Mini Review

FAIR Data and the Future of Scholarly Communication

Anup Basnet Chetry[1]*, Sushma Thapa[2], Dharmaraj Thapa[1] and Youbaraj Poudel[3]

[1]Graduate School of Science and Technology, MidWest University, Surkhet, Nepal

[2]Department of English, Tribhuwan University, Kathmandu, Nepal

[3]Department of Computer Science, Kathmandu University, Nepal

https://www.clinsurggroup.us

Check for updates

Abstract

In the digital age, the way research data is managed, shared, and reused has become a critical factor in shaping the quality, transparency, and impact of scholarly communication. The FAIR data principles—Findability, Accessibility, Interoperability, and Reusability—provide a foundational framework for ensuring that scientific data can be effectively located, understood, and applied by both humans and machines. This mini-review explores how FAIR data enhances global collaboration, accelerates scientific discovery through AI and automation, and fosters open, inclusive research ecosystems. While significant barriers remain—including technical, cultural, and infrastructural challenges—widespread adoption of FAIR is essential for building a future-ready, responsible, and equitable scientific enterprise.

Introduction

In an era where research is increasingly data-driven, the way scientific information is managed, shared, and reused is undergoing a profound transformation. Traditional academic publishing has long prioritized the research article as the primary unit of scientific output. However, as the volume and complexity of data grow exponentially, the limitations of this model have become evident. Most published papers include only summaries or snapshots of data, with raw datasets often remaining inaccessible, poorly documented, or incompatible with other systems. This lack of data transparency contributes to issues of reproducibility, inefficiency, and missed opportunities for interdisciplinary collaboration. Recognizing these challenges, the scientific community has begun to shift its focus toward open science and responsible data management. Central to this shift are the FAIR Guiding Principles—Findability, Accessibility, Interoperability, and Reusability—first articulated by Wilkinson, et al. in 2016 [1]. These principles provide a framework for structuring and sharing research data in a way that enables both human and machine discovery. Unlike open data, which merely implies availability, FAIR emphasizes the quality, usability, and long-term stewardship of data. It is not about making all data public, but rather ensuring that it can be reliably located, understood, and reused under appropriate conditions. The adoption of FAIR principles represents a critical evolution in scholarly communication. It empowers researchers to make their work more transparent, facilitates collaboration across disciplines and borders, and supports the reproducibility and verification of results—cornerstones of scientific progress. Moreover, as artificial intelligence, digital laboratories, and automated analytics become increasingly embedded in research practice, the need for data that is both accessible and machine-actionable has never been greater. In this context, FAIR is not just a best practice—it is a scientific imperative [2,3].

Unlike open data, which merely implies availability, FAIR emphasizes the quality, usability, and long-term stewardship of data. Although FAIR data and open data are often discussed together, they represent distinct yet complementary concepts within modern data governance. Open data primarily

001

emphasizes unrestricted access and availability, whereas FAIR data focuses on ensuring that data are Findable, Accessible, Interoperable, and Reusable, with an emphasis on quality, machine readability, and long-term stewardship. In practical terms, data can be open without being FAIR. For instance, research datasets published as static files (e.g., PDFs or unstructured spreadsheets) on institutional websites may be freely accessible but lack standardized metadata, persistent identifiers, or interoperable formats. As a result, such datasets are difficult to discover through automated systems, cannot be readily integrated with other data sources, and often require substantial manual preprocessing, limiting their reuse. Conversely, FAIR data are not necessarily fully open. In biomedical research, sensitive patient or genomic data are frequently subject to ethical, legal, or privacy constraints. FAIR principles allow these datasets to remain controlled-access while still being richly described with standardized metadata, persistent identifiers, and interoperable vocabularies. This enables authorized users and machine agents to locate, access, and reuse the data responsibly, without compromising confidentiality. A concrete example can be observed in climate science and environmental monitoring. Open climate datasets released without harmonized metadata or common ontologies may be accessible but difficult to integrate across agencies or regions. In contrast, FAIR-aligned climate datasets—such as those coordinated through international frameworks—enable seamless cross-platform integration, facilitating multi-scale modeling and long-term trend analysis. This distinction highlights that FAIRness enhances usability and sustainability, rather than merely increasing data visibility. By emphasizing structured metadata, interoperability standards, and reuse conditions, FAIR data principles extend beyond the scope of open data. Together, these frameworks support transparent, efficient, and responsible data sharing, particularly in data-intensive and interdisciplinary research domains (Table 1).

The need for FAIR in scientific communication

Traditional research outputs have centered on the published article, often treating data as supplementary or even disposable. This approach is incompatible with the reproducibility and transparency demands of modern science. Data that is not properly curated, documented, or accessible becomes a barrier to knowledge reuse. FAIR data practices, by contrast, promote longevity and integrity across the research lifecycle. Moreover, funding agencies, journals, and institutions are beginning to require data management plans and enforce open data mandates, making adherence to FAIR not only good practice but also a compliance issue. Researchers who engage with FAIR principles early are better positioned to meet these evolving standards. Scientific progress depends not only on generating new data but also on building upon existing knowledge. However, without proper data management, even valuable datasets can become functionally invisible—lost in disconnected repositories, poorly labeled spreadsheets, or behind institutional firewalls. The FAIR data principles were introduced to address these very issues, ensuring that research outputs are not only preserved but optimized for reuse, interoperability, and discovery. The "Findable" aspect of FAIR ensures that data and associated metadata can be discovered by both humans and machines. This typically involves assigning a persistent and globally unique identifier, such as a DOI, and ensuring metadata is indexed in searchable resources. In practice, this means researchers can trace datasets back to their origin, understand their structure, and assess their relevance to new investigations. Accessibility goes beyond simply making data publicly available. It involves ensuring that users—whether people or software—can retrieve data using standardized protocols. Even when data is behind restricted access due to privacy or intellectual property constraints, FAIR principles encourage clarity in terms of access conditions. Interoperability is critical for integration and comparison across studies. Inconsistent formats, ambiguous variable names, and a lack of standardized vocabularies have long hindered the synthesis of scientific knowledge. FAIR encourages the use of formal, shared ontologies and structured formats to make data compatible across platforms and disciplines. Finally, Reusability depends on rich metadata, detailed provenance, and clearly defined usage licenses. This allows future researchers to understand the context in which the data was collected,

Table1: Comparison of FAIR Data and Open Data Across Key Dimensions of Scholarly Communication.

Dimension	Open Data	FAIR Data
Primary focus	Unrestricted access to data	Data usability through findability, accessibility, interoperability, and reusability
Access model	Typically fully open to all users	Can be open or controlled-access, depending on ethical and legal constraints
Metadata quality	Often minimal or inconsistent	Rich, standardized, and machine-readable metadata
Findability	Dependent on manual search or repository visibility	Enabled through persistent identifiers (e.g., DOIs) and indexed metadata
Interoperability	Limited; formats and terminologies may vary	High: uses shared standards, ontologies, and vocabularies
Reusability	Often unclear due to missing licenses or context	Explicit reuse conditions, provenance, and versioning
Machine readability	Frequently human-readable only	Designed for machine-to-machine interaction
Support for AI and ML	Requires extensive preprocessing	Directly usable for automated analysis and model training
Data stewardship	Short-term sharing emphasis	Long-term preservation and sustainability
Ethical and governance handling	May overlook privacy or sensitivity issues	Integrates governance, access control, and ethical safeguards
Implementation complexity	Relatively low	Higher; requires infrastructure, expertise, and policy alignment
Representative example	Public datasets shared as PDFs or spreadsheets	GISAID genomic data, FAIR-aligned climate repositories

assess its quality, and confidently apply it to new questions. Reusable data fuels meta-analyses, supports replication, and drives innovation by reducing duplication of effort. In sum, FAIR data transforms research outputs from isolated artifacts into active components of an evolving knowledge ecosystem. It enables cumulative science, fosters interdisciplinary research, and aligns with global trends toward open science and digital transformation.

Rather than treating the distinction between FAIR data and open data as a purely conceptual issue, this manuscript integrates their differences across key stages of the research and scholarly communication lifecycle. This approach highlights how FAIR principles extend the utility of open data practices in real-world applications. In scholarly publishing, open data policies typically require authors to make datasets publicly available as a condition of publication. However, without FAIR-aligned metadata standards, persistent identifiers, and machine-readable formats, such datasets often remain underutilized. FAIR data infrastructures, by contrast, enable automated discovery and citation of datasets, improving reproducibility and increasing the long-term scholarly value of published research outputs.

FAIR as a catalyst for collaboration

The COVID-19 pandemic underscored the importance of data sharing across borders and disciplines. Public health responses, epidemiological modeling, and vaccine development all relied on rapid, open access to data repositories. FAIR principles facilitate such interdisciplinary collaboration by ensuring that datasets are comprehensible and usable beyond their original context. Tools like the GO FAIR initiative and the European Open Science Cloud (EOSC) exemplify the institutional momentum building around FAIR infrastructures. One of the most transformative impacts of FAIR data is its ability to foster meaningful, scalable collaboration across disciplines, institutions, and national boundaries. By ensuring that data is properly structured, documented, and accessible, FAIR principles unlock opportunities for co-creation, verification, and innovation that were previously hindered by data silos and incompatible formats. A striking example of this came during the COVID-19 pandemic, when the urgency for a rapid and coordinated scientific response demanded unprecedented levels of data sharing. Researchers around the globe relied on FAIR-aligned repositories such as GISAID for genomic data and WHO COVID-19 dashboards for clinical and epidemiological tracking. These platforms allowed scientists, policymakers, and health workers to make real-time decisions based on up-to-date information—a feat only possible because the data was standardized, well-annotated, and openly accessible under clear usage agreements. Beyond the pandemic, FAIR data is also advancing collaboration in fields such as climate science, where global data harmonization is crucial for modeling future scenarios. Initiatives like the Group on Earth Observations (GEO) and Copernicus integrate satellite, terrestrial, and oceanographic data from multiple countries and make it available in FAIR-compliant formats. These collective efforts enable climate researchers to perform high-resolution

environmental monitoring, predict extreme weather events, and inform international climate policy. In the life sciences, platforms like ELIXIR, The Human Cell Atlas, and BioSharing use FAIR frameworks to facilitate integrative research across genomics, proteomics, and health data. In environmental and biodiversity science, projects such as GBIF (Global Biodiversity Information Facility) curate species distribution data in standardized formats, enabling ecological research and conservation strategies across continents. These examples illustrate that FAIR is not merely a data management slogan—it is a driver of systemic change. It allows researchers in low-resource settings to participate in global science, democratizes access to high-value datasets, and supports reproducibility across culturally and technically diverse research communities. As global challenges—from pandemics to climate change to sustainable development—require collective intelligence and joint action, FAIR data becomes a critical enabler of the science-society interface.

These platforms allowed scientists, policymakers, and health workers to make real-time decisions based on up-to-date information. While conceptual discussions highlight the transformative potential of FAIR data, empirical evidence further demonstrates its measurable impact on research efficiency, collaboration, and decision-making. Several large-scale initiatives provide quantifiable examples of how FAIR-aligned data infrastructures enhance scientific outcomes. During the COVID-19 pandemic, the Global Initiative on Sharing All Influenza Data (GISAID) enabled the rapid dissemination of SARS-CoV-2 genomic data across the global research community. By mid-2022, the platform had facilitated the sharing of millions of viral genome sequences from more than 200 countries, allowing near real-time monitoring of viral evolution. This FAIR-oriented data-sharing framework reduced variant identification timelines from weeks to days, directly supporting accelerated epidemiological modeling, vaccine optimization, and public health decision-making. The high level of standardized metadata and controlled but interoperable access significantly increased data reuse rates, as reflected by the exponential growth in secondary analyses and cross-institutional studies. Similarly, in climate and environmental sciences, the Group on Earth Observations (GEO) and its Global Earth Observation System of Systems (GEOSS) exemplify the quantitative benefits of FAIR data implementation. GEO-aligned datasets have improved cross-agency interoperability, enabling integration of satellite, in situ, and model-based data streams. Studies report measurable reductions in data acquisition redundancy and shortened research-to-policy timelines, as interoperable datasets allow policymakers and researchers to access harmonized information without duplicative data collection efforts. Furthermore, FAIR-enabled metadata standards have contributed to increased dataset discoverability and reuse, strengthening longitudinal climate assessments and risk modeling. Collectively, these examples demonstrate that FAIR data principles do not merely improve theoretical accessibility but yield tangible performance gains, including faster research cycles, enhanced collaboration efficiency, and improved responsiveness to global challenges. Incorporating quantitative indicators into FAIR case studies

underscores the principles' value as an operational framework for modern, data-intensive science.

Challenges to FAIR implementation

Despite its clear benefits, the adoption of FAIR principles is not without obstacles. Key challenges include:

- **Lack of infrastructure** in many low- and middle-income countries to support data repositories and persistent identifiers.

- **Disciplinary differences** in metadata standards and documentation practices.

- **Limited incentives** for researchers to invest time in curating and sharing data, especially when career advancement remains tied to traditional publication metrics.

Addressing these issues requires a systemic shift: journals must prioritize data quality, universities must reward open practices, and funders must invest in digital infrastructure. Despite the growing recognition of the FAIR principles and their benefits, widespread implementation remains uneven. Many research communities, particularly in low- and middle-income countries (LMICs), face structural, technical, and cultural obstacles that hinder full alignment with FAIR standards. Without deliberate strategies to address these gaps, FAIR risks becoming more aspirational than actionable. One major barrier is the lack of adequate digital infrastructure. Maintaining FAIR-compliant data repositories requires reliable internet connectivity, cloud storage systems, and secure servers—resources that are not uniformly distributed across institutions and regions. In many settings, datasets are still stored on personal devices or shared informally, making them vulnerable to loss, duplication, or poor version control. A second challenge lies in disciplinary diversity and fragmentation. Different fields have different data cultures, with varied standards, terminologies, and ontologies. While genomics and astronomy have established strong traditions of open, structured data, many social sciences, humanities, and engineering disciplines are still developing community-wide frameworks for data sharing. This inconsistency makes it difficult to apply FAIR principles uniformly across all domains of research. Human capacity constraints also limit FAIR adoption. Many researchers—especially early-career scientists—receive limited training in data stewardship, metadata generation, and digital preservation. Without proper incentives or recognition for their efforts, they may view data curation as an administrative burden rather than an integral part of research. Moreover, in academic environments where funding, promotions, and citations are closely tied to traditional publications, data sharing may be seen as low-priority or even risky. Legal and ethical considerations present yet another layer of complexity. Sensitive data—such as patient health records, indigenous knowledge, or environmental data with geopolitical implications—cannot always be made fully open. Navigating the tension between accessibility and protection requires clear data governance policies, robust consent frameworks, and secure

access control systems. Lastly, there remains a disconnect between policy and practice. While many journals, funders, and governments have issued open data mandates, enforcement is often weak. FAIR compliance is rarely monitored, and tools to assess the quality of metadata or the reusability of datasets are still evolving. Addressing these barriers will require coordinated action across stakeholders. Funders and institutions must invest in training and infrastructure. Journals should require data availability statements and enforce metadata standards. Governments must promote open science policies that balance openness with privacy and equity. Crucially, researchers must be recognized and rewarded for sharing high-quality, reusable data.

Challenges, risks, and governance considerations of FAIR data

While the FAIR principles offer substantial benefits for data-intensive research and scholarly communication, their implementation is not without challenges. A balanced evaluation requires acknowledging the technical, financial, and ethical complexities associated with transitioning toward FAIR-aligned data infrastructures. One significant challenge is the **cost of** implementation and maintenance. Developing interoperable data repositories, curating high-quality metadata, and ensuring long-term data preservation require sustained financial investment and skilled personnel. For resource-limited institutions and researchers, particularly in low- and middle-income countries, these requirements may pose barriers to FAIR adoption, potentially exacerbating existing inequities in global research participation. Data governance and stewardship complexity represent another critical concern. FAIR data frameworks necessitate clear policies regarding data ownership, access rights, version control, and responsibility for long-term maintenance. In multi-institutional and cross-border collaborations, aligning governance structures and legal frameworks can be challenging, especially when data are subject to differing national regulations or institutional policies. Ethical considerations are particularly salient in domains involving sensitive or human-related data, such as health, genomics, and social sciences. While FAIR principles support controlled access, improper implementation may increase the risk of data misuse, re-identification, or unintended secondary applications. Robust governance mechanisms, including ethical review processes, access committees, and privacy-preserving technologies, are therefore essential to ensure responsible data reuse. Additionally, the push toward FAIR data may increase the administrative burden on researchers, who must allocate time and resources to data documentation, metadata creation, and compliance with repository standards. Without adequate institutional support, incentives, and training, this burden could discourage data sharing rather than promote it. Despite these challenges, many risks associated with FAIR data implementation are manageable through coordinated policy development, capacity building, and international cooperation. Investments in shared infrastructures, standardized guidelines, and researcher training can mitigate barriers while maximizing the long-term benefits of FAIR data for transparent, reproducible, and equitable scholarly communication.

The link to AI and automation

FAIR data not only benefits human collaboration—it is also key to machine learning and AI. Well-structured, machine-readable data enables automated literature reviews, data mining, and even robotic experimentation. For instance, the concept of *chemputation* in chemistry exemplifies how FAIR-aligned protocols allow digital systems to perform reproducible chemical synthesis. As AI becomes more prevalent in research, FAIR will underpin its effectiveness. As artificial intelligence (AI), machine learning (ML), and automation become deeply integrated into research workflows, the importance of structured, machine-actionable data has never been more apparent. FAIR data serves as the essential fuel for these technologies, allowing algorithms to parse, interpret, and learn from vast and diverse datasets with minimal human intervention. Unlike unstructured or poorly annotated data, FAIR-aligned datasets are embedded with metadata, standardized formats, and clear provenance. This structured richness enables AI models to not only access data but to understand its context, units, relationships, and constraints—critical for generating valid and reproducible outputs. In essence, FAIR transforms data into a digital asset ready for computational discovery. One compelling example of this synergy is the concept of "chemputation"—the automation and digitization of chemical synthesis guided by standardized datasets, machine-readable protocols, and robotic systems. Developed by Cronin and colleagues [4], chemputation relies on FAIR principles to ensure that chemical experiments can be replicated, optimized, and shared between laboratories and even across continents. These systems are not just mimicking human behavior; they are enhancing it by identifying novel reactions, predicting outcomes, and generating hypotheses based on vast and interoperable chemical databases.

Similarly, projects like the Open Research Knowledge Graph (ORKG) are reshaping how scientific literature is processed. By converting research papers into structured semantic data, ORKG facilitates AI-driven literature reviews, comparison of findings across studies, and dynamic updating of knowledge bases. FAIR data ensures that these systems are not biased by incomplete inputs or ambiguous language, making them more robust and reliable. In the health sciences, FAIR-compliant clinical data is transforming precision medicine. Machine learning models trained on interoperable, well-annotated patient datasets are improving diagnostics, predicting treatment outcomes, and personalizing care pathways—all while preserving patient privacy through federated learning and secure data exchange. Even in global development fields such as agriculture and climate adaptation, AI models supported by FAIR weather, soil, and crop data are enhancing decision-making for farmers, governments, and NGOs alike. However, the promise of AI also underscores the risks of non-FAIR data: biased models, irreproducible analyses, and ethical concerns arising from opaque data pipelines. Without FAIR standards, the very tools meant to accelerate discovery could instead amplify inequity, misinformation, and exclusion. Ultimately, the future of science will be increasingly digital, data-intensive, and automated. Ensuring that this future is trustworthy, inclusive, and innovative depends on our collective ability to implement FAIR data principles at scale.

In the context of artificial intelligence and machine learning, the limitations of open but non-FAIR data become particularly evident. Open datasets that lack standardized ontologies or semantic annotations require extensive manual preprocessing, significantly increasing computational and human resource costs. FAIR-compliant datasets, designed for machine-to-machine interaction, facilitate efficient data ingestion, model training, and validation, thereby accelerating AI-driven scientific discovery. The distinction is also apparent in interdisciplinary and policy-relevant research, such as climate science and public health. Open datasets released by different agencies may be accessible but incompatible due to inconsistent formats or terminologies. FAIR-aligned frameworks enable interoperability across institutional and national boundaries, supporting integrated modeling, evidence synthesis, and timely policy interventions. Furthermore, in long-term data stewardship, open data initiatives often prioritize immediate access over sustainability. FAIR data principles emphasize persistent identifiers, version control, and clear reuse licenses, ensuring that datasets remain interpretable and reusable well beyond their initial publication. This distinction is particularly critical for longitudinal studies and large-scale monitoring programs. By consistently integrating FAIR versus open data distinctions across applications—including publishing, AI, interdisciplinary research, and data stewardship—this manuscript demonstrates that FAIR principles provide an operational framework that complements open data, transforming accessibility into actionable and reusable scientific knowledge [5-20].

Conclusion

FAIR data is reshaping the norms of scholarly communication, aligning scientific practice with values of openness, transparency, and reuse. As research grows more digital, collaborative, and automated, FAIR principles will become a cornerstone of responsible science. Investing in FAIR is investing in the future of knowledge. The FAIR data principles represent more than just a technical framework—they embody a cultural and structural shift in how science is conducted, shared, and reused in the 21st century. By making research data Findable, Accessible, Interoperable, and Reusable, FAIR enhances transparency, boosts reproducibility, and enables data-driven discovery at both human and machine scales. As global challenges—ranging from pandemics and climate change to sustainable development—demand more collaborative and responsive science, FAIR data provides the connective tissue across disciplines, institutions, and borders. It empowers researchers from diverse backgrounds, facilitates interdisciplinary insights, and enables responsible innovation powered by AI and automation. Yet, the widespread adoption of FAIR principles is not guaranteed. It requires coordinated efforts from researchers, institutions, funders, journals, and governments to invest in training, infrastructure, policy, and incentives. It also requires a cultural shift where data stewardship is valued as a core part of scientific excellence—not

an afterthought. In a world increasingly defined by information flows, the quality and accessibility of data will shape the pace and impact of scientific progress. Embracing FAIR is no longer optional—it is essential. For science to remain a force for good, FAIR data must be seen as a cornerstone of open, inclusive, and future-ready scholarly communication.

References

1. Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data. 2016;3(1):160018. Available from: https://doi.org/10.1038/sdata.2016.18

2. Chen S, Liu G, Liu Q. Research progress and implementation of FAIR principles for scientific data management. J Libr Inf Sci Agric. 2022;34(8):30. Available from: https://doi.org/10.13998/j.cnki.issn1002-1248.22-0238

3. Schindler S, Dembska M, Bonnarel F, Lacy M, Johnson M, Desai V. Data documentation beyond provenance: Metadata, research data management, FAIR principles. In: Data-Intensive Radio Astronomy: Bringing Astrophysics to the Exabyte Era. Cham: Springer International Publishing; 2024;397-418. Available from: https://doi.org/10.1007/978-3-031-58468-8_14

4. Jati PHP, Lin Y, Nodehi S, Cahyono DB, van Reisen M. FAIR versus open data: A comparison of objectives and principles. Data Intell. 2022;4(4):867-81. Available from: https://doi.org/10.1162/dint_a_00176

5. Carballo-Garcia A, Boté-Vericad JJ. Fair data: History and present context. Cent Eur J Educ Res. 2022;4(2):45-53. Available from: https://doi.org/10.37441/cejer/2022/4/2/11379

6. Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, et al. Addendum: The FAIR Guiding Principles for scientific data management and stewardship. Sci Data. 2019;6(1):6. Available from: https://doi.org/10.1038/sdata.2016.18

7. Rogushina JV, Grishanova IJ. Study of principles, models, and methods of the FAIR paradigm of scientific data management for analysis of big data metadata. Probl Program. 2021;(4):26-35. Available from: https://doi.org/10.15407/pp2021.04.026

8. Wang M, Savard D. The FAIR principles and research data management. In: Research Data Management in the Canadian Context. 2023.

9. Dunning A, De Smaele M, Böhmer J. Are the FAIR data principles fair? Int J Digit Curation. 2017;12(2):177-95. Available from: http://dx.doi.org/10.2218/ijdc.v0i0.0

10. Bühler MM, Calzada I, Cane I, Jelinek T, Kapoor A, Mannan M, et al. Data cooperatives as catalysts for collaboration, data sharing, and the (trans) formation of the digital commons. Preprints. 2023. Available from: https://www.preprints.org/manuscript/202304.0130/v1

11. Chen X, Jagerhorn M. Implementing FAIR workflows along the research lifecycle. Procedia Comput Sci. 2022;211:83-92. Available from: https://doi.org/10.1016/j.procs.2022.10.179

12. Marshall CP, Schumann J, Trunschke A. Achieving digital catalysis: Strategies for data acquisition, storage, and use. Angew Chem Int Ed. 2023;62(30):e202302971. Available from: https://doi.org/10.1002/anie.202302971

13. Jacobsen A, de Miranda Azevedo R, Juty N, Batista D, Coles S, Cornet R, et al. FAIR principles: Interpretations and implementation considerations. Data Intell. 2020;2(1-2):10-29. Available from: https://doi.org/10.1162/dint_r_00024

14. Ugochukwu AI, Phillips PW. Open data ownership and sharing: Challenges and opportunities for application of FAIR principles and a checklist for data managers. J Agric Food Res. 2024;16:101157. Available from: https://doi.org/10.1016/j.jafr.2024.101157

15. Abiteboul S, Stoyanovich J. Transparency, fairness, data protection, neutrality: Data management challenges in the face of new regulation. J Data Inf Qual. 2019;11(3):1-9. Available from: https://doi.org/10.1145/3310231

16. van Reisen M, Stokmans M, Basajja M, Ong'ayo AO, Kirkpatrick C, Mons B. Towards the tipping point for FAIR implementation. Data Intell. 2020;2(1-2):264-75. Available from: https://doi.org/10.1162/dint_a_00049

17. Chen P, Wu L, Wang L. AI fairness in data management and analytics: A review on challenges, methodologies, and applications. Appl Sci. 2023;13(18):10258. Available from: https://doi.org/10.3390/app131810258

18. Mugahid D, Lyon J, Demurjian C, Eolin N, Whittaker C, Godek M, et al. A practical guide to FAIR data management in the age of multi-OMICS and AI. Front Immunol. 2025;15:1439434. Available from: https://doi.org/10.3389/fimmu.2024.1439434

19. Borycz J, Carroll B. Implementing FAIR data for people and machines: Impacts and implications—results of a research data community workshop. Inf Serv Use. 2020;40(1-2):71-85. Available from: https://doi.org/10.3233/ISU-200083

20. Huerta EA, Blaiszik B, Brinson LC, Bouchard KE, Diaz D, Doglioni C, et al. FAIR for AI: An interdisciplinary and international community-building perspective. Sci Data. 2023;10(1):487. Available from: https://doi.org/10.1038/s41597-023-02298-6