

An Explainable Natural Language Processing Framework for Semantic Analysis and Knowledge Retrieval from Islamic and Feminist Text Collections

Shatha Haroush Salih Al Fahdawi¹, Sohaib Abbas Oudah Al-Kubaisi² and
Younus Abdulhadi Khaleel Al Fayyadh^{3*}

¹Researcher, College of Education for Humanities, University of Anbar, Anbar, Iraq

²Professor, College of Education for Humanities, University of Anbar, Anbar, Iraq

^{3*}Faculty Member, United Arab Emirates University, Al Ain, UAE

E-mail: ¹sha23h4005@uoanbar.edu.iq, ²ed.sahib.abbas@uoanbar.edu.iq, ³yalfayyadh@uaeu.ac.ae

ORCID: ¹<https://orcid.org/0009-0005-2495-3089>, ²<https://orcid.org/0000-0002-3912-0662>,

³<https://orcid.org/0000-0002-2868-0998>

(Received 01 June 2026; Revised 10 July 2026, Accepted 20 August 2026; Available online 29 September 2026)

Abstract - As religious, cultural, and social text collections become increasingly digitized, sophisticated computational techniques are needed to extract meaningful information without losing transparency, context, or interpretation. Digital Islamic and feminist text collections have complex linguistic elements, historical contexts, multiple interpretations, and socially sensitive concepts that cannot be adequately processed using traditional keyword-based retrieval systems and non-explainable NLP approaches. The paper proposes an explainable NLP conceptual framework that combines semantic analysis, explainability mechanisms, and knowledge retrieval approaches for culturally sensitive texts. The study uses a qualitative conceptual approach that involves the analysis of 21 academic articles on topics such as NLP, Explainable Artificial Intelligence (XAI), semantic models, information retrieval, Islamic text processing, feminist discourse analysis, and ethical AI. There are five main conceptual themes identified in the qualitative analysis. The designed model is built on five interrelated layers: text collection and processing, natural language processing methods, semantic analysis and representation, explainability techniques, and knowledge extraction and application. In the review, such vital problems like context vagueness, language differences, multi-meaning, implicitness, and biases in analyzing sensitive texts are underlined. Besides that, the conceptual analysis of the feminist text analysis issues shows some major thematic areas like cultural and historical difference, multiple theoretical standpoints, implicitness, evolution of the vocabulary, and NLP bias. These thematic areas are qualitative aspects derived from the analyzed literature and not any empirical statistical data. The result is an understandable and person-oriented Explainable NLP model that ensures reliable knowledge discovery and provides the basis for further digital humanities research.

Keywords: Explainable Artificial Intelligence, Natural Language Processing, Semantic Analysis, Knowledge Retrieval, Islamic Text Analysis, Feminist Text Analysis, Digital Humanities

I. INTRODUCTION

The rapid growth of such areas as Natural Language Processing (NLP) and Artificial Intelligence (AI) opens new horizons for the analysis of complex cultural, religious and social textual resources, but the problem of making such technologies transparent and interpretable is a key point in such sensitive areas. XAI approach aims at creating transparent artificial intelligence tools for improving interpretation and the degree of users' trust to computerized analysis (Arrieta et al., 2020). Large language models and advanced NLP tools reveal their abilities in language comprehension, but the presence of hidden biases, low interpretability and knowledge representation problems proves that it is necessary to use responsible and explainable approaches to text analysis (Bender et al., 2021). In Islamic text analysis, the latest scientific research demonstrates the use of transformer-based techniques in explainable semantic search of the Qur'an. It proves the importance of interpretability in religious concepts and meaning extraction from Arabic texts (Mustafa et al., 2024). Knowledge graphs methodologies show the efficiency of the semantic enhancement technique in the organization and retrieval of Islamic textual knowledge from Hadith collections (Kamran et al., 2026). Nevertheless, linguistic barriers, problems in translation and interpretation are important limitations for using NLP in the area of Islam studies (Al-Shidi et al., 2025).

Likewise, text mining of feminist texts necessitates the use of computational tools that can take into account the dynamic nature of socio-cultural constructs, gender stories, and culturally diverse views. Studies focusing on gender bias in information retrieval systems have underscored the need to devise unbiased text retrieval systems that avoid bias in text analysis (Seyedsalehi et al., 2025). Text analysis using NLP techniques has proven to be effective in recognizing gender representation patterns and socio-cultural themes from large textual data sources for understanding cultural stories (Lucy

et al., 2020). More sophisticated language modeling and semantic analysis techniques have been used to detect socio-cultural content in texts, underlining the need for contextual reasoning and explainable decision-making in text-based applications (Pan et al., 2024). In addition, socio-cultural text analysis through computational means has also revealed the complexity of the socio-cultural and ideological structure of language (Nguyen et al., 2020).

Text mining approaches have found their use in the field of digital social discourse analysis, illustrating the benefits of automatic knowledge extraction regarding human behavior and public discourses (Turenne, 2023). Moreover, machine learning approaches to gender violence narratives have shown the capabilities of natural language processing techniques in uncovering social patterns, along with the necessity of responsible computational approaches (Low et al., 2024). Thus, it is important to create an Explainable NLP approach including semantic analysis, interpretability components, and knowledge extraction methods for Islamic and feminist text collections.

Problem Statement

Due to the swift increase in the number of digital archives with Islamic and feminist texts, in addition to other religious, cultural, and social texts, there is an increasing need for more sophisticated computational models that can enable extraction of knowledge from complex textual sources. There are different linguistic features, historical background, interpretive dimensions, and socially sensitive aspects in Islamic and feminist texts which cannot be captured through the use of traditional keyword-based searches or other text-mining techniques. Current NLP technologies can help us extract semantic relationships within textual data; but many advanced NLP algorithms are based on black-box systems, which means that these technologies lack the required level of explainability. In addition, most knowledge retrieval models do not consider contextual meanings, ideological perspectives, and interpretations associated with these texts. Thus, there is a need for an Explainable NLP technology that combines semantic analysis and knowledge retrieval techniques in a way that enables analysis of Islamic and feminist text collections in an explainable manner.

Research Gap

Despite considerable successes in the field of NLP, semantic analysis, and information retrieval, the approaches currently used are not adequate enough to be applied to text collections that require cultural and social sensitivity, such as Islamic and feminist literature. While traditional text-mining methods are based on superficial language patterns, deep learning NLP models offer better semantic understanding but lack transparency and interpretability. Current XAI approaches concentrate mostly on classification and prediction problems and do not pay much attention to the explanation of semantics and knowledge extraction from humanities-focused textual materials. Also, existing research hardly ever takes into account the integration of contextual semantic modeling,

explainability, and knowledge retrieval processes into a single concept needed to analyze multiple interpretations, historical background, and concepts embedded in Islamic and feminist texts. Hence, there is a gap in researching the theoretical foundation of an explainable NLP approach to be used for complex text analysis.

Research Objective

1. To explore NLP applications for semantic analysis and knowledge mining of Islamic and feminist literature.
2. To investigate explainable AI technologies and how improve the interpretability of NLP-powered textual analysis.
3. To develop a theoretical model of an Explainable NLP approach integrating semantic representation, explainability mechanisms, and knowledge retrieval methods for culturally specific text collections.
4. To discuss the challenges and ethical considerations of using NLP and Explainable AI, and future prospects in digital humanities and textual research.

Research Hypothesis

H1: Explainable Natural Language Processing techniques can enhance the transparency and interpretability of semantic analysis in Islamic and feminist text collections.

H2: Context-aware semantic analysis approaches can improve the accuracy of knowledge identification and retrieval from complex and culturally sensitive textual resources.

H3: Integration of explainability mechanisms with NLP-based knowledge retrieval can support more reliable and human-understandable interpretation of Islamic and feminist textual concepts.

H4: A unified Explainable NLP framework integrating semantic modeling, interpretability approaches, and knowledge retrieval methods can provide an effective conceptual approach for analyzing diverse digital text collections.

Key Contribution

- As part of this research study, an integrated conceptual model has been presented which incorporates NLP, semantic analysis, explainable AI models, and knowledge extraction approaches for the analysis of Islamic and feminist text databases.
- In the current study, the importance of contextual semantic modeling and explanation approaches has been discussed for improving the reliability, interpretability, and transparency of the text analysis process using NLP techniques.

- The current research study focuses on the importance of contextual knowledge extraction approaches for deriving valuable information from multicultural text databases in light of linguistic and socio-cultural aspects.
- The study lays the groundwork for the use of explainable natural language processing in humanities-based studies.

The rest of this paper is structured as follows. In section II, the literature review is presented through the discussion of existing literature about Natural Language Processing, Explainable Artificial Intelligence, semantic analysis, knowledge retrieval, Islamic text processing, and feminist text analysis, with an emphasis on their weaknesses. In section III, the research methodology is provided in accordance with the methodological approach of qualitative conceptual analysis through the discussion of research design, data sources, and thematic analysis process for building the conceptual framework. Section IV describes the conceptual Explainable NLP framework that proposes to apply to Islamic and feminist texts through the discussion of the integration of text preparation, NLP techniques, semantic analysis, explainability techniques, and knowledge retrieval processes. Section V is dedicated to the qualitative discussion of the importance of explainability in sensitive text analysis, semantic issues in Islamic and feminist text collections, and the ethical aspect in applying responsible NLP. The theoretical and practical significance of the proposed framework for the research in digital humanities, information science, and knowledge management is discussed in section VI. Finally, section VII summarizes the results, contributes the key features of the conceptual framework, and offers the future direction for research.

II. LITERATURE REVIEW

Advances in NLP techniques in recent times have revealed the ability of computational techniques to aid in semantic understanding, knowledge discovery, and context-driven interpretation of large sets of texts. Human-centered and explainable computation techniques have become relevant in the analysis of complex textual domains where meaning is contingent upon cultural, social, and historical context. The computational grounded theory approaches have revealed the power of combining automatic text analysis with qualitative interpretation to achieve a deep insight into social-cultural phenomenon in the contextually relevant way (Nelson, 2020). Further, the topic modeling and sentiment analysis methods have proved the efficacy of their use in discovering the patterns of themes and public perception in the context of online discussions related to Islam (González-Baquero et al., 2023). Domain-specific NLP techniques, including named entity recognition, have been useful in improving the identification of concepts, entities, and relationships in the domain of Islamic texts (Tarmizi & Saad, 2025).

Semantic analysis methodologies have become popular for investigating elaborate social narratives, especially in areas

dealing with gender issues, cultural representations, and identity. Semantic analysis of media texts has demonstrated how sentiment analysis and text mining could be employed to investigate public attitudes and perspectives about the social role and representation of women (Ghobain et al., 2026). Additional evidence for the discovery of linguistic structures and patterns that are not easily detected by traditional keyword search can also be derived from linguistic pattern mining methods (Sorato et al., 2020). The use of machine learning in humanistic text analysis should be carefully considered since computational analysis should always go hand-in-hand with interpretation to prevent misrepresentation of complex ideas (Pääkkönen & Ylikoski, 2021).

In the context of feminist text analysis, automatic methods based on NLP have been proposed for detecting gender-based discrimination, sexism, and discriminatory language use in social media and digital communication channels. Such methods illustrate the need for developing explainable models that can differentiate between various types of gender-related content (Rodríguez-Sánchez et al., 2020). Moreover, ethical aspects in intelligent sentiment analysis show the necessity of building transparent, ethical, and responsible systems while analyzing socially sensitive content and human-made narratives (Noor, 2024). Approaches related to computational modeling of digital information systems have also emphasized the significance of reliable knowledge management and secure information processing and analysis to facilitate information retrieval (Trisnawati et al., 2025). Moreover, studies devoted to social harm have shown the significance of integrating theoretical insights and computational analysis of human experience and social narrative (Roy & Sharma, 2025).

Though there have been considerable advances made so far in semantic analysis of text through NLP in areas such as sentiment analysis, entity recognition, and knowledge extraction, there are still many shortcomings that exist in creating an explainable and contextual framework for such text collections that hold cultural significance. While the current efforts tend to be task-oriented (classification, sentiment detection, retrieval), not much effort has been put into creating an integrated approach towards semantic analysis, explainability, and knowledge retrieval. It is thus necessary to create an Explainable Natural Language Processing Framework to help analyze Islamic and feminist text collections.

III. RESEARCH METHODOLOGY: QUALITATIVE CONCEPTUAL ANALYSIS

3.1 Research Design

The research employs the qualitative conceptual research methodology to investigate the relevance of Explainable NLP models for conducting semantic analysis and knowledge extraction from Islamic and feminist texts. Rather than creating and testing a computational model, this study should focus on examining current NLP models, explainability

methods, semantic modeling techniques, and knowledge extraction methods through a conceptual approach. Methodology of the research is grounded on views on natural language processing, XAI, digital humanities, and information science that enable the analysis of ways to help computational approaches support transparency when dealing with sensitive cultural and social texts. Qualitative research methodology gives an opportunity to conduct a deeper analysis of the contextual meaning, differences in languages, interpretation requirements, and ethics connected with analysis of Islamic and feminist texts. Conceptual framework is developed based on associations between existing theories and aspects of text processing, semantic understanding, explainability, and knowledge discovery.

3.2 Data Sources for Conceptual Analysis

This research concept is formulated through secondary data sources found in the existing literature pertaining to NLP, Explainable AI, semantic analysis, information retrieval, and digital humanities. This research considers the research articles, review papers, and conceptual papers related to text processing of religious texts, discourse analysis of social texts, text mining concerning gender issues, and explainable computational methods. The relevant literature includes research papers related to Islamic text analysis, specifically involving Qur'anic semantic search, Hadith knowledge representation, and domain-specific NLP, as well as research papers related to feminist text analysis, gender bias detection, social narrative mining, and ethically-based AI interpretation. These literature sources offer a theoretical base for determining appropriate NLP, explainability, and knowledge retrieval methods for text analysis of complex text collections.

3.3 Procedure of Qualitative Analysis

The qualitative analysis procedure is undertaken using a systematic approach for thematic analysis to find important concepts, relationships, and methodology related to Explainable NLP-based text analysis. The process of analysis involves three major steps, starting with literature identification and selection, where studies that are relevant to Explainable NLP, XAI, semantic modeling, information retrieval, Islamic text processing, and feminist discourse analysis are thoroughly analyzed. Selected studies are analyzed considering the goals, computational methods, explainability techniques used, and their domains to understand how relevant they are to culturally sensitive text analysis. In the second step, thematic categorization and concepts are extracted through thematic organization of identified information. These include semantic understanding, explainability, knowledge retrieval, and domain adaptation.

In this phase, the process explores how various NLP techniques help in context-based meaning extraction and how

the explainability methods help in increasing transparency and interpretability during computational analysis. Finally, in the last phase, the themes and their relationships are synthesized to come up with the Explainable NLP model by combining text preparation, NLP techniques, semantic analysis, explainability methods, and knowledge retrieval procedures with the context of the domain and human expertise. This qualitative process helps the research understand the concept of how explainable NLP can assist in contextually meaningful knowledge discovery from Islamic and feminist texts.

Additionally, based on thematic synthesis of the reviewed literature, a conceptual distribution model was developed to illustrate the relative prominence of major semantic challenges identified in feminist text analysis. The represented categories are qualitative thematic observations derived from the frequency and emphasis of discussions across the selected studies. They should not be interpreted as precise statistical measurements or empirical proportions, but rather as illustrative conceptual weights that support understanding of the major challenges discussed in existing literature.

IV. CONCEPTUAL EXPLAINABLE NLP FRAMEWORK

As depicted in fig. 1 below, the Explainable NLP framework for semantic analysis and knowledge retrieval utilizes Islamic and feminist texts as the input data set. Particularly, the approach begins by identifying the relevant textual material from a specific domain. Such textual material may comprise of religious texts, scholarly work, historical information, feminist texts, and discourse information. Such text material will have to undergo several processes of data pre-processing. This process entails text cleaning, language normalization, metadata creation, and domain-specific data pre-processing. The NLP techniques component of the framework incorporates linguistic processing, contextual language modeling, topic modeling, named entity recognition, and discourse analysis techniques. The concept of semantic analysis is performed using concept extraction, relationship extraction, thematic analysis, and contextual analysis. To obtain better transparency, the explainability layer performs feature importance analysis, attention-based explanation, rule-based explanation, and human-explainable explanation of the computational processes. Human expertise and domain knowledge from the perspective of Islamic scholars, feminists, and information science professionals are used throughout the entire knowledge discovery process for a proper interpretation of the findings. Thus, the output of the knowledge discovery process is an interpretable knowledge discovery process.

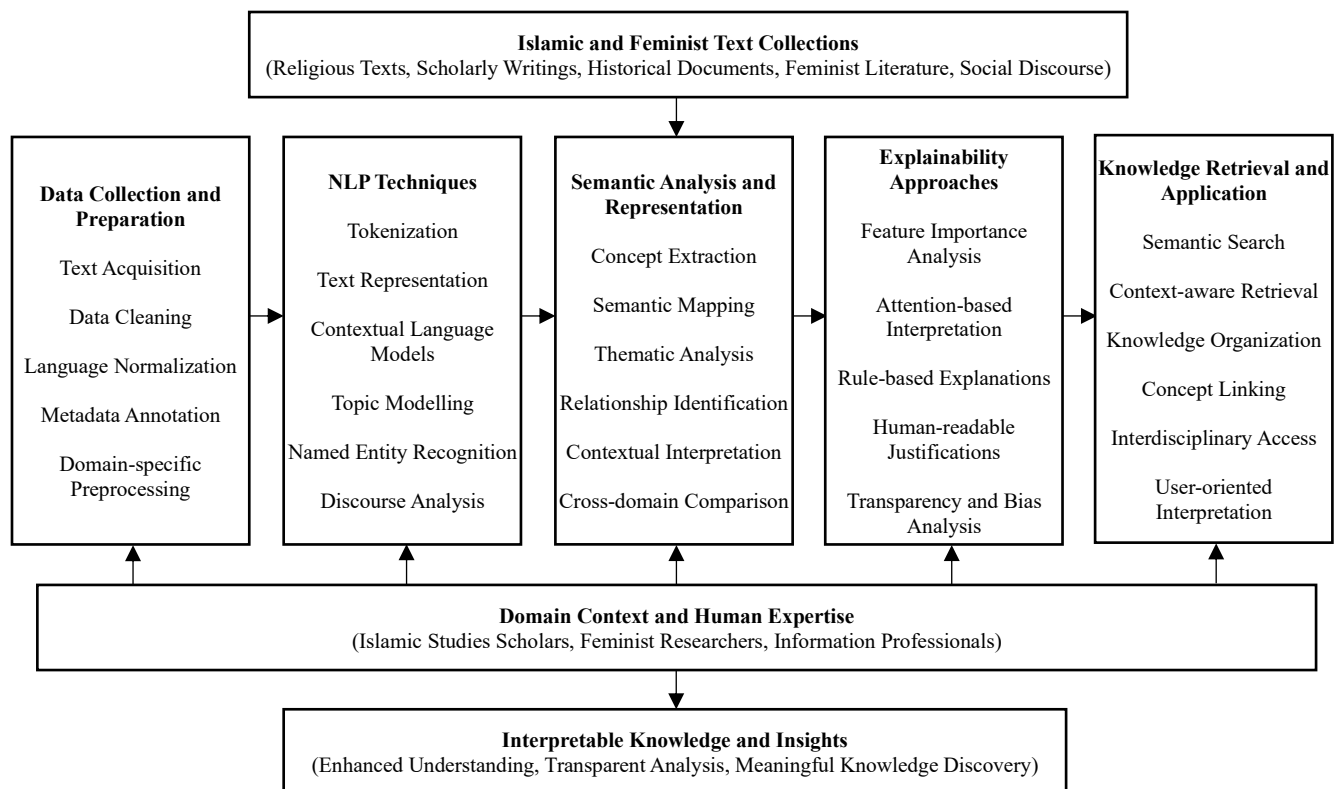


Fig. 1 Overall Architecture for Explainable NLP Framework

V. QUALITATIVE DISCUSSION

5.1 Hypothesis Assessment and Framework Validation

The proposed Explainable NLP framework was assessed conceptually according to the formulated research hypotheses. Hypothesis H1 is verified since the framework applies explainable approaches like feature importance assessment, attention mechanisms, and human-understandable explanations to enhance transparency and explainability in semantic analysis. Hypothesis H2 is resolved by implementing approaches to context-aware semantic analysis that include semantic representation, contextual language models, thematic analysis, and relationship identification, which allow identifying and retrieving better knowledge from culturally sensitive text sources. Hypothesis H3 is proven through applying explainability techniques in combination with knowledge retrieval to help users comprehend how the extracted concepts and knowledge were acquired. Hypothesis H4 is verified by implementing a unified architectural model that incorporates layers of text pre-processing, NLP approaches, semantic analysis, explainability techniques, and knowledge retrieval into a single conceptual model. In this way, the proposed model provides theoretical justification for all research hypotheses through a transparent, contextual, and user-centric approach to analyzing text collections of an Islamic and feminist nature.

5.2 Role of Explainability in Sensitive Text Analysis

Explainability becomes a crucial element in using Natural Language Processing (NLP) approaches on sensitive textual material, in which the outputs need to be carefully interpreted and validated by humans. In contrast to traditional text analysis, in cases of Islamic and feminist text corpora, there exist complex meanings, cultural context, and even several interpretations. Thus, having only the results of retrieval and classification is not enough for such tasks; it is necessary to understand the process of achieving the particular result. Most researchers describe "interpretability aims at understanding and explaining the internal workings of a model in human-understandable terms".

In the context of analysis of Islamic text, explainability makes it possible for researchers to investigate reasons why a certain verse, concept, or interpretation by scholars belongs to a certain semantic category. For instance, when an NLP system makes connections between concepts like ethics, justice, and social responsibility, an explanation mechanism makes it possible for researchers to understand linguistic cues that led to a certain outcome. The same case applies to feminist text analysis, which needs transparent interpretation due to the different nature of concepts that are culturally and theoretically based.

Explainable NLP also helps minimize the problem of algorithmic bias in the processing of sensitive texts. According to Arrieta et al. (2020), "explainable artificial intelligence seeks to render artificial intelligence decisions comprehensible to human beings" (Arrieta et al., 2020). It

becomes especially relevant when processing texts that relate to religious views, social identity, and ideology. For example, in sentiment analysis, a researcher might wrongly interpret a conversation regarding the issue of women's rights or religious practices because of the lack of context. Through feature importance analysis, attention-based explanation, and comprehensible reasoning, explainable NLP allows researchers to verify results and conduct ethical knowledge discovery.

5.3 Semantic Challenges in Islamic Text Collection

The Islamic text corpora are an example of very complex semantic domains that have certain historical, linguistic, and interpretive qualities. The Islamic text corpora contain the Quran, Hadith, classical scholars' work, and modern interpretations, where the meaning of the concept varies based on linguistic construction, history, and scholarship. The traditional NLP methods that rely on word matching usually miss these complex semantic relations to concentrate on direct word occurrences.

For instance, words concerning justice, morality, kinship, and social duty could occur in various Islamic texts with varied meanings. A search engine that focuses on particular keywords can pick up such content but does not know the actual theological or historical meaning of the word. According to the definition, "information retrieval is concerned with finding material that satisfies an information need from large collections." In the context of religion, however, meeting an information need involves more than just retrieval; it involves interpretation.

Translation differences and multilingualism are the additional issues that are faced in the field. Islamic documents are written in classical Arabic, and when scholars study them, they are actually referring to the translation of the document in other languages, such as English and Urdu. Translation differences may cause semantic differences and impact the computer interpretation. For example, one single word in Arabic could be translated differently based on the views of scientists and language context. Explanatory semantic analysis is capable of resolving this issue as it identifies the relationship between concepts and contexts.

Additionally, it is essential that Islamic texts be distinguished carefully between original religious materials, past practices, and personal interpretations. Automation tools lacking context can unwittingly merge divergent viewpoints or provide false associations. Hence, it becomes important to use explainable NLP techniques in addition to domain knowledge.

5.4 Semantic Challenges in Feminist Text Collections

It is evident that the study of feminist text collections requires addressing the semantic problems inherent to them. This is mainly because of the variety of theoretical approaches and terminology, as well as the social character of semantics. Feminist text collections encompass academic work, political texts, activists' writing, literature, and personal experience accounts. In such texts, the same concept may be understood differently depending on the context.

For instance, notions such as "equality," "empowerment," "representation," "gender roles," among others, might be defined differently depending on whether one is talking about liberal feminism, radical feminism, intersectional feminism, or some other feminist ideology. The traditional approach to text retrieval likely finds all these words and won't notice in what context were mentioned. As Bender et al. (2021) argued, "language models are not just learning language; they are learning patterns from the data on which they are trained" (Bender et al., 2021). This means that NLP technologies should account for possible biases and constraints of training data.

An additional task that needs to be addressed is the detection of hidden feminist ideas, which cannot be stated explicitly via particular terminology. This means that literary texts from the past might describe social inequality, discrimination, and gender expectations without applying any feminist terms. Likewise, literary stories about women may convey some meanings that need to be interpreted and understood in context. This is where explainable NLP techniques could help.

Additionally, texts that pertain to feminism tend to represent different cultures and societal contexts, thus making their interpretation difficult and not universal. A concept that relates to the idea of gender equality would mean different things in different geographical locations and times. It is for this reason that semantic analysis, along with explainability, can be used to make sure that researchers are able to analyze the results provided by computations while being sensitive to culture and theory.

Fig. 2 presents a conceptual representation of the major semantic challenges identified in feminist text analysis literature. The figure illustrates the relative prominence of themes such as cultural and historical context variations, theoretical diversity, implicit meanings, evolving terminology, and interpretation biases. These categories are derived from qualitative thematic synthesis of the reviewed studies and are presented as illustrative conceptual weights rather than statistically validated proportions.

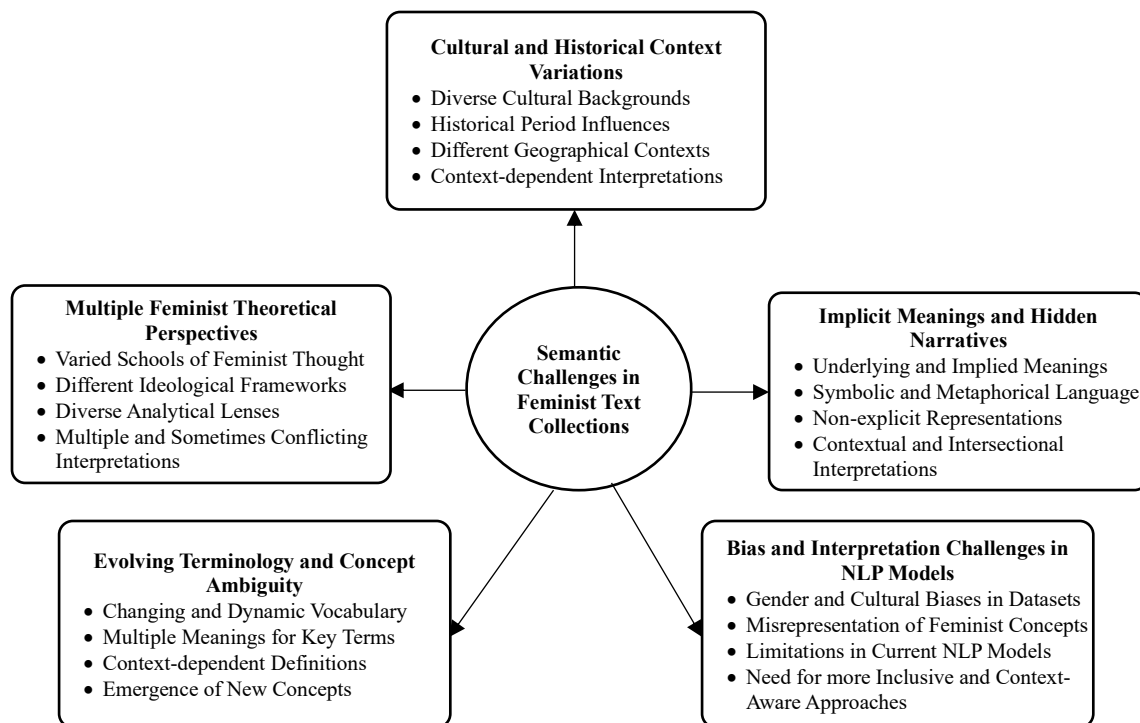


Fig. 2 Conceptual Distribution of Semantic Challenges in Feminist Text Collections

5.5 Ethical Considerations

Using Explainable NLP to analyze Islamic and feminist text corpora raises ethical concerns about cultural sensitivity, bias, privacy, and proper interpretation. Given that such textual resources have religious beliefs, identities, histories, and different ideological perspectives, any form of automated analysis should not create misinterpretations or reinforce biases. In that sense, since NLP technologies may have inherent biases depending on their training datasets, it is important to ensure transparency with regard to how certain concepts, relationships, and interpretations are derived. It is also crucial to involve specialists of the corresponding domains – for instance, Islamic scholars, feminists, and information scientists – to verify the computational results and guarantee that knowledge extracted is contextually and culturally appropriate.

VI. RESEARCH IMPLICATIONS

6.1 Theoretical Implications

This current paper contributes to the theoretical framework regarding the application of NLP, Explainable Artificial Intelligence, semantic analysis, and knowledge retrieval in the field of digital humanities. The theory extends the previous understanding of NLP through interpretation and contextualization of texts, especially those which are socially and culturally sensitive in nature. Whereas most computational approaches tend to focus on information retrieval or information classification, the present theory highlights the link between semantic representation, explainability, and interpretability. This study serves as a basis for future research in the field of explainable

computational approaches in religious studies and feminist discourse analysis, among others.

6.2 Practical Implications

The above-mentioned conceptual model presents practical advice for academics, information specialists, and digital humanities experts conducting research related to Islamic and feminist text corpora. The application of both semantic and explainable approaches helps to achieve better transparency in knowledge generation, higher contextualization, and proper investigation of complex text corpora. This approach provides researchers with an opportunity to identify critical concepts, explore thematic relations, and analyze cultural/social stories being aware of computational algorithms. Furthermore, the conceptual model can be helpful in developing digital repositories, search engines, and knowledge management systems based on NLP and dealing with sensitive texts.

VII. CONCLUSION

This study aims to establish a conceptual understanding of the role explainable natural language processing (NLP) approaches can play in semantic analysis and knowledge retrieval of text collections in Islam and feminism domains. The current research has been addressing the drawbacks of traditional approaches to NLP, especially their failure to accommodate the needs of contextual meaning identification, interpretation diversity, and transparency. A qualitative conceptual research method has been used based on the analysis of 21 scholarly articles dealing with NLP, XAI, semantic modeling, information retrieval, processing of Islamic texts, feminist discourse analysis, and ethics of AI. 5 major conceptual themes have been derived from thematic

analysis: semantic understanding, explainability, knowledge retrieval, domain adaptation, and ethical issues. Based on these results, an Explainable NLP framework with 5 interrelated layers has been designed: text collection and preparation, NLP methods, semantic analysis and representation, explainability methods, knowledge retrieval and application. Qualitative review highlighted the following challenges for Islamic text analysis – linguistic difficulties, multilingual challenges, contextual meaning, and differences in academic views, while feminist text analysis revealed several conceptual challenges, including variations in cultural and historical context, diversity of theoretical perspectives, implicit meanings, evolving terminology, and potential NLP interpretation biases. The results confirm the necessity of using explainability and contextual semantic modeling together with expert knowledge for efficient knowledge discovery based on sensitive texts. Further studies should be aimed at the implementation of the proposed methodology through the use of actual Islamic and feminist data, development of explainable language models tailored to the specific area, and empirical validation with expert evaluation involved.

Declaration

Funding Statement: This research received no specific grant, financial support, or funding from any public, commercial, or not-for-profit organization.

Conflict of Interest: The authors declare that there is no conflict of interest associated with this research, its conceptual framework, or its publication.

REFERENCES

- [1] Al-Shidi, A., Nordin, F. H. B., Mohamed, I. I., Younis, M. A., Elmutasim, I. E., & Abusham, E. E. (2025). A literature review on natural language processing techniques for Qur'anic studies: challenges and insights. *Frontiers in Signal Processing*, 5, 1535166. <https://doi.org/10.3389/frsip.2025.1535166>
- [2] Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Benetot, A., Tabik, S., Barbado, A., ... & Herrera, F. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information fusion*, 58, 82-115. <https://doi.org/10.1016/j.inffus.2019.12.012>
- [3] Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021, March). On the dangers of stochastic parrots: Can language models be too big?. In *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency* (pp. 610-623). <https://doi.org/10.1145/3442188.3445922>
- [4] Ghobain, E., Al-Nofaie, H., Alhazmi, F., Bosli, R., & Shamakhi, M. (2026). A Triangulated Digital Approach to News Sentiment Analysis: Insights from Media Coverage of Saudi Women Enlistment in Military Forces. *Journalism and Media*, 7(1), 50. <https://doi.org/10.3390/journalmedia7010050>
- [5] González-Baquero, W., Amores, J. J., & Arcila-Calderón, C. (2023). The conversation around islam on twitter: Topic modeling and sentiment analysis of tweets about the muslim community in spain since 2015. *Religions*, 14(6), 724. <https://doi.org/10.3390/rel14060724>
- [6] Kamran, A. B., Butt, N. A., & Basharat, A. (2026). Semantic Enrichment of Hadith Corpus—Knowledge Graph Generation from Islamic Text. *Semantic Web*, 17(2), 22104968261431425. <https://doi.org/10.1177/22104968261431425>
- [7] Low, H. Q., Keikhosrokiani, P., & Pourya Asl, M. (2024). Decoding violence against women: analysing harassment in middle eastern literature with machine learning and sentiment analysis. *Humanities and Social Sciences Communications*, 11(1), 497. <https://doi.org/10.1057/s41599-024-02908-7>
- [8] Lucy, L., Demszky, D., Bromley, P., & Jurafsky, D. (2020). Content analysis of textbooks via natural language processing: Findings on gender, race, and ethnicity in Texas US history textbooks. *AERA Open*, 6(3), 2332858420940312. <https://doi.org/10.1177/2332858420940312>
- [9] Mustafa, A. M., Nakhleh, S., Irsheidat, R., & Alruosan, R. (2024). Interpreting Arabic transformer models: A study on XAI interpretability for Qur'anic semantic-search models. *Jordanian Journal of Computers and Information Technology*, 10(4), 352-366. <https://doi.org/10.5455/jjcit.71-1704878720>
- [10] Nelson, L. K. (2020). Computational grounded theory: A methodological framework. *Sociological methods & research*, 49(1), 3-42. <https://doi.org/10.1177/0049124117729703>
- [11] Nguyen, D., Liakata, M., DeDeo, S., Eisenstein, J., Mimno, D., Tromble, R., & Winters, J. (2020). How we do things with words: Analyzing text as social and cultural data. *Frontiers in Artificial Intelligence*, 3, 62. <https://doi.org/10.3389/frai.2020.00062>
- [12] Noor, M. F. (2024). A feminist interpretation of Qur'anic texts by Amina Wadud and Zaitunah Subhan. *INTIHA: Islamic Education Journal*, 1(3), 144-156. <https://doi.org/10.58988/intiha.v1i3.311>
- [13] Pääkkönen, J., & Ylikoski, P. (2021). Humanistic interpretation and machine learning. *Synthese*, 199(1), 1461-1497. <https://doi.org/10.1007/s11229-020-02806-w>
- [14] Pan, R., García-Díaz, J. A., & Valencia-García, R. (2024). Comparing fine-tuning, zero and few-shot strategies with large language models in hate speech detection in english. *CMES-Computer Modeling in Engineering and Sciences*, 140(3), 2849-2868. <https://doi.org/10.32604/cmcs.2024.049631>
- [15] Rodríguez-Sánchez, F., Carrillo-de-Albornoz, J., & Plaza, L. (2020). Automatic classification of sexism in social networks: An empirical study on twitter data. *IEEE Access*, 8, 219563-219576. <https://doi.org/10.1109/ACCESS.2020.3042604>
- [16] Roy, L., & Sharma, A. (2025). Integrated theory of sexual offending framework for harmful sexual behavior in autism and intellectual disability. *Journal of Intellectual Disabilities and Offending Behaviour*, 16(2), 60-70. <https://doi.org/10.47059/jidob/V16/I2/7>
- [17] Seyed-salehi, S., Bigdeli, A., Arabzadeh, N., AlMousawi, B., Marshall, Z., Zihayat, M., & Bagheri, E. (2025). Understanding and mitigating gender bias in information retrieval systems. *Foundations and Trends® in Information Retrieval*, 19(3), 191-364. <https://doi.org/10.1561/1500000103>
- [18] Sorato, D., Goularte, F. B., & Fileto, R. (2020). Short semantic patterns: A linguistic pattern mining approach for content analysis applied to hate speech. *International Journal on Artificial Intelligence Tools*, 29(02), 2040002. <https://doi.org/10.1142/S0218213020400023>
- [19] Tarmizi, S. A., & Saad, S. (2025). Named Entity Recognition on Islamic Texts: A Systematic Review. *International Journal of Advanced Science Computing and Engineering*, 7(3), 98-111. <https://doi.org/10.62527/ijasce.7.3.265>
- [20] Trisnawati, L., Samsudin, N. A. B., Bin Ahmad Khalid, S. K., Bin Ahmad Shaubari, E. F., & Indra, Z. (2025). An Ensemble Semantic Text Representation with Ontology and Query Expansion for Enhanced Indonesian Quranic Information Retrieval. *International Journal of Advanced Computer Science & Applications*, 16(1), 481. <https://doi.org/10.14569/ijacsa.2025.0160148>
- [21] Turenne, N. (2023). Net activism and whistleblowing on YouTube: a text mining analysis. *Multimedia Tools and Applications*, 82(6), 9201-9221. <https://doi.org/10.1007/s11042-022-13777-0>