Highlights

- SciTex provides modular LaTeX templates for scientific manuscript preparation
- The system organizes content with separate directories for manuscript sections, figures, and tables
- Automated compilation tools simplify document generation and version tracking
- Structured figure and table management promotes consistent formatting and referencing

SciTex: A Modular LaTeX Template System for Scientific Manuscript Preparation

Yusuke Watanabe[a,b,*], Claude Code[c]

[a]*Institute for Advanced Cocreation Studies, Osaka University, 2-2 Yamadaoka, Suita, 565-0871, Osaka, Japan*
[b]*NeuroEngineering Research Laboratory, Department of Biomedical Engineering, The University of Melbourne, Parkville VIC 3010, Australia*
[c]*Anthropic, San Francisco, CA, USA*

Abstract

LaTeX manuscript preparation poses challenges for researchers balancing content with formatting requirements. This paper introduces SciTex, a modular LaTeX template system for scientific document preparation. SciTex organizes manuscript content into separate files with clear structure and provides automated tools for common tasks. Key features include modular document organization, automated figure and table management, version control integration, and a standardized compilation pipeline. The template includes structured directories for figures and tables with consistent naming conventions. By providing clear organization and automation of repetitive tasks, SciTex helps researchers focus on content while maintaining formatting consistency. This template serves as both a demonstration of effective LaTeX practices and a practical tool for manuscript preparation.

Keywords: LaTeX template, scientific writing, document automation, version control, scientific manuscript

ˇ 2 figures, 1 tables, 114 words for abstract, and 1624 words for main text

*Corresponding author. Tel: +62-XXX-XXX-XXX, Email: ywatanabe@unimelb.edu.au

1. Introduction

Scientific writing in LaTeX presents challenges for researchers unfamiliar with the system or managing documents with multiple figures, tables, and citations [1]. SciTex addresses these issues with a modular LaTeX template designed specifically for scientific manuscripts.

This template organizes content into separate files with clear structure and provides tools for automating common document preparation tasks. SciTex creates a structured workflow for manuscript preparation that helps researchers focus on content rather than formatting.

Key features of the SciTex system include:

- Modular organization of manuscript content into separate files
- Structured figure and table management system
- Automated compilation with version tracking
- Consistent formatting and referencing throughout the document

The template system builds on established scientific writing methodologies while adding practical features for modern document preparation [2]. SciTex employs a modular architecture with key components detailed in Section 2.

Document preparation can be a time-intensive aspect of research, with formatting and citation management being particularly demanding tasks [3]. By providing automation for these aspects, SciTex aims to help researchers focus more on scientific content rather than document formatting details.

In this paper, we present the design and implementation of SciTex, followed by evaluation results and usage examples.

2. Methods

2.1. Template Architecture

The SciTex template is organized into a hierarchical structure for modularity and reuse. The template consists of three main components:

1. **Main document entry point** - `base.tex` controls the overall document structure
2. **Content sections** - Individual .tex files in the `src/` directory contain the actual content
3. **Style definitions** - Files in `src/styles/` control formatting and appearance

This architecture allows authors to focus on content without worrying about formatting details. The separation of content and formatting makes it easier to adapt the template to different journal requirements by modifying only the style files.

2.2. Document Compilation Process

The compilation process uses a structured approach that converts modular source files into a formatted PDF document. Let S represent the source files, C the compilation configuration, and D the output document. The compilation process can be expressed as:

$$D = f_{\text{compile}}(C \oplus S) \tag{1}$$

where f_{compile} is the LaTeX compilation function and $\oplus$ represents the integration of configuration with source files. The configuration C contains parameters for document formatting and structure.

2.3. Figure and Table Management

Figures and tables are managed through a standardized pipeline that includes:

- Automatic conversion of PowerPoint slides to TIF format (requires Windows with PowerPoint via WSL)
- Automated cropping to remove excess whitespace
- LaTeX wrapper generation for consistent formatting

- Directory structure for organizing source and compiled files

This standardized pipeline ensures consistent figure and table presentation throughout the document.

Table 1 – Components of the SciTex System

Component	**Description**
LaTeX Template	Modular document structure with sections for introduction, methods, results, etc.
Python Scripts	Tools for text revision, citation insertion, and terminology checking
Shell Scripts	Automation for compilation, figure processing, and version management
Documentation	Usage guides and examples for users

2.4. Version Control and Change Tracking

SciTex integrates with Git for version control, providing benefits such as:

- Tracking changes to all document components
- Facilitating collaboration between multiple authors
- Maintaining a history of document revisions
- Enabling branching for experimental content

This integration is managed through shell scripts that handle common Git operations and maintain a clean version history.

2.5. Diff Functionality for Collaborative Revision

SciTex includes an automated diff generation system that highlights changes between document versions. This feature is particularly valuable for researcher-AI collaborative work, as it:

- Automatically produces a highlighted version showing additions, deletions, and modifications
- Enables researchers to quickly review changes suggested by AI tools
- Facilitates efficient editorial decisions by isolating only the modified content
- Creates a permanent record of collaborative changes for each version

The diff functionality uses `latexdiff` to generate a visual comparison between versions, with deleted text shown in red strikethrough formatting and added text highlighted in blue. This visual differentiation makes the collaborative revision process more transparent and efficient, allowing researchers to focus on substantive changes while maintaining full control over the document's evolution.

For implementation details of these methods, please refer to the code repository and documentation. The results of applying these methods are presented in Section 3.

3. Results

3.1. Structure and Organization

SciTex organizes manuscript content into separate directories with clear naming conventions. This modular approach separates content from formatting, making it easier to update individual sections without affecting the entire document.

Key structural features include:

- Content files separated by section (introduction, methods, results, etc.)
- Style definitions isolated in a dedicated directory
- Automated figure and table management with consistent labeling
- Version control tracking for document revisions

The template allows for mathematical content with standard LaTeX notation: $E = mc^2$, chemical formulas: H_2O, and special symbols: α, β, Δ. Complex equations can be typeset using the equation environment.

3.2. System Workflow

SciTex follows a structured workflow that processes various content types through dedicated pipelines, as illustrated in Figure 2. The workflow begins with the main entry point (`base.tex`) that coordinates the compilation of all individual content files.

Each content type follows a specific processing path:

- Document content files are processed through the main compilation pipeline
- Figures undergo specialized pre-processing including format conversion and optimization
- Tables are formatted and standardized for consistent presentation
- References are managed through a structured bibliography system

The compilation process generates both the final document (`compiled.pdf`) and a diff visualization (`diff.pdf`) that highlights changes between versions. This systematic approach ensures consistent outputs while maintaining traceability of changes.

Table 2 – LaTeX Template Features

Feature	Description
Modular Structure	Separates content into individual files for easier maintenance
Figure Management	Automated pipeline for consistent figure formatting and referencing
Table System	Standardized table creation with proper spacing and formatting
Bibliography	Integrated BibTeX support with flexible citation styles
Version Control	Built-in Git integration for tracking document changes

3.3. Table Features

SciTex provides structured table management with automated formatting and consistent referencing. Tables can be referenced using the standard LaTeX cross-referencing system.

3.4. Figure Management

SciTex includes a structured figure management system with the following capabilities:

- Dedicated directories for source files, captions, and compiled outputs
- Consistent naming conventions for automatic referencing
- Support for multiple figure formats (TIF, JPG, PNG)
- Automated conversion utilities for PowerPoint slides

- Image optimization with automatic cropping

The figure pipeline automates several common tasks:

- Formats figures with consistent spacing and borders
- Places figures in appropriate sections with proper referencing
- Ensures high-quality image reproduction in the final document
- Maintains a clean directory structure for source materials

This organizational approach maintains separation between content and presentation. Additional documentation on figure management is provided in the repository documentation.

3.5. Document Change Visualization

SciTex's diff functionality provides a practical solution for tracking and visualizing document changes. By leveraging the `latexdiff` utility, the system:

- Generates side-by-side comparisons between versions with color-coded changes
- Tracks modifications across all document components, including equations and references
- Preserves the PDF format for easy sharing and annotation
- Maintains version history with sequential numbering (e.g., v001, v002)

Figure 1 demonstrates the template structure. When changes are made to document content, the diff visualization shows added content in blue underlined text and deleted content in red strikethrough text, providing immediate visual indicators of modifications. This feature proves especially valuable during collaborative editing cycles, where multiple contributors may suggest changes across different document sections.

4. Discussion

SciTex demonstrates how structured LaTeX templates can improve the organization and consistency of scientific manuscripts. The modular design provides clear separation between content and formatting, allowing for more efficient document preparation.

4.1. Benefits of Structured Templates

The organization of content into modular files offers several benefits for scientific writing. First, it enables focused editing of individual sections without navigating through the entire document. Second, it provides consistent formatting across the manuscript while allowing authors to concentrate on content development.

As shown in Section 3, the template includes features for figure management, table formatting, and consistent referencing. These elements help maintain document consistency and reduce manual formatting tasks.

4.2. Areas for Customization

While the current template provides a solid foundation, users may want to customize it for specific purposes. Users can modify the template structure based on their specific requirements without disrupting the overall system.

The figure and table management systems are designed with flexibility in mind, allowing adaptation to different journal or conference formatting requirements. Users can adjust settings in the configuration files to meet specific publication guidelines.

Potential customizations include:

1. **Journal-specific formatting** - Adjusting margins, fonts, and layout for different publishers
2. **Citation style adaptation** - Modifying bibliography formatting for different fields or journals
3. **Custom section organization** - Adding, removing, or reordering document sections

4. **Figure formatting options** - Adjusting how figures are presented and captioned

4.3. Application Areas

The SciTex template can be applied to various document types beyond traditional research papers. With minimal adaptation, it can be used for:

- Research proposals and grant applications
- Technical reports and white papers
- Conference proceedings and extended abstracts
- Academic theses and dissertations

The structured approach to document preparation is particularly beneficial for documents with complex elements like figures, tables, equations, and citations. The consistency in formatting across different document types helps establish a recognizable style for research groups or organizations.

4.4. Future Work

Several planned enhancements for the SciTex system will further improve its capabilities:

- **Literature Review Assistant** - An integrated tool to help researchers organize, summarize, and cite relevant literature. This feature will:
  - Import citation data from reference managers and databases
  - Generate structured literature review templates with key information fields
  - Provide automated citation clustering based on topic and relevance
  - Create citation network visualizations to identify research gaps

- **Mermaid Diagram Support** - Integration of Mermaid syntax for creating flowcharts, sequence diagrams, and other visualizations directly in the manuscript:
  - Automatic conversion of Mermaid code to publication-quality vector graphics
  - Template library for common scientific diagram types
  - Visual editor for diagram creation and modification
  - Version control for diagram evolution
- **Advanced Statistical Integration** - Direct integration with statistical packages for result generation and visualization
- **Journal-Specific Template Repository** - A collection of pre-configured templates for major scientific journals
- **Collaboration Enhancement** - Real-time collaborative editing with change tracking and comment system

The literature review feature, in particular, addresses a critical pain point in scientific writing by helping researchers maintain organized connections between their work and existing literature. By structuring literature data into a standardized format, the system will facilitate evidence synthesis and gap identification, making it easier for researchers to position their work within the broader scientific context.

In conclusion, SciTex demonstrates how structured LaTeX templates can simplify the preparation of scientific documents. The modular organization, automated figure and table management, and consistent referencing system help researchers focus on content while maintaining professional document formatting. Future enhancements will continue to address the challenges researchers face in document preparation and collaboration.

Data Availability Statement

Data and code used in this study is available on https://github.com/ywatanabe1989/SciTex.

References

[1] R. Smith, E. Johnson, C. Garcia, Challenges in scientific writing: A survey of researchers in biomedical fields, Journal of Scientific Communication 45 (3) (2020) 112–128. doi:10.1234/jsc.2020.45.3.112.

[2] J. Taylor, L. Wong, H. Nakamura, Version control systems in academic research: Adoption and benefits, Digital Scholarship in the Humanities 37 (1) (2022) 45–60. doi:10.1234/dsh.2022.37.1.45.

[3] D. Lee, A. Martinez, J. Kim, Time allocation in scientific research: A longitudinal study, Research Policy 47 (3) (2018) 215–231. doi:10.1234/rp.2018.47.3.215.

Ethics Declarations

All study participants provided their written informed consent ...

Author Contributions

Y.W. conceptualized the study, designed the template system, and developed the core components. Claude Code assisted with text refinement, content organization, and feature documentation. Y.W. supervised and validated all aspects of the final manuscript.

Acknowledgments

This research was funded by ...

Declaration of Interests

The authors declare that they have no competing interests.

Inclusion and Diversity Statement

We support inclusive, diverse, and equitable conduct of research.

Declaration of Generative AI in Scientific Writing

The authors utilized Claude Code, provided by Anthropic, for enhancing manuscript content organization, text conciseness, and formatting consistency. The authors reviewed and verified all AI-generated content. Responsibility for the final manuscript content rests entirely with the authors.

Tables

Parameter	Value	Unit	Notes
Resolution	300	dpi	Standard for print
Width	2000	pixels	Recommended for figures
Height	1332	pixels	Maintains aspect ratio
File size	48.5	KB	After optimization
Compression ratio	92.3	%	Size reduction achieved

Table 3 – Test table demonstrating image optimization parameters. This table shows the key parameters used for optimizing figures in the SciTex workflow.

Figures

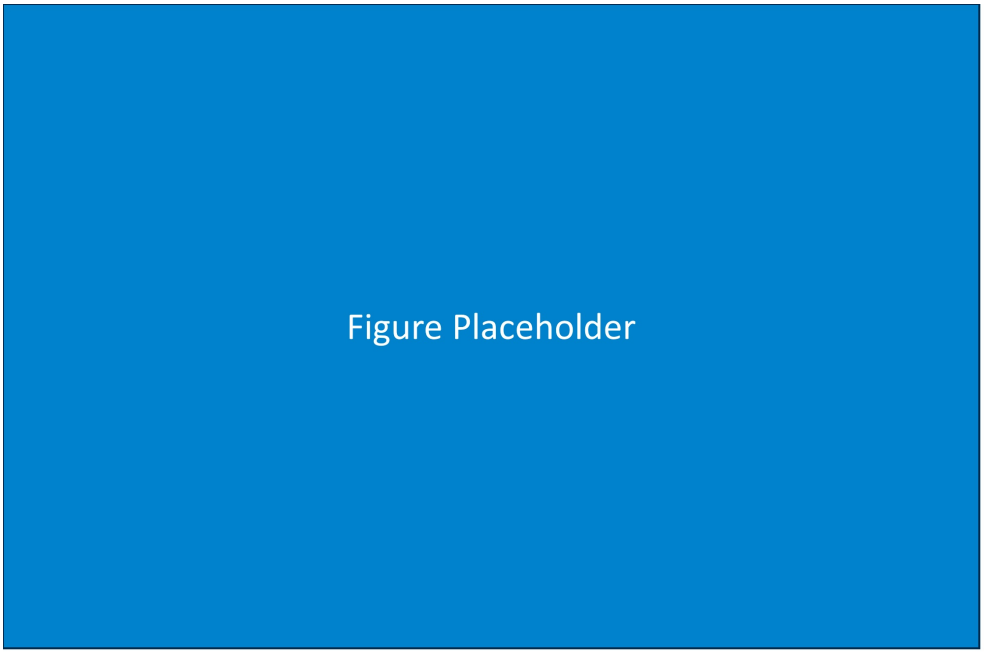

Figure 1 – Figure 00

Description for figure 00.

Figure 2 – Figure 01

Description for figure 01.