

Personal-Name Metadata Accuracy in Indian Institutional Repositories: An Analysis of DC.creator and DC.contributor

Husna Jabeen¹; Dr. N. S. Harinarayana²

Research candidate, Department of Studies in Library and Information Science, University of Mysore, Mysore¹; Professor and chairman, Department of Studies in Library and Information Science, University of Mysore, Mysore, Karnataka, India²

husnalisc@gmail.com; harinarayana@lisc.uni-mysore.ac.in

ABSTRACT

Personal-name metadata is important for author identification, resource discovery, and retrieval in institutional repositories. This study examines the accuracy of personal-name metadata in the DC.creator and DC.contributor elements of 15 selected Indian institutional repositories representing academic institutions and research institutes. Considering name formatting, punctuation, spacing, capitalization, a total of 14,410 Dublin Core metadata records constituted the final analytical dataset. The accuracy assessment focused on English-language personal names, and related deviations. Of the 14,028 DC.creator records assessed, 9,055 (64.55%) were correctly formatted, while 4,973 (35.45%) showed formatting deviations. For DC.contributor, 7,134 of the 11,411 records assessed (62.52%) were correctly formatted, while missing punctuation was the major error, affecting 3,878 records (33.98%). The findings also revealed variation across repositories and institutional categories, indicating a need for standardized name-entry practices, authority control, and improved metadata-quality mechanisms.

KEYWORDS: Personal-name metadata; Metadata accuracy; Institutional repositories; Dublin Core; DC.creator; DC.contributor; Metadata quality; India; Authority control

1. INTRODUCTION

Institutional repositories have become an important component of scholarly communication by collecting, preserving, managing, and providing access to the intellectual output of academic and research institutions. Universities, research institutes, and other scholarly organizations use institutional repositories to disseminate theses, dissertations, journal articles, conference papers, reports, datasets, and other forms of digital scholarly content. The effectiveness of these repositories depends not only on the availability of digital resources but also on the quality of the metadata used to describe them, as inaccurate or inconsistent metadata can adversely affect resource discovery and retrieval (Chandruppa&Harinarayana, 2016;Park &Tosaka, 2010).

Among descriptive metadata elements, personal names are particularly important because they establish intellectual responsibility for a resource and serve as important access points for retrieval. In Dublin Core-based institutional repositories, the **DC.creator** element identifies the person or entity primarily responsible for creating a resource, while the **DC.contributor** element records persons or entities contributing to it (**Dublin Core Metadata Initiative (DCMI, 2020)**). Personal names, however, may be represented in different forms. Variations in name order, punctuation, spacing, capitalization, and treatment of initials can result in multiple representations of the same individual and may affect consistent author identification and retrieval (**Salo, 2009**). Similarly (Chandrappa and Harinarayana, 2016) identified punctuation, spacing, capitalization, and personal-name entry errors in metadata records, demonstrating the importance of accurate and standardized name representation.

Standardized personal-name representation can support author searching, record linking, authority control, and improved metadata quality (**Eichenlaub, 2018; Goslen, 2011; Morgan Salo, 2009;**). The present study therefore investigates the accuracy of personal-name metadata in selected Indian institutional repositories, with particular reference to the **DC.creator** and **DC.contributor** elements. By examining correctly formatted names and deviations involving name structure, initials, punctuation, spacing, and capitalization, the study seeks to identify major problems affecting personal-name metadata and suggest measures for improving metadata quality.

2. NEED FOR THE STUDY

Names of creators and contributors are among the principal access points through which users search for scholarly resources. When the same individual is represented differently across records, retrieval may become fragmented. For example, variations in punctuation, spacing, initials, capitalization, or name order may cause records belonging to the same author to appear under different forms (**Chandrappa&Harinarayana, 2016; Salo, 2009**).

This problem is particularly relevant to institutional repositories because metadata may be created manually, imported from other systems, migrated from older databases, or harvested from heterogeneous sources. Repository software can provide metadata fields, but the existence of a field does not by itself ensure that the value entered into that field is accurate or standardized (**Park &Tosaka, 2010**).

Assessment of personal-name metadata is important for identifying the extent to which institutional repositories follow standardized practices in recording creator and contributor names. Such assessment can identify variations and errors in name representation and provide a basis for improving authority control, metadata-entry guidelines, validation procedures, and overall metadata quality in institutional repositories (**Park &Tosaka, 2010; Salo, 2009; Goslen, 2011**).

3. REVIEW OF LITERATURE

Metadata quality is an important factor in the effective organization, discovery, and retrieval of digital resources. Chandrappa and Harinarayana (2016) investigated the accuracy of metadata records in Mysore University Library by comparing them with corresponding Library of Congress records. The study classified metadata errors into major and minor categories, including omitted or incorrect entries, misspellings, punctuation errors, inadequate spacing, and incorrect capitalization. The findings revealed a considerably higher error rate in Mysore University

Library records than in the Library of Congress records. Importantly, the study also identified substantial errors in personal-name access points, demonstrating the need for greater accuracy and consistency in name representation. However, the study focused on library catalogue records, whereas the present study examines personal-name metadata specifically in the DC.creator and DC.contributor elements of institutional repositories.

Salo (2009) examined **name authority control in institutional repositories** and highlighted the problems caused by variations in author names. The study observed that inconsistent forms of personal names can create difficulties in identifying, retrieving, and aggregating works associated with the same individual. It emphasized the need for improved name authority control and disambiguation mechanisms in institutional repositories. This study is directly relevant to the present research, as variations in the representation of creator and contributor names can affect the accuracy and consistency of repository metadata.

Park and Tosaka (2010) investigated **metadata quality control in digital repositories and collections**, focusing on quality criteria and mechanisms used to maintain metadata quality. Their study identified accuracy and consistency among the important criteria for assessing metadata and emphasized the role of metadata guidelines, staff training, manual review, and quality-control mechanisms. Their findings provide a broader framework for evaluating the accuracy of personal-name metadata in DC.creator and DC.contributor elements.

Goslen (2011) investigated **name authority control and its relationship with metadata quality in institutional repositories**. The study found that name authority control was regarded as an important component of metadata quality and could contribute to improved collocation and retrieval of resources associated with individual authors. The study demonstrates the close relationship between standardized personal-name representation and the overall quality of institutional repository metadata.

Morgan and Eichenlaub (2018) examined **author identifiers and name authority control in two institutional repositories** by comparing author representation across systems such as LCNAF, ORCID, VIAF, ISNI, and Scopus. The study demonstrated variations in author coverage across different authority and identifier systems and suggested that institutional repositories could improve personal-name metadata by linking local author information with external authority sources and persistent identifiers.

Downey (2019) assessed author identifiers in the context of **name authority control in an institutional repository** and examined the potential use of LCNAF, VIAF, and ORCID. The study highlighted the usefulness of external authority sources and persistent identifiers for improving author identification and reducing ambiguity caused by variant forms of personal names. The findings support the adoption of standardized authority-control mechanisms for creator and contributor metadata.

The reviewed studies demonstrate that accurate and standardized representation of personal names is important for metadata quality, author identification, resource discovery, and retrieval in institutional repositories. Previous research has primarily concentrated on **name authority control, author disambiguation, authority workflows, and persistent identifiers**. However, comparatively limited attention has been given to the **quantitative**

assessment of formatting accuracy and specific error patterns in DC.creator and DC.contributor metadata across Indian institutional repositories. The present study addresses this gap by examining English-language personal-name metadata in selected Indian institutional repositories.

4. OBJECTIVES OF THE STUDY

The study aims to:

1. To examine the representation of personal names in the DC.creator and DC.contributor metadata elements of selected Indian institutional repositories.
2. To assess the accuracy of personal-name formatting in the DC.creator element.
3. To identify the major formatting errors and deviations in the DC.contributor element.
4. To compare personal-name metadata practices between academic institutions and research institutes
5. To identify the repository-level variations in personal-name metadata accuracy.

5. METHODOLOGY

The study adopted a descriptive and evaluative research design to examine personal-name metadata in selected Indian institutional repositories. Fifteen repositories representing academic institutions and research institutes were identified through OpenDOAR and ROAR and selected based on the availability of metadata harvesting through the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH). Metadata was harvested using MarcEdit and organized for analysis.

A stratified random sampling approach was adopted to ensure representation from each selected repository. Additional sampling was undertaken for larger repositories, along with manual audit samples for detailed metadata assessment. The final analytical dataset comprised 14,410 metadata records from 15 institutional repositories.

The present article focuses specifically on English-language personal-name metadata recorded in the DC.creator and DC.contributor elements. The assessment criteria were informed by previous studies on metadata accuracy and personal-name representation (Chandruppa&Harinarayana, 2016; Salo, 2009). Accuracy was assessed in terms of name structure and order, punctuation, spacing, capitalization, and treatment of initials. Records conforming to the predefined operational criteria were classified as correctly formatted, while deviations were categorized according to the type of formatting problem observed. Frequencies and percentages were used for the analysis, together with repository-wise and institutional category-wise comparisons and representative examples of formatting patterns and errors.

6. ANALYSIS AND INTERPRETATION

The collected metadata was organized and analysed using frequencies and percentages. Repository-wise analysis was used to identify differences among individual repositories, while institutional category-wise analysis was used to compare academic institutions and research institutes.

Personal-Name Metadata Accuracy in Indian Institutional Repositories: An Analysis of DC.creator and DC.contributor

Table 1 Overall Accuracy of Personal-Name Formatting in the DC.creator Element

Accuracy status	Number of records	Percentage (%)
Correctly formatted	9,055	64.55
Incorrectly formatted	4,973	35.45
Total	14,028	100

The results show that 9,055 records (64.55%) contained correctly formatted creator names, while 4,973 records (35.45%) exhibited formatting deviations. Thus, approximately two-thirds of the examined English-language creator metadata followed the adopted formatting criteria. Correctly formatted names generally followed a standardized surname-first pattern with appropriate initials, punctuation, and spacing. For example, “**Kapadia, C. R.**” and “**Gandhi, Ashutosh S.**” represent correctly formatted creator names in which the surname and personal name or initials are clearly structured and punctuation is consistently applied.

In contrast, 4,973 records (35.45%) contained formatting deviations, indicating that more than one-third of the creator metadata did not conform to the adopted format. Several types of deviations were observed. For instance, “**A.K., Hemanth Kumar**” and “**G., Bhagavannarayana**” illustrate inconsistent representation of initials and personal names, particularly in the placement of initials and punctuation. Spacing-related deviations were evident in examples such as “**Aadhithyan , A K**” and “**Saraswat, V.K.**”, where spacing around punctuation or between initials was inconsistent with the adopted formatting pattern. Other deviations involved name order and capitalization. Examples such as “**Fey, G.T.K.**” and “**MUTHIAH, SARAVANAN**” demonstrate variation in the treatment of initials and capitalization, with the latter showing the personal name entirely in capital letters.

The findings show that DC.creator metadata contained different types of formatting deviations, particularly in punctuation, spacing, capitalization, initials, and name representation. These variations may affect the consistent representation and retrieval of author names in institutional repositories.

Table 2 Institutional Category-Wise Accuracy of English-Language Personal-Name Formatting in the DC.creator Element

Institutional category	Correct format, n (%)	Incorrect format, n (%)	Total, N (%)
Academic institutionss	1,018 (68.83)	461 (31.17)	1,479 (100.00)
Research institutes	8,037 (64.04)	4,512 (35.96)	12,549 (100.00)
Total	9,055 (64.55)	4,973 (35.45)	14,028 (100.00)

The institutional category-wise analysis shows that academic institutions had a slightly higher proportion of correctly formatted DC.creator personal names than research institutes. Among the 1,479 English-language creator records from academic institutions, 1,018 (68.83%) were correctly formatted and 461 (31.17%) were incorrectly formatted. Among the 12,549 records from research institutes, 8,037 (64.04%) were correctly formatted and 4,512 (35.96%) were incorrectly formatted. Overall, 9,055 (64.55%) of the 14,028 records were correctly formatted. The

findings indicate that correct formatting predominated in both institutional categories, although academic institutions showed a modestly higher proportion of correctly formatted creator names.

Table 3 Repository-Wise Accuracy of English-Language Personal-Name Formatting in the DC.creator Element

Repository	Correctly formatted, n (%)	Incorrectly formatted, n (%)	Total, N (%)
CECRI_TAMILNADU	99 (42.49%)	134 (57.51%)	233 (1.66%)
CFTRI_MYSORE	126 (58.60%)	89 (41.40%)	215 (1.53%)
CMFRI_KOCHI	356 (34.30%)	682 (65.70%)	1,038 (7.40%)
IAS_BANGALORE	5,142 (99.96%)	2 (0.04%)	5,144 (36.67%)
ICRISAT_HYDERABAD	0 (0.00%)	685 (100.00%)	685 (4.88%)
IISC_BANGALORE	109 (26.52%)	302 (73.48%)	411 (2.93%)
MDRF_CHENNAI	5 (5.00%)	95 (95.00%)	100 (0.71%)
NEERI_NAGPUR	16 (11.43%)	124 (88.57%)	140 (1.00%)
NEIST_ASSAM	2 (1.71%)	115 (98.29%)	117 (0.83%)
NIRT_CHENNAI	44 (22.45%)	152 (77.55%)	196 (1.40%)
NISCAIR_NEWDELHI	1,938 (50.00%)	1,938 (50.00%)	3,876 (27.63%)
NITR_ROURKELA	348 (72.35%)	133 (27.65%)	481 (3.43%)
NML_JAMSHEDPUR	102 (19.84%)	412 (80.16%)	514 (3.66%)
NPL_NEWDELHI	207 (71.13%)	84 (28.87%)	291 (2.07%)
UOM_MYSORE	561 (95.57%)	26 (4.43%)	587 (4.18%)
Total	9,055 (64.55%)	4,973 (35.45%)	14,028 (100.00%)

The repository-wise analysis reveals considerable variation in the accuracy of English-language personal-name formatting in the DC.creator element. Of the 14,028 records assessed, 9,055 (64.55%) were correctly formatted, while 4,973 (35.45%) showed formatting deviations. IAS Bangalore recorded the highest proportion of correctly formatted creator names, with 5,142 of 5,144 records (99.96%) correctly formatted. UOM Mysore also showed a very high level of correct formatting (95.57%), followed by NITR Rourkela (72.35%) and NPL New Delhi (71.13%).

In contrast, ICRISAT Hyderabad recorded no correctly formatted creator names, with all 685 records (100.00%) classified as incorrectly formatted. High proportions of formatting deviations were also observed in NEIST Assam (98.29%), MDRF Chennai (95.00%), NEERI Nagpur (88.57%), NML Jamshedpur (80.16%), NIRT Chennai (77.55%), IISc Bangalore (73.48%), and CMFRI Kochi (65.70%). NISCAIR New Delhi showed an equal distribution, with 50.00% correctly formatted and 50.00% incorrectly formatted records. Overall, the findings indicate substantial repository-level variation in DC.creator personal-name formatting and suggest that standardized name-entry practices were not uniformly followed across the selected institutional repositories.

Table 4 Overall Accuracy and Error Categories in DC.contributor Metadata

Accuracy category	Frequency, n	Percentage (%)
Correct format	7,134	62.52
Incorrect format	80	0.7
All capitals	99	0.87
Missing punctuation only	3,878	33.98
Missing space only	203	1.78
Missing punctuation and space	17	0.15
Total	11,411	100

Note. Accuracy was assessed only for the 11,411 records containing English-language DC.contributor metadata. A further 2,961 records contained no contributor metadata, while 38 multilingual contributor records were excluded from the English-language formatting assessment.

The table presents the accuracy of English-language contributor metadata recorded in the DC.contributor element. Of the 11,411 English-language contributor records assessed, 7,134 records (62.52%) were correctly formatted, while the remaining records showed different types of formatting deviations.

Missing punctuation was the most frequently observed problem, affecting 3,878 records (33.98%). Missing spaces occurred in 203 records (1.78%), while 99 records (0.87%) were represented entirely in capital letters. A further 80 records (0.70%) were classified under other incorrect formatting, and 17 records (0.15%) contained both missing punctuation and spacing problems.

Overall, nearly two-thirds of the English-language contributor records were correctly formatted, while 37.48% showed some form of formatting deviation. Missing punctuation constituted the major accuracy problem. Such variation in contributor-name representation may affect consistent name identification, authority control, searching, retrieval, and metadata interoperability.

A further 2,961 records contained no DC.contributor metadata and were therefore not included in the formatting-accuracy assessment. The 38 multilingual contributor records were also excluded from this English-language analysis.

Table 5 Institutional Category-Wise Accuracy and Error Categories of English-Language Personal-Name Formatting in the DC.contributor Element

Accuracy category	Academic Institute, n	Percentage (%)	Research Institute, n	Percentage (%)	Total, N	Percentage (%)
Correct format	651	98.64	6,483	60.3	7,134	62.52
Incorrect format	3	0.45	77	0.72	80	0.7
All capitals	0	0	99	0.92	99	0.87

Missing punctuation only	6	0.91	3,872	36.02	3,878	33.98
Missing space only	0	0	203	1.89	203	1.78
Missing punctuation and space	0	0	17	0.16	17	0.15
Total	660	100	10,751	100	11,411	100

The institutional category-wise analysis shows a clear difference in DC.contributor personal-name formatting between academic institutions and research institutes. Among the 660 English-language contributor records from academic institutions, 651 (98.64%) were correctly formatted, while only 9 records (1.36%) showed formatting deviations. In contrast, among the 10,751 records from research institutes, 6,483 (60.30%) were correctly formatted, while 4,268 (39.70%) contained one or more formatting deviations. Missing punctuation was the most prominent problem among research-institute records, accounting for 3,872 records (36.02%), whereas only 6 records (0.91%) from academic institutions showed this deviation. Other problems, including incorrect formatting, all-capital representation, missing spaces, and combined punctuation-and-spacing errors, occurred mainly in research-institute records. Overall, the findings indicate that correctly formatted contributor names were proportionally more common in academic institutions, while formatting deviations were considerably more prevalent among the research institutes examined.

Table 6 Repository-Wise Accuracy and Error Patterns of English-Language Personal-Name Formatting in the DC.contributor Element

Institutional Repository	Correct format, n (%)	Incorrect format, n (%)	All capitals, n (%)	Missing punctuation only, n (%)	Missing space only, n (%)	Missing punctuation and space, n (%)	Total assessed, N (%)
CECRI_TAMILNADU	74 (33.18)	0 (0.00)	0 (0.00)	0 (0.00)	149 (66.82)	0 (0.00)	223 (1.95)
CFTRI_MYSORE	167 (95.98)	0 (0.00)	0 (0.00)	0 (0.00)	7 (4.02)	0 (0.00)	174 (1.52)
CMFRI_KOCHI	10 (1.74)	0 (0.00)	0 (0.00)	564 (98.26)	0 (0.00)	0 (0.00)	574 (5.03)
IAS_BANGALORE	4,543 (99.93)	3 (0.07)	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	4,546 (39.84)
ICRISAT_HYDERABAD	0 (0.00)	3 (0.50)	0 (0.00)	593 (99.50)	0 (0.00)	0 (0.00)	596 (5.22)
IISC_BANGALORE	-	-	-	-	-	-	-
MDRF_CHENNAI	1 (1.10)	0 (0.00)	0 (0.00)	90 (98.90)	0 (0.00)	0 (0.00)	91 (0.80)
NEERI_NAGPUR	52 (38.24)	1 (0.74)	0 (0.00)	79 (58.09)	4 (2.94)	0 (0.00)	136 (1.19)

Personal-Name Metadata Accuracy in Indian Institutional Repositories: An Analysis of DC.creator and DC.contributor

NEIST_ASSAM	2 (1.72)	0 (0.00)	94 (81.03)	20 (17.24)	0 (0.00)	0 (0.00)	116 (1.02)
NIRT_CHENNAI	48 (25.95)	3 (1.62)	0 (0.00)	133 (71.89)	1 (0.54)	0 (0.00)	185 (1.62)
NISCAIR_NEWDELHI	1,315 (38.94)	27 (0.80)	2 (0.06)	1,975 (58.48)	41 (1.21)	17 (0.50)	3,377 (29.59)
NITR_ROURKELA	139 (93.92)	3 (2.03)	0 (0.00)	6 (4.05)	0 (0.00)	0 (0.00)	148 (1.30)
NML_JAMSHEDPUR	27 (6.05)	1 (0.22)	1 (0.22)	417 (93.50)	0 (0.00)	0 (0.00)	446 (3.91)
NPL_NEWDELHI	244 (85.02)	39 (13.59)	2 (0.70)	1 (0.35)	1 (0.35)	0 (0.00)	287 (2.52)
UOM_MYSORE	512 (100.00)	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	512 (4.49)
Total	7,134 (62.52)	80 (0.70)	99 (0.87)	3,878 (33.98)	203 (1.78)	17 (0.15)	11,411 (100.00)

Note: Percentages for individual repositories were calculated using the total number of English-language DC.contributor records assessed within each repository. Percentages in the Total row were calculated using the overall assessed records (N = 11,411). DC.contributor was not applicable to IISc Bangalore because the repository primarily contained thesis records in which contributor metadata was not used for the present assessment.

The repository-wise analysis of DC.contributor metadata reveals considerable variation in personal-name formatting accuracy. Of the 11,411 English-language contributor records assessed, 7,134 (62.52%) were correctly formatted. UOM Mysore recorded 100.00% correctly formatted records, followed by IAS Bangalore (99.93%), CFTRI Mysore (95.98%), and NITR Rourkela (93.92%). A representative correctly formatted example is “Anbukulandainathan, M.; Arulraj, I.; Chandrasekaran, R.; Pattabiraman, R.”

Missing punctuation was the most prominent error, affecting 3,878 records (33.98%) overall and occurring at particularly high levels in ICRISAT Hyderabad, MDRF Chennai, CMFRI Kochi, and NML Jamshedpur. For example, “Nair, K R Manmadhan; Kripa, V” illustrates missing punctuation between initials. CECRI Tamil Nadu showed a high proportion of missing-space errors, while all-capital formatting was prominent in NEIST Assam. Overall, the findings indicate substantial repository-level variation and show that punctuation-related deviations constituted the major accuracy problem in DC.contributor personal-name metadata.

7. MAJOR FINDINGS

The study revealed considerable variation in the accuracy of English-language personal-name metadata across the selected institutional repositories. In the DC.creator element, 9,055 of 14,028 records (64.55%) were correctly formatted, while 4,973 (35.45%) showed formatting deviations. Academic institutions recorded a slightly higher

proportion of correctly formatted creator names (68.83%) than research institutes (64.04%). Repository-wise analysis also showed substantial variation, with IAS Bangalore recording the highest level of correct formatting.

In the DC.contributor element, 7,134 of 11,411 records (62.52%) were correctly formatted, while missing punctuation was the most prominent error, affecting 3,878 records (33.98%). The institutional category-wise analysis showed that 98.64% of assessed academic-institution contributor records were correctly formatted, compared with 60.30% of research-institute records.

Overall, the findings indicate considerable variation in personal-name formatting across repositories and institutional categories, highlighting the need for greater standardization in DC.creator and DC.contributor metadata.

DISCUSSION

The findings confirm that personal-name metadata is an important component of metadata quality in institutional repositories. Creator and contributor names serve as important access points for identifying scholarly responsibility and retrieving works associated with individuals. Variations in name order, initials, punctuation, spacing, and capitalization may therefore affect the uniform representation of personal names across repository records.

The DC.creator analysis shows that 64.55% of the English-language creator records were correctly formatted, while 35.45% contained formatting deviations. The institutional category-wise comparison further shows that academic institutions had a slightly higher proportion of correctly formatted creator names (68.83%) than research institutes (64.04%). Incorrect formatting was correspondingly higher among research institutes (35.96%) than academic institutions (31.17%). These results indicate a modest difference between the two institutional categories, while the repository-wise findings demonstrate much greater variation in creator-name formatting practices.

The DC.contributor findings show that punctuation was the most prominent accuracy problem. Of the 11,411 English-language contributor records assessed, 3,878 (33.98%) contained missing punctuation. For example, “Nair, K R Manmadhan; Kripa, V” illustrates the absence of expected punctuation between initials. Other problems, including missing spaces, all-capital formatting, incorrect name structure, and combined punctuation-and-spacing errors, occurred less frequently.

The findings correspond with earlier studies emphasizing the importance of name authority control, standardized personal-name representation, and persistent identifiers for improving author identification and retrieval. They also demonstrate an important distinction between completeness and accuracy: the presence of a creator or contributor value does not necessarily mean that the name is accurately or consistently represented. Records without contributor data were therefore excluded from the revised accuracy assessment because absence relates to completeness rather than accuracy.

Overall, the results demonstrate that personal-name formatting was not uniform across the selected repositories or institutional categories. The observed differences may reflect variations in metadata-entry and quality-control practices; however, the present study did not directly investigate the specific causes of these differences.

RECOMMENDATIONS

Based on the findings, the following measures are recommended for improving personal-name metadata in Indian institutional repositories:

1. **Standardize name-entry practices.** Repositories should establish clear rules for surname and forename order, initials, punctuation, spacing, and capitalization in DC.creator and DC.contributor fields.
2. **Introduce authority control.** Standardized forms of creator and contributor names should be maintained, with variant forms linked where possible.
3. **Use automated validation mechanisms.** Repository systems should incorporate checks for missing punctuation, missing spaces, unusual capitalization, and other predictable formatting deviations.
4. **Conduct regular metadata audits and staff training.** Periodic review of creator and contributor fields, combined with training in standardized name representation, can help identify recurring problems and improve metadata quality.
5. **Document repository-level metadata policies.** Institutions should maintain clear metadata guidelines or application profiles defining the expected use and formatting of DC.creator and DC.contributor elements.

CONCLUSION

The study demonstrates considerable variation in the accuracy of personal-name metadata across the selected Indian institutional repositories. In the DC.creator element, 9,055 of the 14,028 English-language records assessed (64.55%) were correctly formatted, while 4,973 records (35.45%) contained formatting deviations. Academic institutions showed a slightly higher proportion of correctly formatted creator names than research institutes, while notable differences were also observed across individual repositories.

A similar pattern was evident in the DC.contributor element. Of the 11,411 English-language contributor records assessed, 7,134 (62.52%) were correctly formatted, whereas 4,277 records (37.48%) showed one or more formatting deviations. Missing punctuation was the most frequently observed problem, affecting 3,878 records (33.98%). Other deviations, including incorrect name structure, all-capital representation, missing spaces, and combined punctuation-and-spacing errors, occurred comparatively less frequently.

Overall, the findings indicate that the presence of creator and contributor metadata does not necessarily ensure accurate and standardized personal-name representation. Variations in name structure, punctuation, spacing, capitalization, and treatment of initials may reduce consistency in author identification and retrieval. Institutional repositories should therefore adopt standardized name-entry guidelines, strengthen authority-control practices, introduce appropriate validation mechanisms, provide staff training, and conduct periodic metadata audits. These measures can improve the reliability and uniformity of DC.creator and DC.contributor metadata and contribute to more effective author identification, resource discovery, and metadata interoperability in institutional repositories.

REFERENCES

- [1] Chandrappa, & Harinarayana, N. S. (2016). Investigation into diligence in metadata records of Mysore University Library. *DESIDOC Journal of Library & Information Technology*, 36(6), 431–436.
<https://doi.org/10.14429/djlit.36.6.10608>

- [2] Downey, M. (2019). Assessing author identifiers: Preparing for a linked data approach to name authority control in an institutional repository context. *Journal of Library Metadata*, 19(1–2), 117–136.
<https://doi.org/10.1080/19386389.2019.1590936>
- [3] DCMI Usage Board. (2020). *DCMI metadata terms*. Dublin Core Metadata Initiative.
<https://www.dublincore.org/specifications/dublin-core/dcmi-terms/>
- [4] Goslen, A. S. (2011). *Name authority control in institutional repositories and its relationships to metadata quality* [Master's paper, University of North Carolina at Chapel Hill]. Carolina Digital Repository.
<https://doi.org/10.17615/cxer-ny36>
- [5] Morgan, M., & Eichenlaub, N. (2018). Author identifier analysis: Name authority control in two institutional repositories. *Proceedings of the International Conference on Dublin Core and Metadata Applications, 2018*, 113–116. <https://doi.org/10.23106/dcmi.952139036>
- [6] Park, J.-R., & Tosaka, Y. (2010). Metadata quality control in digital repositories and collections: Criteria, semantics, and mechanisms. *Cataloging & Classification Quarterly*, 48(8), 696–715.
<https://doi.org/10.1080/01639374.2010.508711>
- [7] Salo, D. (2009). Name authority control in institutional repositories. *Cataloging & Classification Quarterly*, 47(3–4), 249–261. <https://doi.org/10.1080/01639370902737232>
-