

Business Design Requirement (BDR)

Project: BRI Research Paper Plagiarism Detection Platform

1. Project Overview

Bambhari Research Institute (BRI) aims to build a **research integrity and plagiarism detection platform**, inspired by tools like Grammarly (desktop reference), but focused on **research papers, technical documents, and academic content**.

The platform will allow **students, working professionals, and researchers** to upload research papers. The system will automatically analyze the content to **detect plagiarism, similarity percentage, and originality score**.

The goal is to **protect intellectual property, ensure originality, and improve research quality**.

2. Project Motive (Why This Project Exists)

Primary Motive

- Ensure **original research submission** at BRI
- Prevent **copyright violations and unethical copying**
- Educate researchers on **proper citation and originality**

Secondary Motive

- Build trust with journals, universities, and companies
 - Create a **BRI-owned plagiarism verification ecosystem**
 - Provide researchers with **self-check tools before final submission**
-

3. Target Users

User Type	Purpose
Students	Validate academic projects & papers
Working Professionals	Ensure compliance before publishing
Researchers	Check originality before journal submission
BRI Admin	Review, approve, and audit submissions

4. Core Problem Statement

Currently, many research papers:

- Are partially or fully copied
- Lack proper citations
- Violate copyright laws unknowingly

Manual checking is:

- Time-consuming
- Inconsistent
- Not scalable

Solution: An automated, accurate plagiarism detection platform.

5. Functional Requirements

5.1 User Module

- User registration & login
- Upload research papers (PDF, DOCX)
- View plagiarism report
- Download detailed analysis

5.2 Upload & Processing Module

- Accept document upload
- Extract text from document
- Preprocess text (remove stop words, normalize)

5.3 Plagiarism Detection Engine

The system should:

- Compare uploaded content with:
 - Public research papers
 - Journals & articles
 - Internet sources
 - BRI internal research database

Accuracy Rule (Critical Requirement)

- **0-12% similarity** → Original / Acceptable
- **18-20% similarity** → Warning (Needs review)
- **Above 20% similarity** → Flag as *Plagiarized*

This threshold must be clearly visible to users and admins.

6. Plagiarism Logic (How the System Works)

Step-by-Step Flow

1. User uploads research paper
 2. System extracts text
 3. Content split into chunks (sentences / paragraphs)
 4. Each chunk compared using:
 - Text similarity algorithms
 - N-gram matching
 - Semantic similarity (ML-based)
 5. Matched sources identified
 6. Similarity percentage calculated
 7. Final plagiarism score generated
-

7. Output Report Structure

Each plagiarism report should show:

- Overall similarity percentage
 - Originality score
 - Matched sources with links
 - Highlighted copied text
 - Section-wise similarity breakdown
 - Final verdict:
 -  Original
 -  Needs Review
 -  Plagiarized
-

8. Admin & Review Panel

Admin capabilities:

- View all submissions
- Review flagged papers

- Approve / Reject research papers
 - Export reports
 - Maintain internal research database
-

9. Non-Functional Requirements

- High accuracy & consistency
 - Secure document storage
 - Fast processing (under 2–3 minutes per paper)
 - Scalable for large datasets
 - Data privacy & confidentiality
-

10. Technology Direction (High-Level)

Frontend

- Web-based UI (clean, Grammarly-like experience)
- Upload progress & real-time status

Backend

- Text extraction services
- Plagiarism detection APIs / ML models
- Scoring & threshold logic

Database

- Online Surfing
 - Google Research Paper
 - Multiple Research Preview
-

11. Success Metrics

The project is successful if:

- Users clearly understand plagiarism percentage
 - False positives are minimized
 - Research originality improves over time
 - Admin review effort is reduced
-

12. Key Takeaway for the Team

This project is not just a website.

It is a **research integrity platform** that:

- Protects originality
- Educates researchers
- Builds credibility for BRI

Every feature should answer one question:

"Does this help ensure genuine, original research?"

13. Next Steps for Development Team

1. Finalize plagiarism threshold logic
 2. Design user flow (Upload → Analyze → Report)
 3. Build MVP with sample documents
 4. Validate accuracy with real research papers
 5. Improve ML models iteratively
-

End of BDR Document

Important Notice: Redistribution of this document is strictly prohibited.