Software Specifications Requirements

for

Solution to prevent copyright infringement/piracy/plagiarism of NCERT text books

Prepared by:

Shantanu Mishra
2100290120152

KIET Group of Institutions
Delhi-NCR, Ghaziabad

March 05, 2024

Revision History

Name	Date	Reason For Changes	Version

1. Introduction

1.1 Purpose:

This document outlines the requirements for the development of a plagiarism detection tool to enhance academic integrity. It specifies the features and functionalities that will enable the tool to analyze paper characteristics and identify potential acts of plagiarism in academic publications.

1.2 Document Conventions:

The following conventions are used in this document:

- Font: Times New Roman, size 12.
- Priority: Priorities assigned to higher-level requirements are assumed to be inherited by detailed requirements.

1.3 Intended Audience and Reading Suggestions

Intended Audience:

- Developers
- Project Managers
- Marketing Staff
- Users
- Testers
- Documentation Writers

Reading Suggestions:

Readers are advised to begin with the overview sections and proceed to sections most pertinent to their role, ensuring a comprehensive understanding of the plagiarism detection tool's requirements.

1.4 Product Scope

The plagiarism detection tool is designed to analyze paper characteristics, including watermarks, paper quality, text formatting, and physical dimensions, to determine the originality of academic publications. The tool aims to address existing limitations in detecting copied tables, figures, or formulae across diverse academic disciplines.

References

- Teddi Fishman's comprehensive definition of plagiarism
- Any additional documents related to the vision and scope of the plagiarism detection tool.

Overall Description

Product Perspective

The software tool for identifying pirated NCERT textbooks operates as a standalone application designed to assist users in detecting unauthorized reproductions and plagiarism. It integrates with scanning functionalities to analyze paper characteristics, text formatting, and other identifiable features of NCERT textbooks. The tool aims to provide a comprehensive solution for verifying the authenticity of educational content across various publishers.

Product Function

The main functions of the software tool include:

- Scanning NCERT textbooks from different publishers using various methods such as mobile phones, scanners, or inputting files.

- Analyzing paper characteristics such as watermarks, paper quality, and dimensions to differentiate between original and pirated publications.

- Analyzing text formatting for inconsistencies and comparing formatting against known NCERT standards.

- Comparing the number of pages in scanned textbooks against known NCERT textbooks to detect deviations indicative of unauthorized reproductions.

- Verifying the authenticity of scanned textbooks based on the analysis results and providing clear indications of the likelihood of piracy.

User Classes and Characteristics

The user classes for the software tool include:

- Administrators: Responsible for managing the tool, updating databases, and configuring settings. Administrators possess advanced knowledge of the tool's functionalities and have access to administrative controls.

- Users: Individuals or organizations utilizing the tool to scan and verify NCERT textbooks. Users may have varying levels of technical expertise and may use the tool for specific purposes such as academic research or educational content validation.

Operating Environment

The software tool operates in a diverse range of environments, including:

- Mobile Devices: Compatible with smartphones and tablets running iOS or Android operating systems.

- Computers: Compatible with desktop and laptop computers running Windows, macOS, or Linux operating systems.

- Internet Connectivity: Requires internet connectivity for accessing updated databases, performing online verification checks, and receiving software updates.

Design and Implementation Constraints

The design and implementation of the software tool are subject to the following constraints:

- Compatibility: The tool must be compatible with a wide range of devices, operating systems, and scanning technologies to ensure accessibility and usability.

- Performance: The tool should perform scanning and analysis efficiently, minimizing processing time and resource usage.

- Security: The tool must implement robust security measures to protect scanned data and ensure confidentiality.

- Scalability: The tool should be designed to accommodate future enhancements and updates, allowing for scalability and flexibility.

User Documentation

The software tool comes with comprehensive user documentation, including:

- User Manuals: Detailed instructions on how to use the tool, including scanning procedures, analysis techniques, and result interpretation.

- FAQs: Frequently Asked Questions to address common queries and troubleshooting issues.

- Tutorials: Step-by-step guides and video tutorials demonstrating key functionalities and best practices for using the tool effectively.

Assumptions and Dependencies

The software tool operates under the following assumptions and dependencies:

- Availability of Scanning Devices: Users have access to scanning devices such as mobile phones or scanners to capture images or input scanned files.

- Access to Internet Connectivity: Users have access to internet connectivity for accessing updated databases, performing online verification checks, and receiving software updates.

- Compliance with Copyright Laws: Users are expected to comply with copyright laws and regulations when using the tool to verify the authenticity of NCERT textbooks.

External Interface Requirements

User Interfaces

The user interface of the software tool for identifying pirated NCERT textbooks shall include the following components:

Scan Options:

- The scan options interface shall provide users with choices for scanning NCERT textbooks using various methods such as:

  - Mobile Phones: Users can utilize the built-in camera functionalities of their mobile phones to capture images of textbooks.

  - Scanners: Users can connect scanners to their computers for scanning physical textbooks directly.

  - File Upload: Users can upload scanned files saved in formats such as PDF or image files.

- Each scan option shall be accompanied by clear instructions and prompts to guide users through the scanning process.

Results Display:

- Upon completion of scanning and analysis, the results display interface shall present the findings regarding the authenticity of the scanned textbook.

- The results shall be presented in a user-friendly format, utilizing clear and concise language to convey the analysis outcomes.

- Visual indicators such as color-coded labels or icons may be used to denote the likelihood of piracy or the presence of discrepancies.

Settings Management:

- The settings management interface shall allow users to configure various parameters and preferences related to the scanning and analysis process.

- Users can access settings to update databases, configure scanning parameters (e.g., resolution, file format), and customize result display options.

- The interface shall provide options for saving user preferences and settings for future use.

Hardware Interfaces

The software tool shall interface with various hardware components to facilitate scanning and analysis functionalities:

Mobile Phones:

- The tool shall utilize the camera functionalities of mobile phones to capture images of NCERT textbooks for scanning purposes.

- Compatibility with smartphones running iOS or Android operating systems shall be ensured.

- The interface with mobile phones shall support features such as autofocus, image stabilization, and flash control to optimize image quality.

Scanners:

- The tool shall interface with scanners connected to computers to scan physical NCERT textbooks directly.

- Compatibility with a wide range of scanner models and manufacturers shall be ensured, including flatbed and document scanners.

- The interface with scanners shall support functionalities such as color depth adjustment, scanning resolution selection, and image format conversion.

File Upload Devices:

- The tool shall support inputting scanned files saved in formats such as PDF, JPEG, PNG, or TIFF.

- Users can upload scanned files from external storage devices such as USB drives, SD cards, or cloud storage platforms.

- The interface with file upload devices shall ensure compatibility with common file formats and storage mediums.

Software Interfaces

The software tool shall integrate with various software components and APIs to enhance its functionality:

Scanning APIs:

- The tool may interface with scanning APIs or libraries to access advanced scanning functionalities.

- Integration with scanning APIs shall enable features such as image preprocessing, text recognition, and metadata extraction.

- Compatibility with popular scanning SDKs (Software Development Kits) and platforms shall be considered for seamless integration.

Comparison Algorithms:

- The tool shall interface with databases or comparison algorithms to compare scanned textbooks against known NCERT standards.

- Integration with comparison algorithms shall facilitate the detection of discrepancies in paper characteristics, text formatting, and other features.

- Compatibility with established comparison algorithms and standards shall be ensured for accurate analysis.

Authentication Systems:

- The tool may interface with authentication systems for user login and access control functionalities.

- Integration with authentication systems shall enable features such as user authentication, role-based access control, and session management.

- Compatibility with common authentication protocols (e.g., OAuth, LDAP) and frameworks shall be considered for interoperability.

Communications Interfaces

The software tool may require communication interfaces for exchanging data and notifications:

Messaging Systems:

- The tool may interface with messaging systems for real-time communication between users and administrators.

- Integration with messaging systems shall support features such as instant messaging, notifications, and alerts.

- Compatibility with popular messaging protocols (e.g., XMPP, MQTT) and platforms shall be considered for seamless communication.

Email Services:

- The tool may interface with email services for sending notifications, reminders, and communication updates.

- Integration with email services shall support features such as email composition, address book integration, and attachment handling.

- Compatibility with standard email protocols (e.g., SMTP, IMAP) and providers shall be ensured for reliable communication.

Notification Systems:

- The tool may interface with notification systems for delivering real-time updates and alerts to users.

- Integration with notification systems shall support features such as push notifications, in-app messaging, and notification preferences management.

- Compatibility with popular notification protocols (e.g., Firebase Cloud Messaging, Apple Push Notification Service) and platforms shall be considered for effective communication.

System Features

Scan Functionality

Mobile Phone Scanning

- Users can utilize the built-in camera functionalities of their mobile phones to capture images of NCERT textbooks.

- The scanning interface shall provide options for adjusting camera settings such as resolution, focus, and flash.

- After capturing the image, users can preview and confirm the scan before proceeding with analysis.

Scanner Integration

- The tool shall interface with scanners connected to computers to scan physical NCERT textbooks directly.

- Users can select scanning parameters such as color depth, resolution, and file format through the scanner interface.

- Scanned images shall be processed and analyzed automatically upon completion of the scanning process.

File Upload

- Users can upload scanned files saved in formats such as PDF, JPEG, PNG, or TIFF.

- The file upload interface shall support drag-and-drop functionality for ease of use.

- Uploaded files shall be processed and analyzed in real-time to verify the authenticity of the scanned textbook.

Analysis of Paper Characteristics

Watermark Detection

- The tool shall analyze paper characteristics to detect watermarks embedded in NCERT textbooks.

- Advanced image processing techniques shall be employed to enhance watermark visibility and accuracy.

- Watermark detection results shall be presented to users as part of the analysis report.

Paper Quality Analysis

- The tool shall assess paper quality based on factors such as texture, weight, and color consistency.

- Algorithms for image analysis and pattern recognition shall be utilized to identify variations in paper quality.

- Paper quality analysis results shall be included in the analysis report for user reference.

Dimension Comparison

- The tool shall compare the dimensions of the scanned textbook against known NCERT standards.

- Measurements such as length, width, and thickness shall be analyzed to identify deviations.

- Dimension comparison results shall be presented as part of the analysis report for user evaluation.

Text Formatting Analysis

Font and Typography Recognition

- The tool shall analyze text formatting features such as font type, size, and style.

- Optical character recognition (OCR) algorithms shall be utilized to extract text from scanned images.

- Text formatting analysis shall include comparisons against known NCERT standards to detect discrepancies.

Spacing and Alignment Detection

- The tool shall assess text spacing, alignment, and layout consistency across pages.

- Algorithms for text layout analysis shall be employed to identify irregularities in spacing and alignment.

- Spacing and alignment detection results shall be included in the analysis report for user review.

Style and Formatting Verification

- The tool shall verify text styles and formatting elements such as headings, paragraphs, and bullet points.

- Style and formatting verification shall ensure adherence to NCERT guidelines and standards.

- Detected deviations in style and formatting shall be highlighted in the analysis report for user awareness.

Comparison of Number of Pages

Page Count Analysis

- The tool shall compare the number of pages in the scanned textbook against known NCERT textbooks.

- Automated page counting algorithms shall be utilized to accurately determine the total number of pages.

- Page count analysis results shall be presented in the analysis report for user validation.

Pagination Verification

- The tool shall verify pagination consistency and sequence integrity within the scanned textbook.

- Algorithms for page numbering analysis shall be employed to identify missing or duplicate pages.

- Pagination verification results shall be included in the analysis report for user verification.

Authenticity Verification

Plagiarism Detection

- The tool shall compare the analyzed textbook against a database of known NCERT textbooks to detect plagiarism.

- Advanced similarity detection algorithms shall be utilized to identify plagiarized content and unauthorized reproductions.

- Plagiarism detection results shall be presented in the analysis report along with recommendations for further action.

Authenticity Score Calculation

- The tool shall calculate an authenticity score based on the analysis results of paper characteristics, text formatting, and other factors.

- The authenticity score shall indicate the likelihood of piracy or the presence of discrepancies in the scanned textbook.

- Authenticity score interpretation guidelines shall be provided to users for understanding and decision-making.

Report Generation

- The tool shall generate a comprehensive analysis report summarizing the findings of the authenticity verification process.

- The analysis report shall include detailed information on paper characteristics, text formatting, page count, plagiarism detection, and authenticity score.

- Users can access and download the analysis report for documentation and reference purposes.

Other Nonfunctional Requirements

Performance Requirements

Scanning Performance

- The tool shall be capable of processing scanned textbooks within a reasonable time frame, with a maximum processing time of 30 seconds per textbook.

- Scanning performance shall be optimized to handle high volumes of scanning requests simultaneously without degradation in speed or efficiency.

Analysis Speed

- The analysis process shall be completed promptly, with results available to users within 1 minute after scanning completion.

- Performance optimizations such as parallel processing and caching shall be implemented to expedite the analysis speed.

Scalability

- The tool shall be scalable to accommodate a growing user base and increasing workload demands.

- Scalability shall be achieved through modular design, cloud-based infrastructure, and efficient resource allocation strategies.

Response Time

- The user interface shall respond to user inputs promptly, with an average response time of less than 500 milliseconds.

- Response time optimizations such as asynchronous processing and client-side caching shall be implemented to minimize latency.

Safety Requirements

Data Integrity

- The tool shall ensure data integrity by implementing safeguards against data corruption, manipulation, or loss.

- Measures such as data encryption, checksum verification, and transaction logging shall be employed to maintain data integrity.

Error Handling

- The tool shall provide robust error handling mechanisms to detect, report, and recover from errors gracefully.

- Error messages shall be informative and user-friendly, guiding users on how to resolve issues encountered during scanning or analysis.

User Privacy

- User privacy shall be safeguarded by adhering to data protection regulations and best practices.

- Personally identifiable information (PII) shall be handled securely, with strict access controls and encryption protocols in place.

Backup and Recovery

- The tool shall support automated backup and recovery mechanisms to prevent data loss in the event of system failures or disasters.

- Regular backups shall be performed and stored in secure off-site locations for disaster recovery purposes.

Emergency Shutdown

- The tool shall include an emergency shutdown feature to halt all scanning and analysis processes in case of emergencies or critical issues.

- The emergency shutdown procedure shall be accessible to administrators and authorized personnel for immediate action.

Security Requirements

Authentication and Authorization

- The tool shall enforce user authentication and authorization mechanisms to control access to sensitive functionalities and data.

- Role-based access control (RBAC) shall be implemented to assign specific permissions and privileges to users based on their roles.

Data Encryption

- All sensitive data, including scanned textbooks and user credentials, shall be encrypted during transmission and storage.

- Encryption algorithms such as AES (Advanced Encryption Standard) shall be used to ensure data confidentiality and integrity.

Vulnerability Management

- The tool shall undergo regular security assessments and vulnerability scans to identify and address potential security vulnerabilities.

- Patch management procedures shall be established to apply security updates and fixes in a timely manner.

Audit Logging

- All user interactions and system activities shall be logged for audit trail purposes.

- Audit logs shall include details such as user actions, timestamps, and IP addresses for traceability and accountability.

Secure Communication

- Secure communication protocols such as HTTPS shall be used to encrypt data transmitted between clients and servers.

- SSL/TLS (Secure Sockets Layer/Transport Layer Security) certificates shall be employed to establish secure connections and prevent eavesdropping.

Software Quality Attributes

Reliability

- The tool shall demonstrate high reliability, with minimal downtime and system failures.

- Robust error handling, fault tolerance mechanisms, and redundant infrastructure shall be implemented to ensure system reliability.

Maintainability

- The tool shall be designed for ease of maintenance and future enhancements.

- Modular architecture, clean code practices, and comprehensive documentation shall facilitate ongoing maintenance activities.

Usability

- The user interface shall be intuitive and user-friendly, catering to users with varying levels of technical expertise.

- Usability testing shall be conducted to gather user feedback and improve interface design for enhanced usability.

Performance Efficiency

- The tool shall optimize resource utilization and minimize resource consumption to achieve efficient performance.

- Performance profiling and tuning shall be performed to identify and address performance bottlenecks.

Interoperability

- The tool shall support interoperability with external systems and services through standardized interfaces and protocols.

- Compatibility testing shall be conducted to ensure seamless integration with third-party components and platforms.

Business Rules

Compliance with Copyright Laws

- Users are expected to comply with copyright laws and regulations when using the tool to verify the authenticity of NCERT textbooks.

- The tool shall include reminders and warnings regarding copyright infringement to promote ethical usage practices.

Authorized Use

- The tool shall only be used for legitimate purposes such as educational content validation and research.

- Unauthorized use of the tool for malicious activities or copyright violations shall be strictly prohibited.

Data Confidentiality

- Users shall adhere to strict confidentiality agreements regarding the handling and dissemination of scanned data and analysis results.

- Confidential information such as user credentials and sensitive data shall be protected from unauthorized access.

Other Requirements

Legal and Regulatory Compliance

Copyright Compliance

- The software tool shall comply with copyright laws and regulations governing the use and reproduction of educational materials, including NCERT textbooks.

- Users shall be required to acknowledge and agree to abide by copyright restrictions before using the tool.

Data Privacy Compliance

- The tool shall adhere to data privacy regulations such as GDPR (General Data Protection Regulation) and CCPA (California Consumer Privacy Act).

- User consent shall be obtained for the collection, processing, and storage of personal data, with transparent disclosure of data usage practices.

Accessibility Requirements

User Accessibility

- The user interface of the software tool shall be designed to be accessible to users with disabilities, including those with visual or motor impairments.

- Accessibility features such as screen reader compatibility, keyboard navigation, and alternative text for images shall be implemented.

Language Support

- The tool shall support multiple languages to accommodate users from diverse linguistic backgrounds.

- Language localization options shall be provided to allow users to switch between preferred languages for interface display and communication.

Performance Optimization

Resource Efficiency

- The software tool shall optimize resource utilization to minimize memory usage, CPU load, and bandwidth consumption.

- Resource-efficient algorithms and data structures shall be employed to maximize performance while conserving system resources.

Load Balancing

- Load balancing mechanisms shall be implemented to distribute workload evenly across servers and prevent performance bottlenecks.

- Automatic scaling and dynamic resource allocation shall be utilized to handle fluctuations in user demand effectively.

Documentation Requirements

User Manual

- A comprehensive user manual shall be provided to guide users on how to use the software tool effectively.

- The user manual shall include step-by-step instructions, troubleshooting tips, and best practices for optimal usage.

Technical Documentation

- Technical documentation shall be prepared for developers and system administrators to understand the software architecture, implementation details, and configuration instructions.

- Documentation shall be regularly updated to reflect changes and enhancements to the software tool.

Glossary

-Analysis Report: A document generated by the software tool summarizing the findings of the authenticity verification process, including details on paper characteristics, text formatting, page count, plagiarism detection, and authenticity score.

-Authentication: The process of verifying the identity of users accessing the software tool, typically through username/password authentication or other authentication mechanisms such as biometrics or multi-factor authentication.

-**Authorization:** The process of granting or denying users access to specific functionalities or data within the software tool based on their roles and permissions.

-**Plagiarism Detection:** The process of identifying plagiarized content and unauthorized reproductions within scanned textbooks by comparing them against known NCERT textbooks and standards.

-**Scalability:** The ability of the software tool to handle increasing workload demands and accommodate a growing user base without sacrificing performance or reliability.

-User Interface (UI): The graphical interface through which users interact with the software tool, including features such as scan options, results display, and settings management.

-Watermark: A recognizable pattern or image embedded in paper to identify the authenticity or origin of a document, often used as a security measure in printed materials such as NCERT textbooks.