis underscores the need for repository frameworks that go beyond

https://doi.org/10.23106/dcmi.952576294

access and discoverability to embed cultural, ethical, and representational accountability within metadata infrastructures.

4.2. Iteration 2: DeepSeek (DeepThink R1)

4.2.1. FAIR alignment

Figure 3 illustrates the second-round FAIR subdimension scores for nine repositories, revealing a shift in evaluative emphasis and the application of more critical scoring criteria. Compared to Iteration 1, DeepSeek seems to have stricter judgments in areas such as Metadata Provenance (R_MP), Vocabulary Reuse (I_VR), and Interoperability Standards (I_US) resulted in overall lower or more variable scores across several repositories.

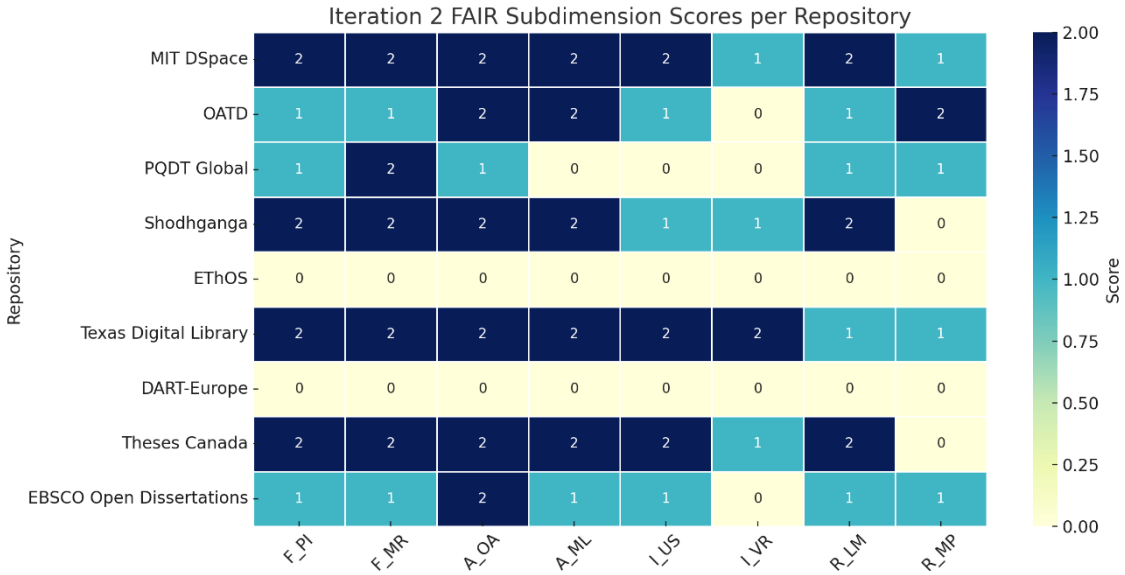

Figure 3: FAIR Scores for evaluated ETD Repositories in Iteration 2

EThOS and DART-Europe received uniformly low scores (0 across all subdimensions), reflecting either a reassessment of previously assumed capacities or the lack of publicly visible FAIR-supporting infrastructure such as persistent identifiers, open protocols, or metadata richness. These results highlight an ongoing issue in aggregated or institutionally inconsistent repositories where technical compliance is not uniformly enforced.

In contrast, MIT DSpace, Texas Digital Library, and Theses Canada maintained high performance across all eight subdimensions, reaffirming their robust infrastructure for persistent identifiers, licensing metadata, and protocol support. Shodhganga remained strong overall, though it dropped to a zero in R_MP, signaling a reassessment of its metadata tracking visibility.

PQDT Global's scores reflect a particular vulnerability in interoperability (I_US = 0, I_VR = 0) and accessibility (A_ML = 0), underscoring the limitations of proprietary metadata frameworks that lack openness and standardization. OATD's scores, while moderate overall, revealed improvement in R_MP (score of 2), possibly indicating better provenance mechanisms in selected source repositories.

https://doi.org/10.23106/dcmi.952576294

The Iteration 2 analysis captures a more critical and refined appraisal of FAIR compliance, emphasizing that the presence of access or metadata fields alone is insufficient—what matters is standardization, persistence, and machine-actionable transparency.

4.2.2. CARE alignment

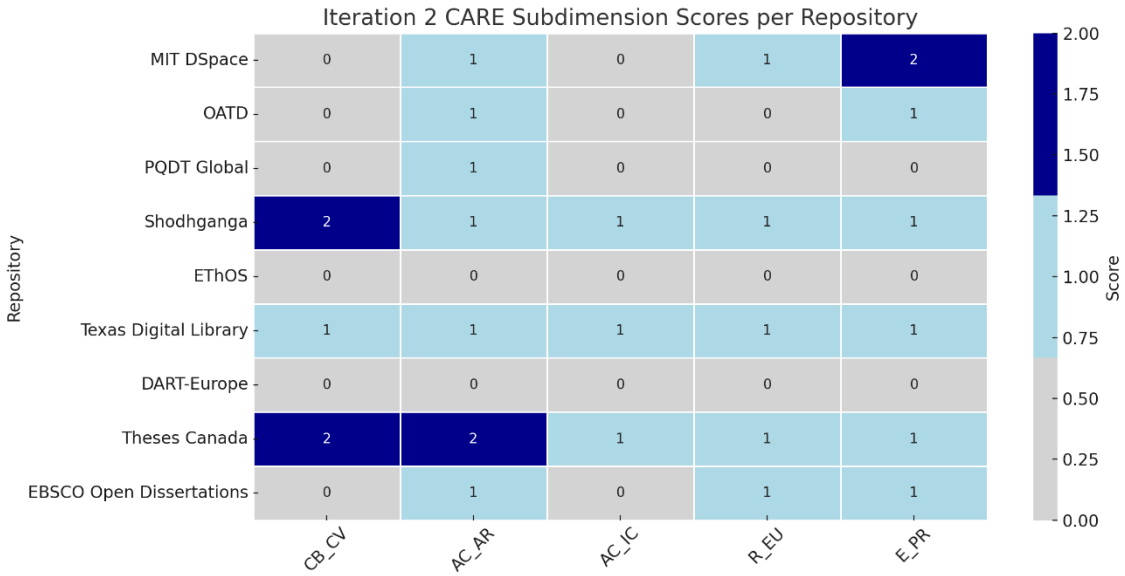

Figure 4: CARE Scores for evaluated ETD Repositories in Iteration 2

Figure 4 presents the second iteration of CARE evaluations, showing a wider distribution of values than in Iteration 1, though still reflecting foundational limitations in CARE principle integration across most repositories. It seems this iteration applies a refined lens to each subdimension, particularly in identifying the explicit absence or token presence of community, ethical, and representational metadata.

MIT DSpace remains among the stronger repositories, but its score decreased slightly in Authority to Control (AC_AR = 1) and Responsibility (R_EU = 1), suggesting only partial support for ethical data use and licensing flexibility. OATD, PQDT Global, and EBSCO all scored lowest in multiple subdimensions, confirming that aggregation models and proprietary infrastructures remain underdeveloped in terms of ethical metadata frameworks and culturally sensitive governance.

Interestingly, Shodhganga and Theses Canada continued to perform better than their peers, achieving scores of 1 or 2 in most subdimensions. However, even these repositories do not demonstrate full CARE compliance—particularly in the absence of structured Indigenous tags or ethical tracking mechanisms beyond basic provenance.

The zero scores for EThOS and DART-Europe indicate a complete lack of observable CARE-related metadata elements or governance provisions. These repositories reflect the systemic neglect of CARE principles in digital scholarship infrastructures, particularly in regions where repository development is driven by policy mandates rather than community engagement.

https://doi.org/10.23106/dcmi.952576294

4.3. Iteration 3: Grok 3 (with DeepSearch)

4.3.1. FAIR alignment

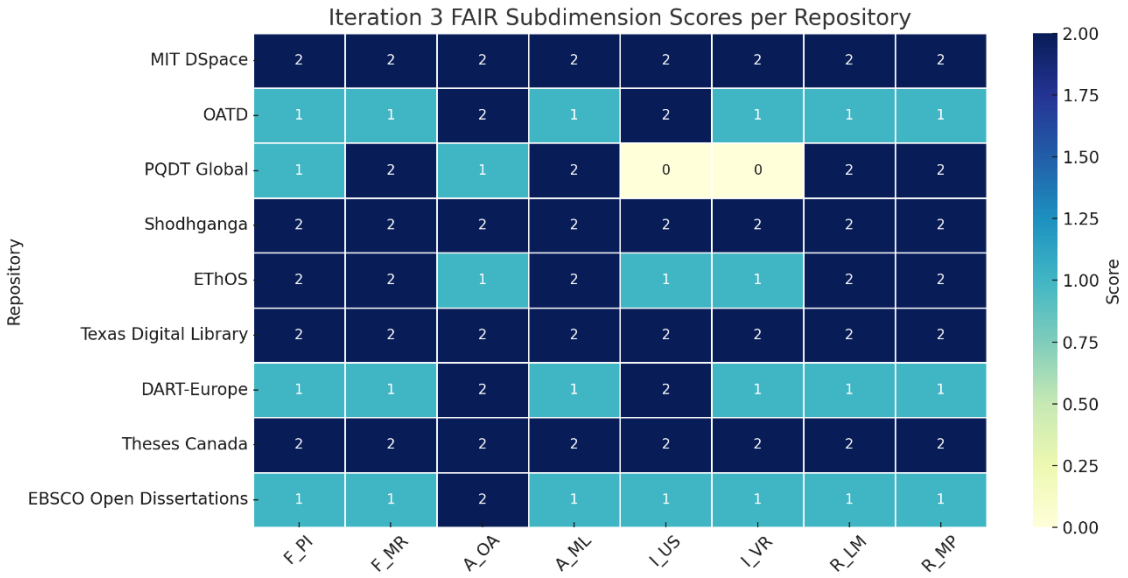

Figure 5: FAIR Scores for evaluated ETD Repositories in Iteration 3

Figure 5 highlights the third iteration of FAIR subdimension scoring, revealing a significant recovery and recalibration in the assessment of technical compliance. MIT DSpace, Shodhganga, Texas Digital Library, and Theses Canada each scored 2 across all subdimensions, reaffirming their comprehensive FAIR alignment. Their metadata infrastructures support persistent identifiers, interoperability standards, controlled vocabularies, and clear reuse policies, positioning them as reference models for other institutional repositories.

EThOS and DART-Europe, which scored poorly in Iteration 2, showed a dramatic rebound, particularly EThOS, which scored 2 in six out of eight dimensions. This improvement may be reflecting either the discovery of previously unassessed infrastructure or a revised interpretation of documentation and accessibility mechanisms.

PQDT Global displayed improvement in accessibility (A_ML = 2) and reusability (R_LM, R_MP = 2), suggesting enhancements in its metadata output or better articulation of licensing and provenance practices. However, its interoperability scores (I_US, I_VR) remained at 0, reiterating the limitations of proprietary metadata.

OATD and EBSCO maintained moderate scores, typically scoring 1 across most categories. This pattern is consistent with their aggregator model, which limits uniform implementation of technical FAIR standards.

This iteration reflects mature evaluator calibration, yielding scores that more accurately distinguish between technically capable, standards-compliant repositories and those still needing substantial development in openness, interoperability, and machine-actionable metadata.

https://doi.org/10.23106/dcmi.952576294

4.3.2. CARE alignment

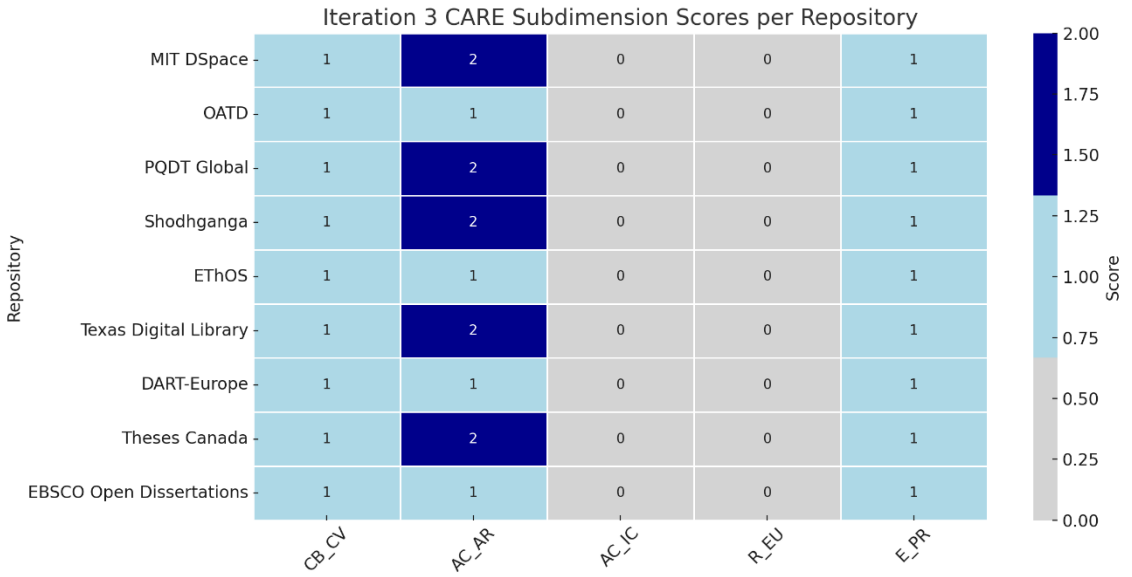

Figure 6: CARE Scores for evaluated ETD Repositories in Iteration 3

Figure 6 presents the third-round CARE subdimension scores, marking the most consistent scoring pattern observed across all iterations. Unlike previous iterations where some repositories showed strong community visibility or author rights, Iteration 3 reveals a relatively flat landscape, with most repositories achieving mid-level scores (1–2) in CB_CV and AC_AR, and consistently low or zero in AC_IC and R_EU.

MIT DSpace, PQDT Global, Shodhganga, and Theses Canada received a score of 2 in AC_AR (author rights), reflecting clear licensing policies and user-defined access control. However, no repository scored above 0 in AC_IC (Indigenous/Community Tags) or R_EU (Ethical Usage), suggesting a systemic failure to incorporate culturally contextual metadata or ethical usage disclosures.

The consistent presence of score 1 in E_PR (Ethical Provenance) across all repositories indicates a basic but generic awareness of provenance tracking, often rooted in standard metadata versioning rather than culturally specific or ethically grounded practices.

The uniformity of this iteration's CARE scores suggests two things: (1) reviewers may have converged on more consistent criteria, and (2) the repositories, regardless of geographical or institutional scope, lack the structural elements necessary for CARE compliance beyond basic academic licensing and attribution mechanisms.

4.4. Comparative Analysis (ChatGPT vs DeepSeek vs Grok)

Figure 7 shows the standard deviation in FAIR and CARE scores across the three evaluators (CGPT, DPSK, GRK) for each repository. FAIR Scores show that there's more disagreement for repositories like *EThOS*, *PQDT Global*, and *Shodhganga*, indicating varied interpretations of their FAIR compliance. CARE Scores in general, show low variation, suggesting a consensus on the limited CARE compliance across repositories. One noticeable observation can be seen that OATD's FAIR scores are unanimous.

https://doi.org/10.23106/dcmi.952576294

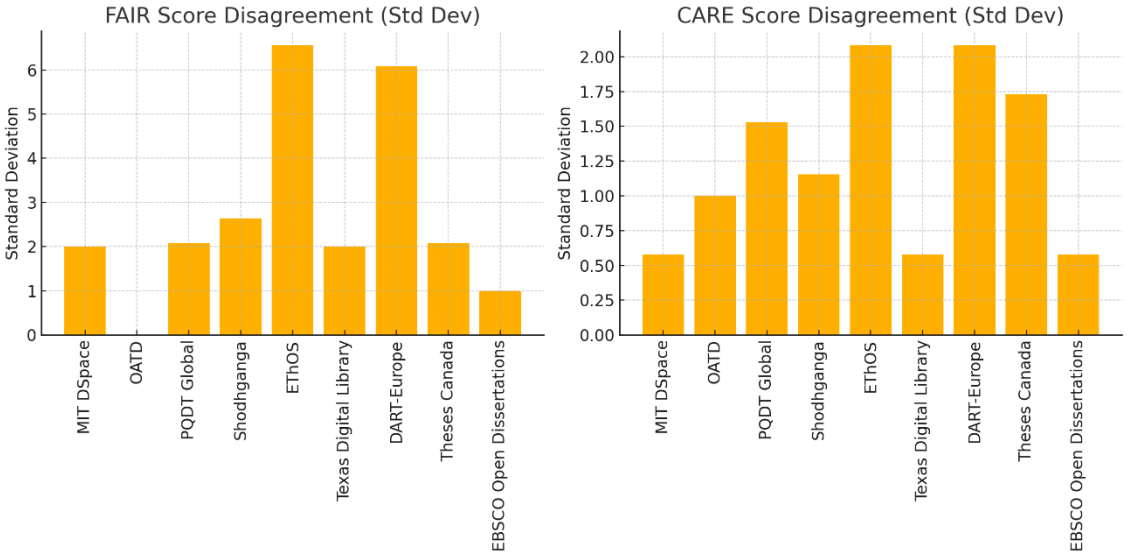

Figure 7: FAIR and CARE Scores for ETD Repositories

Table 2

Subdimension No.	Subdimension	Mean_STD_Across_Repositories
7	R_MP	2.716587
6	R_LM	0.748006
4	I_US	0.57735
8	CB_CV	0.5132

Table 2 highlights subdimensions with the greatest disagreement among evaluators. **R_MP (Metadata Provenance)** shows the **highest variability** (std dev ≈ 2.72), suggesting inconsistent interpretations of how well repositories track metadata provenance. Other FAIR subdimensions like **R_LM** and **I_US** show moderate disagreement. CARE subdimensions (e.g., **CB_CV**) show generally lower variability, reflecting more evaluator agreement.

4.5. Schema-Based Insights

Dublin Core repositories benefit from simplicity and interoperability, enabling consistent FAIR scores. However, they lack fields or extensions to support CARE principles without customizations. Non-DC repositories, while sometimes richer in detail (e.g., MODS), rarely operationalize CARE fields natively, relying instead on external editorial controls.

4.6. Institutional vs Aggregator Contrast

Institutional repositories have ability to offer better support for authorial control and licensing, aligning modestly with CARE principles. Aggregators prioritize normalization and scale, resulting in minimal ethical or contextual metadata.

5. Discussion

These findings confirm that ETD repositories have evolved to meet technical standards of interoperability and access, particularly those aligned with digital preservation and scholarly

https://doi.org/10.23106/dcmi.952576294

communication infrastructure. However, they fall short in integrating socially responsive and community-aware metadata practices.

Across both schema groups, CARE subdimensions such as Collective Benefit and Ethics scored uniformly low. While DC-based repositories enable basic author control, they lack formal support for ethical or community-specific metadata. Non-DC repositories showed slightly better implementation of authority-to-control via licensing and usage restrictions, yet typically enforced at the platform level, not embedded in the schema.

Sometimes relying on minimal metadata profiles, aggregator platforms ignore licensing or cultural background and limit the author's ability to change or improve metadata after submission. By greater control over metadata policies and more author involvement, institutional repositories are more likely to follow CARE principles than opposite.

6. Recommendations

- Schema Extension, CARE-related fields (e.g., cultural identity, ethical use statements) should be added to DC/MODS profiles.
- Ethical Licensing, structured fields for access restrictions, Indigenous consent, and community benefit sharing should be added.
- Author Empowerment, post-submission metadata updates and granular licensing controls should be Allowed.
- CARE training modules can be developed for repository managers on equitable and inclusive metadata practices.
- Framework Adoption, Automated FAIR+CARE audits and visual dashboards for institutional and national repositories can be implemented.
- Hybrid model may be a better approach to apply AI agents for FAIR compliance, complemented by community-led, expert-driven review processes for CARE.
- Development of explainable AI tools that assist but do not replace human judgment, especially in culturally sensitive contexts.

7. Conclusion

Metadata reflects the values embedded in the structure of the academic community, so enhancing its role as a tool for discoverability. This study demonstrates that while FAIR principles are well-established in ETD systems, CARE principles remain inadequately implemented, particularly in scenarios involving large-scale aggregators. To ensure that the future of digital scholarship is both technically interoperable and ethically responsible, it will be necessary to innovate schemas, make a commitment from institutions, and engage in continuing community co-design in order to align metadata with both sets of ideals. To conclude it can be said that AI agents may serve as an efficient tool for evaluation of FAIR and CARE compliance, provided that one uses the principles of 'Human in the loop' and 'structured-detailed evaluation rubrics' for evaluation to minimize the disagreements among various agents. Furthermore, it can be validated over a larger sample of ETD repositories.

https://doi.org/10.23106/dcmi.952576294

Acknowledgements

ChatGPT 4o, Deepseek R1 and Grok 3 have been used in the methodology part of the study. All the content generated out of AI agents has been carefully reviewed by the authors.

References

[1] Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., ... & Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data, 3, 160018. https://doi.org/10.1038/sdata.2016.18

[2] Global Indigenous Data Alliance (GIDA). (2019). CARE Principles for Indigenous Data Governance. Retrieved from https://www.gida-global.org/care

[3] Kunze, J. A., Lagoze, C., & Van de Sompel, H. (2003). The Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH). *DCMI Conference Paper Archive*. Retrieved from https://dcpapers.dublincore.org/files/articles/952137093/dcmi-952137093.pdf

[4] Texas Digital Library. (2022). *ETD Metadata Working Group Report: Evaluation and Recommendation for Vireo 4 Metadata Schema.* Retrieved from https://tdl-ir.tdl.org/bitstreams/6f875047-02ea-4a1d-a816-ffa8dc884404/download

[5] Christen, K. (2012). Does information really want to be free? Indigenous knowledge systems and the question of openness. International Journal of Communication, 6, 2870–2893. https://ijoc.org/index.php/ijoc/article/view/1618

[6] Lagoze, C., Krafft, D. B., Payette, S., & Jesuroga, S. (2006). What is a digital library anymore, anyway? Beyond search and access in the NSDL. D-Lib Magazine, 11(11). https://www.dlib.org/dlib/november05/lagoze/11lagoze.html

https://doi.org/10.23106/dcmi.952576294

A. FAIR+CARE Evaluation Rubric

Principle/Criterion	Indicator	Guiding Question
FAIR / Findable	Persistent Identifiers	Does each thesis have a globally unique and persistent identifier?
	Metadata Richness	Are titles, authors, subjects, and abstracts consistently present and indexed?
FAIR / Accessible	Open Access Policies	Are theses openly accessible with minimal restrictions?
	Metadata Longevity	Is metadata retrievable via stable protocols (e.g., OAI-PMH)?
FAIR / Interoperable	Use of Standards	Are metadata expressed in standard formats (Dublin Core, MODS, ETD-MS)?
	Vocabulary Reuse	Are established vocabularies used (e.g., LCSH, ORCID)?
FAIR / Reusable	Licensing Metadata	Is licensing (e.g., CC-BY) clearly specified?
	Metadata Provenance	Is there information about how/when metadata was created or modified?
CARE / Collective Benefit	Community Visibility	Does the metadata support the discoverability of underrepresented communities?
CARE / Authority to Control	Author Rights & Restrictions	Can authors control access, attribution, or reuse conditions?
	Indigenous/Community Tags	Are there provisions for cultural identity metadata or tribal affiliations?
CARE / Responsibility	Ethical Usage Metadata	Are ethical restrictions or intended use contexts captured in metadata?
CARE / Ethics	Provenance and Representation	Are there mechanisms to trace data origin and correct misrepresentation?

B. Prompt Sequence

Prompt No	Prompt Text
Prompt 1	**We** developed a scoring rubric with 16 FAIR points and 10 CARE points, broken into 4 subdimensions each. Repositories were evaluated based on public (accessible to AI agents) metadata schemas, documentation, and sample records. I am Providing the two rubrics in CSV. First understand this then I will provide you with specific Repository Evaluation Task. FAIR Grading Rubrics,,,,Principle,Criterion,Indicator,Guiding Question,Score (0–2)Findable,Persistent Identifiers,"Use of DOIs, Handles, or other PIDs",Does each thesis have a globally unique and persistent identifier?,,Metadata Richness,"Inclusion of titles,

https://doi.org/10.23106/dcmi.952576294

	authors, subjects, abstracts",Are metadata elements comprehensive and detailed?,Accessible,Open Access Policies,Clear access policies and minimal restrictions,Are theses openly accessible with minimal restrictions?,,Metadata Longevity,"Metadata retrievability via stable protocols (e.g., OAI-PMH)",Is metadata retrievable through standardized protocols?,Interoperable,Use of Standards,"Adoption of standard formats (e.g., Dublin Core, MODS)",Are metadata expressed in recognized standard formats?,,Vocabulary Reuse,"Utilization of established vocabularies (e.g., LCSH, ORCID)",Are controlled vocabularies employed for metadata elements?,Reusable,Licensing Metadata,"Clear specification of licensing (e.g., CC-BY)",Is licensing information explicitly stated in the metadata?,,Metadata Provenance,Documentation of metadata creation and modification,Is there information about how and when metadata was created or modified?,,,,Total,0-16 CARE Grading Rubrics,,,,Principle,Criterion,Indicator,Guiding Question,Score (0–2)Collective Benefit,Community Visibility,Support for discoverability of underrepresented communities,Does the metadata enhance visibility for underrepresented groups?,Authority to Control,Author Rights & Restrictions,Mechanisms for authors to control access and reuse,Can authors specify access restrictions or usage rights?,,Indigenous/Community Tags,Inclusion of cultural identity metadata or tribal affiliations,Are there provisions for indicating cultural or community affiliations?,Responsibility,Ethical Usage Metadata,Inclusion of ethical use statements or intended use contexts,Are ethical considerations documented in the metadata?,Ethics,Provenance and Representation,Mechanisms to trace data origin and correct misrepresentation,Is there a system to track provenance and address misrepresentation?,,,,Total,0-10
Prompt 2	**Following** are the selected repositories you need to evaluate and provide 3 types of analysis in tabulated format, 1. FAIR and CARE alignment 2. Schema-Based Insights 3. Institutional vs Aggregator Contrast. Repository,URL,Type,Metadata Schema,Access Model,Present StatusMIT DSpace,https://dspace.mit.edu/handle/1721.1/7582,Institutional,Dublin Core,Open Access,OATD,https://oatd.org,Aggregator,Dublin Core (varies),Open Access,PQDT Global,https://www.proquest.com/?defaultdiss=true,Commercial/Aggregator,Proprietary Enhanced Metadata,Mixed (subscription & OA),Shodhganga,https://shodhganga.inflibnet.ac.in,National Repository (India),ETD-MS / Dublin Core,Open Access,EThOS (UK),,National Repository (UK),Dublin Core / Institutional Standards,Mixed (mostly OA),Not Active Due to Cyber attackTexas Digital Library (TDL),https://hdl.handle.net/2249.1/9387,Consortium/Institutional,Dublin Core / MODS,Open Access,DART-Europe,https://www.dart-europe.org,Aggregator (Europe),Institutional Standards (mostly Dublin Core),Open Access,ClosedTheses Canada,https://recherche-collection-search.bac-lac.gc.ca/eng/Help/theses,National Repository (Canada),MODS / MARCXML,Open Access,EBSCO Open Dissertations,https://opendissertations.org/,Aggregator,Institutional Standards (varied),Open Access,
Prompt 3	**Provide** subdimension (as per each guiding question) scores of FAIR and CARE for all repositories in tabular form

https://doi.org/10.23106/dcmi.952576294