Challenges faced by the publishing industry

Veeraraghavan Balasubramanian

Abstract

This paper explores the key challenges faced by the publishing industry, especially in the context of LaTeX-based workflows, including manuscript submission complexities, operational hurdles, accessibility issues, supply chain logistics, and technological evolution. It provides insights into industry support for LaTeX, recent trends, and future directions aimed at fostering innovation, collaboration, and inclusivity in scholarly publishing. The discussion emphasizes the critical role of technological advancements such as AI and automation in transforming the landscape.

1 Introduction

The landscape of STM (Scientific, Technical, and Medical) publishing content involves various challenges, especially when using LaTeX as the main authoring, proofing, and typesetting tool. The primary goal for publishers is to produce clear, well-organized, and high-quality files that can be easily shared and accessed across different platforms such as PDFs, EPUBs, web pages, and XML. Ensuring that content is extracted accurately, remains consistent, accessible, and that the formatting is preserved across all formats can be complex. Achieving this often requires specialized workflows and tools.

This study aims to highlight some of the key challenges faced in the publishing industry and to share best practices for effectively addressing them.

2 Key challenges in the publishing industry

2.1 Managing manuscript submissions and workflow complexity

Author content creation: Many researchers, especially in STM fields, prefer LaTeX for its precise formatting of mathematical equations and references. However, converting LaTeX documents into publisher-specific formats (e.g., Word, XML) often requires complex workflows, leading to potential errors and inconsistencies.

In this scenario, the author is developing a manuscript using LaTeX, primarily concentrating on the content and its formatting. He is putting considerable effort into customizing the layout and appearance to match his specific design vision. However, from a publisher's perspective, there are additional requirements, such as adhering to specific page layouts and structuring the content into semantic tags within XML. These publishing needs often necessitate adjustments and conversions beyond the author's initial formatting to ensure the content meets industry standards and is suitable for various platforms.

Submission platforms: Existing submission portals often struggle with handling large LaTeX files, macro-heavy documents, and package dependencies. Compatibility issues can cause delays, and user interfaces may not be optimized for LaTeX users.

Workflow coordination: The peer review process involves multiple stakeholders — authors, reviewers, editors, and production teams. Synchronizing feedback, revisions, and approvals across different formats and platforms is complex, often resulting in bottlenecks and delays.

Images/multimedia: High-quality images, datasets, and multimedia content enhance scientific value but pose challenges in maintaining quality and functionality across formats like PDF, HTML, and EPUB. Embedding interactive figures and ensuring compatibility without loss of fidelity is difficult.

2.2 Operational challenges

Coordination of processes: Managing multiple review cycles, copyediting, typesetting, and final approval stages requires seamless communication. Discrepancies in formatting and style guidelines across journals or conferences add layers of complexity.

Maintaining fidelity of LaTeX files: Ensuring that macros, custom packages, and formatting are preserved during revisions and conversions is crucial. Loss of formatting or content integrity can compromise scientific accuracy and professionalism.

Meeting deadlines: Automating repetitive tasks has reduced turnaround time, but last-minute revisions still cause bottlenecks. Rapid compilation and review cycles are essential to meet publication schedules.

2.3 Accessibility challenges

Web Content Accessibility Guidelines (WCAG): Compliance for PDFs involves ensuring that PDF documents are accessible to all users, including those with disabilities. To achieve WCAG compliance for PDFs involves:

- Proper tagging and logical structure
- Searchable, selectable text
- Alt text for images
- High contrast and clear fonts
- Keyboard navigation and bookmarks

doi.org/10.47397/tb/46-2/tb143balasubramanian-challenges

- Embedding metadata
- Testing with accessibility tools and screen readers

Following WCAG 2.1 and PDF/UA standards is essential for Equality Act Accessibility (EAA) and Americans with Disabilities Act (ADA) compliance. Additionally, regularly reviewing and updating PDFs helps sustain accessibility, and offering alternative formats ensures wider access. Limitations include the ongoing development of PDF creation modules in LaTeX, which are still evolving. Providing proper training for content creators on accessible document creation is also vital to maintain ongoing WCAG compliance.

Making content accessible: Adding alt text, MathML, and semantic markup enhances accessibility but also increases production effort and costs. Publishing faces the challenge of justifying these value-added activities and securing the necessary funding. To address this, solutions such as zero-level automation and working solutions to provide opportunities for authors to review and correct automatically generated alt-text are being explored.

Content conversion: Converting LaTeX to XML, HTML, or EPUB while preserving scientific accuracy and accessibility features remains complex, often requiring human oversight and intervention in many steps of the process.

2.4 Supply chain and Logistics

File transfer: Securely transferring large LaTeX files, figures, and bibliographies to publishers or printers across borders can lead to delays.

Timely distribution: During global crises, such as pandemics or widespread disruptions, delays in transferring large files and logistical challenges can significantly extend publication timelines. This can hinder the timely dissemination of research, findings, and critical information, potentially impacting academic progress, public awareness, and the advancement of knowledge. Therefore, ensuring the integrity and security of file assets at every stage of the publication process is essential to mitigate delays and maintain efficient workflows.

Version control: Managing multiple simultaneous revisions and ensuring that the most recent, approved version is used is a complex process. It requires effective version control, clear communication among team members, and meticulous tracking of changes to prevent errors, inconsistencies, or outdated content from being published. Coordinating revisions across different stages, formats, and platforms adds to the challenge, making it essential to implement robust workflows and tools to maintain accuracy and streamline the publication process.

Quality assurance: Remote validation and prepress checks are crucial steps to ensure quality before publication; however, they can sometimes miss errors. This oversight may result in the need for reprints, which can be costly and cause delays. Therefore, combining thorough review processes with multiple validation stages is important to minimize errors and ensure the final product meets quality standards.

External resources: Managing external images, macros, supplementary assets, and linked resources for compatibility introduces significant logistical complexity. Ensuring all assets are correctly linked, compatible across different platforms, and up-to-date requires careful coordination and meticulous management to prevent broken links, missing content, or formatting issues in the final publication.

3 Industry support and adoption

Overleaf user base (2024): Over 22 million users worldwide, including researchers and students [1].

Journal policies: Over 80% of top-tier journals (e.g., IEEE, ACM, Springer, Elsevier) recommend or require LaTeX submissions [2].

Research adoption: Approximately 65% of researchers in physics, mathematics, and computer science rely on LaTeX. However, there is limited research available that precisely quantifies the total usage of LaTeX papers, and no reliable or definitive data currently exists to support or cite this figure.

Educational inclusion: Many universities embed LaTeX into curricula; over 70% of new researchers are trained in LaTeX [3].

Reference management: BibTeX and BibLaTeX continue to be essential, with ongoing community support [4].

4 Future outlook

LaTeX support: LaTeX module development needs to be accelerated to ensure access to the latest content in repositories, facilitate efficient feedback and review processes, and support ongoing repository maintenance. Streamlining these efforts can help reduce delays, improve content accuracy, and enhance overall productivity in the publication workflow.

Innovation through technology: Publishing organizations are exploring AI and automation technologies, which are expected to reduce manuscript processing times by up to 40%–60%. This can lead to faster publication cycles, increased efficiency, and a more streamlined workflow in the publishing process.

Collaborative ecosystems: Developing interoperable systems and forming strategic partnerships will help streamline workflows across various stages such as submission platforms, structured content extraction, authoring, reviewing, and formatting. This integration can enhance efficiency, reduce manual effort, and ensure smoother coordination throughout the publishing process.

Accessibility and inclusivity: Prioritizing accessible formats, multimedia, and adherence to standards will help expand the reach of published content. Additionally, the LaTeX development team should work to improve solutions that address dependencies on accessibility features after PDF extraction, ensuring better integration and enhanced accessibility throughout the publication workflow.

5 Conclusion

After 30 years of experience as a LaTeX coder, developing tools using LaTeX engines, adjusting fonts, implementing accessibility features for publishers, handling XML conversions, and automating typesetting processes, I believe that the publishing industry can effectively manage LaTeX as a comprehensive typesetting and content orchestration platform. Despite the numerous challenges related to LaTeX workflows and accessibility, the developments highlighted in this paper — such as AI, automation, and collaborative platforms — offer promising solutions. These advancements are paving the way for a more efficient, inclusive, and sustainable future in publishing.

References

[1] Overleaf (2024). *Overleaf platform.* `www.overleaf.com`

[2] Springer (2024). *Springer publications.* `www.springer.com`

[3] Harvard University (2024). *Harvard University publications.* `www.harvard.edu`

[4] CTAN (2024). *Comprehensive TeX Archive Network.* `ctan.org`

◇ Veeraraghavan Balasubramanian
Senior Product Manager, Elsevier