DOI : 10.18843/rwjasc/v13si/03
DOI URL : http://dx.doi.org/10.18843/rwjasc/v13si/03

# Emerging Trends in Digital Libraries: A Study of Open Access Resources

**Dr. Shankar Kadam**
*Librarian*
KBH Law College, LVH Marg, Malegaon Camp, Nashik 423105

**ABSTRACT**

**Purpose:** This study examines emerging trends in Open Access (OA) within the context of digital libraries, focusing on how OA has transformed the discovery, preservation, and dissemination of scholarly resources. It aims to highlight the significant developments in OA practices, policies, and infrastructures between 2015 and 2020. **Methodology:** The paper synthesizes empirical studies and meta-research published during 2015–2020. It draws upon large-scale analyses, policy reviews, and case studies to evaluate the expansion of OA content, the development of discovery tools, and the role of institutional and governmental mandates in shaping digital library practices. **Findings:** The analysis reveals several notable trends: (a) rapid growth of OA repositories, Gold/Hybrid journals, and preprint servers; (b) strengthening of infrastructures and discovery services such as DOAJ, CORE, and Unpaywall; (c) increasing influence of mandates and policies on OA compliance; (d) integration of advanced services within digital libraries, including metadata enhancement, user experience innovations, and text-and-data mining. Alongside these advancements, challenges remain in ensuring sustainability, addressing global inequities, maintaining quality, and securing long-term preservation. **Implications:** For librarians, repository managers, policymakers, and funders, the findings underscore the need for sustainable models, greater collaboration, and continuous innovation in digital library services. The results suggest that future success depends on balancing growth with equity and long-term preservation strategies. **Originality:** This study contributes to the discourse on digital libraries by framing OA not only as a matter of increased content availability but also as a qualitative shift in infrastructure, policy enforcement, and technological innovation. It consolidates insights from the 2015–2020 period to inform future directions in digital library development.

**Keywords:** *Open Access, Digital Libraries, Institutional Repositories, Unpaywall, CORE, DOAJ, Open Science, Policy Impact*

## INTRODUCTION:

Digital libraries have evolved from static repositories of digitized heritage materials into dynamic platforms supporting the full life cycle of scholarly communication. In recent years, Open Access (OA) has emerged not merely as an adjunct but as a defining characteristic of how contemporary digital libraries operate. Rather than just delivering digitized books or archival content, modern digital libraries increasingly serve as the gateways to global, freely accessible research output (journals, preprints, datasets, educational resources). Digital libraries increasingly centre their services on open content, not only digitized heritage but also scholarly outputs and educational resources made available through Open Access (OA) pathways. The Open Access ecosystem matured from a fragmented collection of journals and repositories to an interconnected infrastructure of aggregators, discovery tools, and policy levers that collectively influence discoverability and reuse, *(Knoth, P., et al., 2018); (Piwowar, H., et al., 2018)*.

Between 2015 and 2020, the Open Access ecosystem matured significantly. The number of openly accessible articles rose, discovery infrastructure (aggregators, APIs) improved, institutional mandates strengthened, and digital libraries began integrating value-added services (e.g., text mining, enriched metadata). Yet this period also revealed persistent challenges: financial models, sustainability, quality control, equity across regions, and long-term preservation.

This study reviews and synthesizes relevant literature from 2015 to 2020, focusing on how Open Access developments influence digital library practice. The objective is to identify key trends and implications, rather than achieve comprehensiveness. The structure proceeds from content growth to infrastructure, policy, services, challenges, and implications for libraries.

**Methodology and Scope:**

Given the broad and evolving nature of Open Access and digital libraries, this research adopts a literature synthesis approach. It draws upon empirical large-scale studies, meta-research analyses, policy reviews, and case studies published. Key sources include bibliometric studies (e.g., *(Piwowar, H., et al., 2018)*, infrastructure documentation, *(Unpaywall, 2020); (UNESCO, 2019); (CORE blog, 2018); (DOAJ (Directory of Open Access Journals), 2019)*, policy assessments, *(Huang, C. K., et al., 2020); (Swan, A., Gargouri, Y., Hunt, M., & Harnad, S., 2015)* and examinations of digital library innovations (e.g., *(Schultz, T. A., 2019)*. The selection aimed to balance quantitative analyses and qualitative insights, focusing on works that examine cross-institutional or ecosystem-level trends rather than small local case studies. The review organizes trends into thematic categories to facilitate understanding of how Open Access affects digital library ecosystems.

**Growth and Distribution of Open Access Content:**

One of the landmark reviews is by, *(Piwowar, H., et al., 2018)*, who estimated that roughly 28% of all journal articles (with Crossref DOIs) were openly accessible in their sampling, with growth being driven particularly by Gold and Hybrid Open Access pathways, *(Piwowar, H., et al., 2018)*. They also observed a "reader advantage": users accessing articles via Unpaywall encountered Open Access versions about 47% of the time, indicating that Open Access availability is disproportionately higher among recently consulted content, *(Piwowar, H., et al., 2018)*.

Observers documented acceleration in 2018–2019, with many Open Access indicators rising faster than historical trends, *(Morrison, H., 2019); (Unpaywall, 2020)*, metrics). Aggregation and discovery tools such as Unpaywall (launched 2017) and CORE (which by 2018 claimed to be the largest Open Access aggregator) materially improved the discoverability of Open Access content and enabled libraries to surface legally available copies to users, *(Piwowar, H., et al., 2018); (CORE blog, 2018)*.

Projections based on models of Open Access adoption suggest continued growth, by 2025 it is estimated that 44% of journal articles will be Open Access and 70% of article views will be directed to Open Access content, *(Piwowar, H., Priem, J., & et al., 2019)*, signalling that access patterns may outpace mere publication counts.

As Open Access matured, finer distinctions of Open Access types matter more. Green Open Access (repository deposit), Gold Open Access (fully open journals), Hybrid Open Access (author pays in mixed journals), and Bronze Open Access (free-to-read without explicit license) each play significant roles. *(Piwowar, H., et al., 2018)*, noted that the most common mechanism encountered was Bronze Open Access, where content is made free by publishers, but without clear licensing. This raises concerns about reuse rights and permanence, *(Piwowar, H., et al., 2018)*.

Preprints (e.g., arXiv, bioRxiv) also gained visibility within Open Access ecosystems during this period. Users and libraries began to treat preprint servers as integral to discovery and open scholarship pipelines. While many preprints are not peer-reviewed, they expand access especially in fast-moving sciences.

The diversification of Open Access models means digital libraries must handle multiple Open Access categories and assist users in understanding license constraints and provenance.

**Infrastructure and Discovery: Aggregators, Tools, and Metadata:**

A significant trend is the rise of aggregator and discovery tools that harvest, index, and serve Open Access content globally.

- Unpaywall, launched around 2017, functions as a browser extension and backend API to detect Open Access versions of articles. *(Schultz, T. A., 2019)*, assesses the effectiveness of Open Access -finding tools, finding that Unpaywall outperformed several competitors in locating freely available copies, *(Schultz, T. A., 2019)*. Unpaywall continues to maintain a large database of Open Access articles and provides services to digital libraries for linking to Open Access content, *(Dhakal, K., 2019); (Unpaywall, 2020)*.
- CORE aggregates full-text OA content from institutional repositories and Open Access journals. In 2019, CORE released a browser extension ("CORE Discovery") to help users identify Open Access versions of articles even behind paywalls, *(CORE blog, 2018)*. The aggregation and recommender infrastructure of CORE position it as a key back-end for many library systems, *(Knoth, P., et al., 2018)*.
- DOAJ (Directory of Open Access Journals) remains a trusted directory of quality Open Access journals, serving as a curation and filtering resource for libraries assessing journal quality and indexing.

These tools enable digital libraries to integrate Open Access content into their discovery layers, link resolvers, and resource guides, increasing findability of Open Access content.

Open Access infrastructure depends heavily on metadata quality and interoperability. Many repositories expose content via OAI-PMH, but inconsistent metadata, missing license information, or non-standard structures hinder harvesting, *(Akbaritabar, A. & Stahlschmidt, S., 2019)*. In their analysis, *(Akbaritabar, A. & Stahlschmidt, S., 2019)*, highlight that matching Open Access status across databases (Crossref, Unpaywall, DOAJ) often reveals contradictions (with >25% contradictory cases), reflecting the complexity of reliably determining Open Access status.

Digital libraries increasingly invest in metadata clean-up, linkage of dataset and publication metadata, license normalization, and schema alignment (e.g., adopting Schema.org/CreativeWork attributes) to bridge institutional repositories with global aggregators.

For digital libraries infrastructure developments meant they could integrate Open Access content into local discovery layers, link resolvers, and discovery services more reliably. However, success depended on consistent, high-quality metadata and repository compliance with harvesting protocols (OAI-PMH, schema standards), an area where many repositories continued to have uneven practices, *(Charalampous, A. & Knoth, P., 2017)*.

**Policy, Mandates, and Institutional Behaviour:**

Between the year 2015 - 20, many funders and institutions strengthened Open Access policies, requiring deposit, linking Open Access compliance to evaluation, and enforcing monitoring mechanisms. The PASTEUR4OA project, *(Swan, A., Gargouri, Y., Hunt, M., & Harnad, S., 2015)*, emphasized that stronger

mandates (e.g., mandatory deposit, linking deposit to performance evaluation) led to higher compliance rates.

Policy interventions had a decisive influence on Open Access uptake. The PASTEUR4OA analysis, *(Swan, A., Gargouri, Y., Hunt, M., & Harnad, S., 2015)*, showed that institutional and funder mandates with stronger conditions (e.g., must-deposit, linking deposit to evaluation) produced higher repository deposit rates.

Huang and others, *(Huang, C. K., et al., 2020)*, conducted a meta-research study showing that institutions with enforced Open Access policies tend to produce significantly more Open Access outputs compared to institutions with non-enforced or voluntary policies, illustrating policy efficacy.

Digital libraries and institutional repositories became active partners in policy implementation. Libraries often act as Open Access compliance agents, offering deposit workflows, liaising with research offices, integrating with evaluation systems, and providing analytics dashboards to monitor Open Access uptake. In this role, libraries shift from passive storage to advocacy, support, and enforcement.

Another policy-related trend is the "flipping" of formerly subscription journals to Open Access models. *(Momeni, F., Fraser, N., Peters, I., & Mayr, P., 2019)*, analyse a sample of flipped journals and observe that, for many, journal impact factors improved post-flip, though article-level citation advantages were not always clear. Flipping requires careful planning, institution support, and sustainable funding models.

**Service and Technical Innovations in Digital Libraries: (Metadata, UX, and Text/Data mining):**

The growth of Open Access content and aggregator services enables digital libraries to offer TDM services and APIs to researchers. CORE's API and dataset dumps support large-scale text mining and machine access, *(Knoth, P., et al., 2018)*. Libraries can harness Open Access corpus infrastructure to build in-house analytics, topic clustering, recommendation systems, and semantic enrichment.

However, successful TDM requires clear license metadata (e.g., CC licenses), machine-readability, and rights clarity. Variability in licensing (especially in Bronze Open Access) and inconsistent metadata complicate automated reuse.

Digital libraries began embedding recommendation systems, related article suggestions, full-text indexing where Open Access allows, and visual discovery overlays (e.g., highlighting Open Access availability in search results). These UX improvements help users discover more content without needing domain expertise in Open Access.

Libraries also experiment with integrated dashboards showing Open Access availability metrics, usage statistics of Open Access vs closed content, and linking user behaviour to Open Access services.

Open Educational Resources (OER) proliferated alongside Open Access scholarly content. UNESCO's 2019 Recommendation on OER catalysed library involvement in curating, discovering, and promoting educational Open Access content, *(UNESCO, 2019)*. Many digital libraries began to present OER side-by-side with research Open Access content, supporting both research and teaching missions.

**Challenges and Constraints:**

Despite growth, the period exposed key challenges. First, quality control and the proliferation of predatory Open Access outlets raised concerns about trust and curation, *(Tennant, J. P., et al.)*. DOAJ and other quality lists helped but cannot cover all low-quality actors. Second, financial sustainability questions intensified, APC (article processing charge) models shifted costs rather than eliminating them, creating

inequities for authors without funds, *(Piwowar, H., et al., 2018); (Aspesi, C., 2019)*. Third, disparities persisted between well-resourced institutions (capable of supporting green and gold Open Access) and under-resourced institutions or regions where repository infrastructure, funding, or policy support were weak, *(Huang, C. K., et al., 2020)*.

Digital libraries are thus tasked with balancing discovery/aggregation goals against curation and equitable access strategies, including advocating for diamond/green Open Access models and community-owned infrastructure.

- **Sustainability and business models:**
  While Open Access removes reader paywalls, it does not eliminate costs. APC (Article Processing Charge) models shift costs to authors or funders, raising equity concerns for underfunded researchers. Some journals use hybrid or partial OA models that perpetuate barriers. The instability of funding for Open Access infrastructure (repositories, aggregator services) is also a concern.
- **Equity and capacity disparities:**
  Not all institutions or regions are equally equipped to participate in Open Access infrastructure. Well-resourced universities in North America or Europe often lead, while institutions in low- and middle-income countries may lack repository support, paper funding, or policy backing. This leads to uneven Open Access engagement and visibility.
- **Quality control, predatory publishing, and trust:**
  The proliferation of predatory publishers and low-quality Open Access outlets undermines trust in Open Access. Libraries must continue to vet content using tools like DOAJ, Beall's lists (deprecated), and publisher vetting protocols, while educating users about legitimacy.
- **Preservation and long-term access:**
  Open access is not identical with permanence. Ensuring long-term access to Open Access content (especially if publishers change policies) requires robust preservation strategies: LOCKSS, CLOCKSS, institutional and national preservation programs, and coordinated archiving architectures. Smaller repositories often lack the resources to implement full preservation mechanisms.
- **Metadata inconsistencies and OA status ambiguity:**
  As noted earlier, deriving correct Open Access status is nontrivial due to metadata discrepancies, licensing ambiguity, contradictory indexing, or missing license fields, *(Akbaritabar, A. & Stahlschmidt, S., 2019)*. This complicates discovery, linking, analytics, and compliance.

**Implications for Digital Libraries:**

Drawing on the trends above, we highlight several key implications for digital libraries:

1. ***Integrate Open Access discovery infrastructure:*** Libraries should embed services like Unpaywall and CORE into discovery layers and link resolvers to surface Open Access versions transparently to users.
2. ***Invest in metadata quality and interoperability:*** Accurate licensing, consistent schemas, and enriched metadata are foundational to effective harvesting, linking, and reuse.
3. ***Position libraries as Open Access compliance hubs:*** Librarians should proactively support deposit, policy adherence, monitoring, and researcher outreach, shifting from custodial to strategic roles.
4. ***Offer value-added services:*** TDM services, content recommendations, usage analytics, and semantic enrichment can differentiate library services powered by Open Access infrastructure.
5. ***Support sustainable Open Access models:*** Libraries should advocate and participate in models beyond APC (e.g., diamond Open Access, institutional subsidies, consortial publishing) and consider shared infrastructure strategies.

6. ***Collaborate on preservation and archiving:*** Digital libraries should partner with national/international preservation platforms to ensure long-term availability of Open Access content.
7. ***Promote equity and capacity building:*** Digital libraries in resource-poor contexts deserve support in acquiring infrastructure, skills, and policy advocacy to fully participate in Open Access ecosystems.

## CONCLUSION:

Since last some years, Open Access evolved from a peripheral ideal to a central structural pillar of modern digital libraries. The quantitative growth of Open Access content was accompanied by critical qualitative shifts: maturation of aggregator infrastructure (CORE, Unpaywall), the strengthening of Open Access mandates, service innovations in digital libraries, and more nuanced challenges in sustainability, equity, trust, and preservation.

Digital libraries now have a central role in shaping how Open Access is delivered and experienced: not merely as repositories, but as orchestration points linking content, policy, tools, and user services. The journey ahead demands attention to sustainable funding, robust metadata practices, collaborative preservation, and equitable participation across institutions.

By embracing these trends and challenges, digital libraries can help realize Open Access's promise: making scholarly and educational content more accessible, reusable, discoverable, and enduring.

## REFERENCES:

Akbaritabar, A., & Stahlschmidt, S. (2019). Merits and limits: Applying open data to monitor open access publications in bibliometric databases. *arXiv*. Retrieved from https://arxiv.org/abs/1902.03937

Aspesi, C. (2019). The changing academic publishing industry (research report). *University of Nebraska Digital Commons*. Retrieved from https://digitalcommons.unl.edu/scholcom/1100

Charalampous, A., & Knoth, P. (2017). Classifying document types to enhance search and recommendations in digital libraries. *arXiv*. Retrieved from https://arxiv.org/abs/1707.04134

CORE blog. (2018, december 18). CORE becomes the world's largest open-access aggregator — 2018 edition. *CORE blog*. Retrieved from https://blog.core.ac.uk/2018/12/18/core-becomes-the-worlds-largest-open-access-aggregator-or-how-about-them-stats-2018-edition/

Dhakal, K. (2019). Unpaywall: A tool for discovering legally available OA articles. *Journal of the Medical Library Association, 107(3)*, 455-457. Retrieved from https://jmla.pitt.edu/ojs/index.php/jmla/article/view/650

DOAJ (Directory of Open Access Journals). (2019). Directory of Open Access Journals — publisher & journal survey and statistics. *DOAJ (2019 reports)*. Retrieved from https://doaj.org

Huang, C. K., Neylon, C., Hosking, R., Montgomery, L., Wilson, K. S., Ozaygen, A., & Brookes-Kenworthy, C. (2020). Meta-research: Evaluating the impact of open access policies on research institutions. *eLife, 9, e57067*. Retrieved from https://doi.org/10.7554/eLife.57067

Knoth, P., Cancellieri, M., Anastasiou, L., Pontika, N., Pride, D., & Cukier, J. (2018). CORE: A global aggregation service for open access papers (project & usage report, 2018). *CORE project communications*. Retrieved from https://core.ac.uk

Momeni, F., Fraser, N., Peters, I., & Mayr, P. (2019). From closed to open access: A case study of flipped journals. *arXiv*. Retrieved from https://arxiv.org/abs/1903.11682

Morrison, H. (2019, December 31). Dramatic growth of Open Access, (data notes & indicators). *Scholars Portal Dataverse*. doi:https://doi.org/10.5683/SP2/CHLOKU

Piwowar, H., Priem, J., & et al. (2019). The future of OA: A large-scale analysis projecting open access publication and readership. *bioRxiv*. Retrieved from https://www.biorxiv.org/content/10.1101/795310v1.full-text

Piwowar, H., Priem, J., Larivière, V., Alperin, J.P., Matthias, L., Norlander, B., . . . Haustein, S. (2018). The state of OA: A large-scale analysis of the prevalence and impact of Open Access articles. *PeerJ, 6, e4375*. doi:https://doi.org/10.7717/peerj/4375

Schultz, T. A. (2019). Assessing the effectiveness of open access finding tools. *ITAL: International Journal of Librarianship*. Retrieved from https://ital.corejournals.org/index.php/ital/article/view/11009

Swan, A., Gargouri, Y., Hunt, M., & Harnad, S. (2015). Open Access policy: Numbers, analysis, effectiveness (PASTEUR4OA findings). *PLOS/PASTEUR4OA project report*. Retrieved from https://arxiv.org/abs/1504.02261

Tennant, J. P., Waldner, F., Jacques, D. C., Masuzzo, P., Collister, L. B., & Hartgerink, C. H. J. (n.d.). The academic, economic and societal impacts of Open Access: An evidence-based review. *F1000Research, 5, 632.* doi:https://doi.org/10.12688/f1000research.8460.1

UNESCO. (2019). Recommendation on Open Educational Resources (OER). Retrieved from https://unesdoc.unesco.org/ark:/48223/pf0000386595

Unpaywall. (2020). Unpaywall metrics and dataset. *(public snapshots, 2017–2020)*. Retrieved from https://unpaywall.org

******