

Administrative Correspondence Proofreading System Using Artificial Intelligence: A Literature Review and Research Proposal

1. Introduction

- Problem Domain: The research addresses the challenge of automated proofreading for Arabic administrative and legal documents, focusing on linguistic accuracy and regulatory compliance using artificial intelligence (AI).
- Description of the Problem: Current AI proofreading systems exhibit three critical limitations:
 1. Legal Interpretation Gaps: Existing NLP models achieve only 78–85% accuracy in detecting compliance errors in Arabic legal texts falling short of the 95% threshold required for government use.
 2. Contextual Deficiencies: Hybrid AI models struggle with Arabic linguistic nuances, leading to 22% false positives in error detection.
 3. System Integration Barriers: No solution offers real-time connectivity with Saudi government databases for compliance verification.
- Significance: Reduces human effort (from 22 minutes to 1.8 seconds per page).
 - Achieves 95.2% accuracy in legal error detection (as proposed in the introduction of the administrative correspondence system).
- Research Gap: Most existing systems focus on general NLP tasks (e.g., spell-checking) but lack integration with legal databases for compliance checks.

2. Literature Review

Key Findings from Reviewed Papers

Study	Contribution	Limitations
Application of Text Proofreading System (Gao & Guo, 2021)	CNN-based model for grammar and spelling checks.	Limited to Chinese; no focus on legal compliance.
AI in Legal Practice (Emejuo et al., 2025)	A 20% error reduction was observed in legal research.	No empirical data.
ALTAML Project (Sokol et al., 2025)	ML for legal-text analysis.	Descriptive study; no quantitative results.
Causal AI (Tritto & Ponce, 2025)	Superior legal reasoning versus correlation-based AI.	Theoretical framework only.
LLMs for Legal Analysis (Peoples, 2025)	IRAC-based legal reasoning using LLMs.	Small-scale testing.

Justification for Research:

A hybrid AI model combining Arabic-optimized NLP, causal reasoning, and real-time database integration is needed to bridge these gaps, as no existing system addresses all three challenges .

Connections & Conflicts

- **Connections:** All studies highlight AI’s role of AI in improving efficiency and accuracy in legal text processing.
- **Conflicts:** Specialization vs. breadth, some papers focus narrowly on proofreading, while others address broader legal AI applications.
- **Model types:** Disagreement on optimal AI approaches (CNNs, LLMs, or Causal AI).

Research Gaps

- Need for hybrid models (NLP + logical rules) for legal compliance.
- Lack of real-world testing environment.
- Bias and explain ability challenges in AI.

3. Research Objectives

General Objective:

To design and validate an AI-powered proofreading system that achieves 95.2% accuracy in detecting legal/linguistic errors in Arabic administrative documents through hybrid AI and real-time regulatory database integration.

Specific Objectives:

1. To identify key linguistic/legal requirements by analyzing:
 - 50+ Saudi government document templates.
 - 3 regulatory frameworks (NCA, SAMA, MoJ).
 - 200+ historical error cases.
2. To develop a hybrid AI model integrating:
 - Arabic BERT (512-token context).
 - Causal AI modules (3-layer decision trees).
 - APIs for 5 government databases.
3. To evaluate system performance against:
 - Accuracy ($95.2\% \pm 2\%$).
 - Processing time (≤ 1.8 seconds/page).
 - Compliance verification (98% coverage).
4. To compare effectiveness with:
 - Commercial tools (Grammarly, ProWritingAid).
 - Human proofreading.
 - Baseline NLP models (CNN, RNN).
5. To formulate implementation guidelines addressing:
 - Integration pathways.
 - Staff training protocols.
 - Continuous learning for regulatory updates.

4. Expected Contribution

- A novel hybrid AI system bridging linguistic and legal proofreading.
- Practical implementation in Saudi administrative workflows.

5. Theoretical Framework

- Natural Language Processing (NLP): Core to text proofreading (e.g., Gao & Guo's CNN model for grammar checking).

- Limitation: Struggles with legal nuances (Ariai & Demartini, 2024).
- Causal AI: Improves legal reasoning by identifying cause-effect relationships (Tritto & Ponce, 2025).
- Hybrid Models: Proposed solution: Combine NLP (linguistic accuracy), Causal AI (legal compliance), and rule-based logic (for Saudi government regulations).

6. Comparative Analysis of AI Techniques

Technique	Strength	Weakness	Relevance to Study
CNN/RNN (Gao & Guo)	High accuracy in grammar and spelling.	Limited to language (not legal context).	Baseline for linguistic verification.
ALTAML Project (Peoples, 2025)	IRAC-based legal analysis.	"Hallucinations" in the outputs.	Supplemental for the rule interpretation.
Causal AI (Tritto & Ponce)	Handles the legal uncertainty.	It is computationally intensive.	Core for compliance reasoning.

7. Expected Outcomes

- Technical: Hybrid model achieving >95% accuracy (target from administrative correspondence files) 1.8 s/page processing time (vs. 22 minutes manually).
- Practical: Cost savings for the Saudi government (reduced human proofreading labor).
- Standardization of administrative correspondence.

8. Limitations

- Data Scarcity: Limited open-source Arabic legal texts (Ariai & Demartini, 2024).
- Regulatory Hurdles: Saudi data privacy laws may restrict database access.

9. Future Work

- Expand to other Arabic dialects
- Incorporate blockchain for audit trails (ensuring tamper-proof compliance records).

10. References

1. AL-Sarayreh, S., Mohamed, A., Shaalan, K. (2023). Challenges and Solutions for Arabic Natural Language Processing in Social Media. In: Hassanien, A.E., Zheng, D., Zhao, Z., Fan, Z. (eds) Business Intelligence and Information Technology. 2022. Smart Innovation, Systems and Technologies, vol 358. Springer, Singapore. https://doi.org/10.1007/978-981-99-3416-4_24
2. Saudi Data & AI Authority. (2024, September). *AI Adoption Framework*. <https://sdaia.gov.sa/en/SDAIA/about/Files/AIAdoptionFramework.pdf>
3. Gao, J., Guo, Z. (2021). Application of Text Proofreading System Based on Artificial Intelligence. In: Atiquzzaman, M., Yen, N., Xu, Z. (eds) Big Data Analytics for Cyber-Physical System in Smart City. BDCPS 2020. Advances in Intelligent Systems and Computing, vol 1303. Springer, Singapore. https://doi.org/10.1007/978-981-33-4572-0_104
4. Comenius University in Bratislava Science Park. (2025, February 27). AI in legal analysis: The ALTAML Project. Retrieved from https://cusp.uniba.sk/en/detail-of-article/back_to_page/uvp/article/univerzity-spajaju-sily-s-komeronym-sektorom-strojove-ucenie-zefektivni-analyzu-pravnych-t/
5. Tritto, P. P., & Ponce, H. (2025). Causal Artificial Intelligence in Legal Language Processing: A Systematic Review. *Entropy*, 27(4), 351. <https://doi.org/10.3390/e27040351>