

# Fit-for-Purpose PID Adoption in a National Research Data Platform: Balancing FAIR Interoperability and CARE-aligned Governance in DataON

Hea Lim Rhee<sup>1</sup>  

<sup>1</sup> Korea Institute of Science and Technology Information

## Abstract

Persistent identifiers (PIDs) are widely recognized as essential metadata infrastructure for enabling the findability, citation, and reuse of research data. However, practical guidance on how national-scale platforms should adopt and govern PIDs—particularly when balancing global interoperability with locally accountable stewardship—remains limited. This poster presents the PID adoption framework of DataON, Korea's national research data platform operated by KISTI. DataON employs a “fit-for-purpose” strategy combining multiple identifier layers under two operational scenarios: direct registration and API-based integration. I analyze this framework through the dual lens of the FAIR Principles and governance values drawn from the CARE Principles for Indigenous Data Governance, and describe DataON's strategic direction toward PID-based knowledge graphs.

## 1. Introduction

Persistent identifiers (PIDs) have become foundational metadata infrastructure in research data ecosystems. By providing globally unique, stable, and machine-resolvable references to datasets, researchers, and organizations, PIDs address problems of dataset ambiguity, link rot, and unreliable citation [1]. The FAIR Guiding Principles explicitly require globally unique and persistent identifiers for data and metadata, while the CoreTrustSeal certification evaluates whether repositories manage identifiers reliably over time [2].

Yet the practical challenge of PID implementation at a national scale extends beyond technical registration. Decisions about what to identify, who assigns identifiers, how external identifiers are ingested, and what governance structures oversee these processes are fundamentally policy questions. The CARE Principles for Indigenous Data Governance—emphasizing Collective Benefit, Authority to Control, Responsibility, and Ethics—offer a valuable conceptual framework for such governance dimensions [3]. While CARE originated in Indigenous data sovereignty, its emphasis on accountable authority and purposeful stewardship resonates with challenges faced by national research data platforms balancing global openness with locally responsible management [4]. This poster presents DataON's PID framework and analyzes it through the complementary lenses of FAIR and CARE.

## 2. DataON: Platform Context

DataON is Korea's national research data platform, operated by KISTI. Launched in 2020, it supports the full research data lifecycle—from registration and publication to

## Poster

Available Aug 01, 2026

Correspondence to  
Hea Lim Rhee  
rhee.healim@kisti.re.kr



Copyright © 2026 Rhee. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license, which enables reusers to distribute, remix, adapt, and build upon the material in any medium or format, so long as attribution is given to the creator.

access and reuse—serving researchers, institutions, and national R&D programs [5]. DataON holds CoreTrustSeal certification, confirming that it manages data and identifiers according to international standards for trustworthy stewardship. PIDs are treated as a baseline requirement: every dataset entering the system is assigned or mapped to at least one persistent identifier.

### 3. Fit-for-Purpose PID Architecture

DataON's PID strategy follows a “fit-for-purpose” principle: identifiers are assigned selectively based on utility for reuse, citation, and management. The primary target is the research dataset and its versions; related entities receive PIDs only when there is clear operational need.

The platform manages PIDs under two scenarios. In *Case 1 (Direct Registration)*, DataON assigns and manages identifiers during ingestion. Three layers operate: Data Registry Identifiers (DRIDs) for internal lifecycle tracking; DOIs assigned through the Korea DOI Center for global citation; and Service Identifiers (SVCIDs) for platform-level management. Researchers also input their National Researcher Number directly. In *Case 2 (API-based Integration)*, DataON ingests datasets with existing PIDs from partner institutions. DOIs already assigned by affiliates are preserved rather than re-minted, respecting originating stewardship. Depositors' local identifiers are retained for provenance, and ORCID iDs are received to link datasets to researchers across boundaries. DataON thus functions as an integrator, accepting identifiers, preserving their meaning, and embedding them within a national metadata framework.

### 4. Analyzing PID Governance through FAIR and CARE

The FAIR lens highlights how each identifier layer contributes to findability (globally resolvable DOIs and SVCIDs mapped to landing pages), accessibility (stable resolution), interoperability (cross-system harvesting via standardized PIDs), and reusability (citation paths and version tracking). The multi-layered approach ensures FAIR compliance at each operational level.

The CARE lens illuminates governance dimensions FAIR does not fully address. While developed for Indigenous data governance, CARE's core values—particularly Authority to Control and Responsibility—resonate with national platform governance. In DataON, Authority to Control is reflected in national coordination of DOI assignment through the Korea DOI Center, the sovereign determination of which entities warrant global versus local identifiers, and preservation of partner institutions' identifier authority in Case 2. Responsibility is embodied in CoreTrustSeal-certified stewardship and long-term persistence commitments. Collective Benefit manifests in serving Korea's research community while contributing to the global PID ecosystem. Rather than claiming direct CARE implementation, I propose that CARE's governance-oriented principles offer a useful analytical complement to FAIR's data-centric focus [4].

## 5. Strategic Direction: PID-Based Knowledge Graphs

DataON's strategic direction extends PID infrastructure toward knowledge graph construction. KISTI manages PIDs for research outputs, researchers, and organizations. When these identifiers serve as globally resolvable nodes with semantically defined relationships—createdBy (dataset DOI to ORCID), fundedBy (to grant identifier), hostedBy (to organizational ROR)—a knowledge graph emerges, enabling impact tracking, research lineage reconstruction, and provenance verification at a national scale [6].

## 6. Challenges and Lessons Learned

Key challenges included encouraging researcher adoption (technical infrastructure alone proved insufficient without clear incentives), developing versioning policies, remediating legacy datasets predating PID practices, and aligning national governance requirements with evolving international standards. The central lesson is that PID implementation is as much about policy and workflow as technology [7].

## 7. Conclusion

DataON's experience demonstrates that PIDs function most effectively when treated as infrastructure—embedded from the start, governed by clear policy, and sustained by institutional commitment. The dual-lens analysis through FAIR and CARE-aligned values reveals that robust PID adoption requires both machine-actionable interoperability and human-centered governance. National platforms bridge local governance and international standards; their identifier design choices carry implications for trust, discovery, and reuse at scale.

## References

- [1] M. D. Wilkinson, M. Dumontier, and I. J. Albersberg, "The FAIR Guiding Principles for scientific data management and stewardship," *Scientific Data*, vol. 3, p. 160018, 2016, doi: [10.1038/sdata.2016.18](https://doi.org/10.1038/sdata.2016.18).
- [2] CoreTrustSeal Standards and Certification Board, "CoreTrustSeal Trustworthy Data Repositories Requirements 2026–2028," Zenodo, 2026. [Online]. Available: <https://www.coretrustseal.org/why-certification/requirements/>
- [3] S. R. Carroll, I. Garba, O. L. Figueroa-Rodríguez, J. Holbrook, and R. Lovett, "The CARE Principles for Indigenous Data Governance," *Data Science Journal*, vol. 19, no. 1, p. 43, 2020, doi: [10.5334/dsj-2020-043](https://doi.org/10.5334/dsj-2020-043).
- [4] S. R. Carroll, E. Herczog, M. Hudson, K. Russell, and S. Stall, "Operationalizing the CARE and FAIR Principles for Indigenous data futures," *Scientific Data*, vol. 8, no. 1, p. 108, 2021, doi: [10.1038/s41597-021-00892-0](https://doi.org/10.1038/s41597-021-00892-0).
- [5] Y. Shin, S. Shin, J. Um, and D. Seo, "Development of a National Research Data Platform for Sharing and Utilizing Research Data," *Journal of Information Science Theory and Practice*, vol. 10, no. Supplement, pp. 25–38, 2022, doi: [10.1633/JISTaP2022.10.S.3](https://doi.org/10.1633/JISTaP2022.10.S.3).
- [6] M. Fenner, D. Lowenberg, and M. Jones, "Code of Practice for Research Data Usage Metrics," *PeerJ Preprints*, vol. 7, p. e26505v1, 2019, doi: [10.7287/peerj.preprints.26505v1](https://doi.org/10.7287/peerj.preprints.26505v1).
- [7] H. Cousijn, R. Braukmann, and M. Connolly, "Connected Research: The Potential of the PID Graph," *Patterns*, vol. 2, no. 1, p. 100180, 2021, doi: [10.1016/j.patter.2020.100180](https://doi.org/10.1016/j.patter.2020.100180).