

**Applying Natural Language Processing to Advance Research in Hanbali
Semantic and Contextual Analysis of Fiqh Disagreement**

Mofaq Mastour Alfai

Contents

Introduction.....	3
Literature Review	5
(i) The Challenges of Jurisprudential Differences	5
(ii) The Challenge of Arabic Natural Language Processing.....	8
Methodology	11
(i) Data Collection and Preprocessing	11
(ii) Semantic and Contextual Analysis	14
Timeline	15
References.....	16

Introduction

Fiqh is an Arabic term for ‘Islamic jurisprudence.’ In literal terms, *fiqh* means to understand and to comprehend. Although shari’ah is often used to refer to Islamic law, shari’ah is more general and inclusive of both spiritual beliefs and actions as revealed in the Quran and Sunnah. Still, *fiqh* refers to understanding such law as interpreted and derived by legal experts (Sungkar, 2024). *Fiqh* is based on the primary sources: The Holy Qur’an and the Sunnah of the Prophet, and secondary sources: *Ijma’* (scholarly consensus) and *Qiyas* (analogical reasoning). This discipline guides worship, buying and selling, family issues, and criminal law. It has evolved over centuries since Muslim jurists developed methodologies to understand the Qur’an and Sunnah, resolve differences of interpretation, and deal with new situations that occur with time (Faizi and Ali, 2024). Consequently, different schools of thought appeared, out of which four schools of thought—Hanafi, Maliki, Shafi’i, and Hanbali—are most prominent. Each of these schools of *fiqh* has its unique principles of understanding and interpreting the Islamic law.

Among these schools, the Hanbali school is traced back to the jurisprudential principles of Imam Ahmad Ibn Hanbal (d. 241 AH). The principles of the Hanbali school are known for their strong adherence to the Scripture (The Qur’an), Prophetic traditions (*Ahadiths*), fatwas of companions (*athaar*), or reports of the righteous predecessors (*salaf*). The key texts that expound Hanbali Fiqh—*Al-Umdah*, *Al-Muqni’*, *Al-Kafi*, and *Al-Mughni* are penned by a distinguished jurist, Ibn Qudāmah (Ahmad and Ijaz, 2022). Among these texts, *Al-Mughni* is more comprehensive since it is based on the previous texts, expounding and comparing the opinions of jurists of other schools. It evaluates their agreements and disagreements while assessing their rationales and juristic merits. The author expresses his preferred opinion conclusively.

Natural Language Processing (NLP) aims to extract information from a large amount of data, analyze it, and understand it. As the texts of Fiqh are written in classical Arabic, they require a deep understanding of grammar, context, and rationales. This powerful tool has also been applied to Islamic Jurisprudence (Fiqh). NLP helps to identify differences of opinions among jurists, point out the source of specific rulings, and comprehend specific terminologies used by them over time. It determines the explicit and implicit meanings of jurisprudential expressions and highlights linguistic variations in those texts.

Similarly, to understand the linguistic features of Arabic and legal texts, the language models AraBERT and AraLegal-BERT were trained on millions of Arabic texts. These models are perfectly embedded for name entity recognition (NER), question answering (QA), and organizing Arabic legal texts, which classify rulings in fatāwā (Antoun et al., 2020). However, these models are mostly trained on modern Arabic text, and are not much integrated into the Hanbali school's legal texts.

This study aims to investigate the differences in opinions of the four schools of thought—Hanafi, Maliki, Shafi'i, and Hanbali, primarily focusing on the Hanbali jurisprudence. The study integrates an NLP model to compare their rationales and inter-school arguments through the lens of the seminal text, *Al-Mughni*.

Objectives

- To highlight the main differences in the fiqh rulings of the four schools of thought, Hanafi, Maliki, Shafi'i, and Hanbali
- To classify the methodology and reasoning used by each school to support their rulings
- To understand the methodology and reasoning employed by the Hanbali School concerning these rulings outlined in *Al-Mughni*

- To integrate a Natural Language Processing (NLP) model into *Al-Mughni* to evaluate the rationale behind similarities and differences of the Hanbali School with other Jurists and highlight inter-school differences.

Questions

1. What are the key differences in Fiqh rulings among the Hanafi, Maliki, Shafi'i, and Hanbali jurisprudence?
2. How do these schools employ the primary and secondary sources to support their rulings in areas where they disagree, especially the Hanbali School?
3. What is the distinctive methodology of the Hanbali School to assess the disagreements in light of *Al-Mughni*?
4. How do the NLP techniques evaluate the seminal text, *Al-Mughni*, and how can we understand the Hanbali School's approach to these rulings, and differences with others?

Literature Review

(i) The Challenges of Jurisprudential Differences

The seminal Hanbali texts penned by Ibn Qudāmah, *Al-ʿUmdah*, *Al-Muqniʿ*, *Al-Kāfi*, and *Al-Mughnī* function as a legal corpus that represents the consolidation of Hanbali doctrines and the development of legal reasoning within the school.

The four legal texts show *Al-ʿUmdah* as their most basic and direct version. The authors created this text to serve as an introductory legal resource for new students of law. The work presents one definitive ruling for each matter without showing evidence or different opinions or disagreements within the school. The main goal of *Al-ʿUmdah* is to teach Imam Ahmad's *muʿtamad* positions while avoiding complex reasoning and comparative analysis (Al-Saadi,

2022). The chapter on prayer (ṣalāh) in Al-‘Umdah contains only the conditions and integrals without any reference to textual sources.

The Al-Muqni‘ presents more complex material, which includes multiple Hanbali school opinions. The text presents two to three different perspectives about each issue because it draws from multiple transmissions of Ahmad ibn Hanbal and his student interpretations. The text fails to present evidence together with detailed juristic reasoning. The text Al-Muqni‘ presents two different opinions about the waiting period (‘iddah) of a pregnant widow, but it does not explain the supporting evidence for either position (Azzam, 2023).

The text of Al-Kāfi presents additional information about Hanbali legal proofs which include both hadiths and Qur’anic verses. The text combines the fundamental perspectives of the school by selecting the most authoritative (mu‘tamad) position while documenting alternative, weaker positions. The text explains the basis of legal decisions and demonstrates advanced analytical approaches. Al-Saadi (2022) explains in Al-Kāfi how Ahmad's students disagreed about zakāh for trade goods and supports the preferred view through Prophetic narrations and qiyās from other zakāh categories.

The works of Ibn Qudāmah rank Al-Mughnī as the most advanced and comprehensive. The work presents Hanbali legal opinions together with evaluations of the Sunni schools Hanafi, Maliki, and Shafi‘i. Ibn Qudāmah supports his legal preferences with textual evidence and logical arguments, and historical precedents for each position, while defending Hanbali positions through comparative reasoning (Saeed & Ijaz, 2022). The Hanbali position regarding prayer combination during travel receives a detailed presentation in Al-Mughnī, followed by Shafi‘i and

Maliki objections supported by hadiths and counterargument refutations through isnād analysis and contextual reasoning (Azzam, 2023).

The discussion about praying while traveling according to Al-Mughnī shows the Hanbali perspective, which allows it through authentic hadith reports from the Prophet (peace be upon him). The Shāfi‘ī school maintains more stringent requirements for prayer combination because it demands movement during both prayer times. The hadith from Ibn ‘Abbās states that the Messenger of Allah performed combined Ṣuḥr and ‘Aṣr prayers in Madinah without any concerns about rain or fear (Muslim, 705). According to Ibn Qudāmah the Shāfi‘ī position creates excessive limitations which contradict the straightforward interpretation of the text. The Hanbali position stands as the stronger view according to Azzam (2023) because it aligns with both textual interpretation and the practical needs of travelers.

The commercial law section of Al-Mughnī explores sharikat al-‘inān, which refers to a capital investment partnership. The Hanbali position, according to Ibn Qudāmah, permits unequal capital contributions and profit sharing through mutual agreement. The Hanafī position requires profit distribution to follow the exact ratio of capital investments unless the parties specify otherwise. The contract exists through mutual agreement according to Ibn Qudāmah, because the Qur’an and Sunnah lack explicit prohibition against unequal profit distribution. The verse determining the contracts: “O you who believe! Fulfill [all] contracts” (Qur’an 5:1) is about establishing permissibility unless there is explicit prohibition. Thus, the Hanafī thought is criticized by Ibn Qudāmah since it imposes a rational principle that could conflict with mutual consent. Consequently, they prioritize the text and will of the parties above abstract equity (Al-

Saadi, 2022). Through this example, Ibn Qudāmah demonstrates his method of analyzing both his school's ruling and others by using Qur'anic texts and legal maxims and rational critique.

The Hanbali ruling, according to Al-Mughnī, requires washing a dog-contaminated container seven times with one earth-based wash as the hadith states: “If a dog licks the utensil of one of you, let him wash it seven times, the first with dust” (Muslim, 279). The Shāfi‘īs share the same position regarding this matter, but the Mālikīs maintain a different stance by restricting impurity to dangerous dogs while disregarding the requirement of earth-based washing.

According to Ibn Qudāmah, the hadith contains general language that lacks specific qualifications about the dog's purpose. The methodological flaw of using contextual interpretation instead of explicit text becomes evident when the text remains clear according to Saeed & Ijaz (2022). Through his work, *Al-Mughnī* demonstrates his ability to analyze both legal decisions and their underlying knowledge bases.

(ii) The Challenge of Arabic Natural Language Processing

The current research and studies have utilized Natural Language Processing (NLP) for fatwa classification (Al-Khalifa et al., 2020), legal term semantic modeling (Zaghoulani & Charfi, 2021), and Islamic ruling knowledge graph construction (Al-Ajlan et al., 2022). For this purpose, digital corpora, including annotated fatwa databases, court rulings, or modern Arabic translations of classical texts, were used.

However, the vast majority of Islamic legal heritage remains in unstructured, non-digitized formats, for example, in classical Arabic, written in complex orthography, or preserved in handwritten form. This creates significant problems for NLP-based analysis (Yousef et al., 2023).

Similarly, the absence of extensive annotated datasets characterizes Arabic Fiqh in contrast to English legal texts. The process of manual annotation proves to be both time-consuming and demanding both linguistic and jurisprudential expertise, which restricts the use of supervised learning approaches (Zaghouani & Charfi, 2021).

Moreover, several classical texts are preserved as scanned images or poorly typeset PDFs. Although Optical Character Recognition (OCR) technologies have been integrated into the Arabic texts, handwritten or premodern print scripts, they are not as advanced as the Latin-based OCR (Sakr et al., 2022). Thus, in many cases, the accuracy plunges to only 70% which reduces the usability of such data for downstream NLP tasks.

The structure of jurisprudential texts differs from contemporary document organization. The discussion structure in these texts combines rulings with evidence and counterarguments and analogical reasoning without following a linear pattern. The current NLP pipelines struggle to parse these texts because they make it challenging to identify claim-evidence relationships and decision trees (Yousef et al., 2023).

Natural Language Processing (NLP) shows great potential when applied to classical jurisprudential texts such as *Al-Mughnī* by Ibn Qudāmah for both Islamic legal studies and computational linguistics. The systematic organization of Hanbali positions with opinions from other Sunni schools in *Al-Mughnī* makes it an optimal selection for corpus-based semantic analysis.

The initial requirement demands the transformation of *Al-Mughnī* into a machine-readable format. The process demands precise Arabic OCR processing of scanned images or PDFs, followed by manual or model-assisted correction according to Sakr et al. (2022). The text

of Al-Mughnī requires segmentation into chapters (*kitāb*) and sub-issues (*mas’alah*) based on legal themes such as prayer, purification, and trade.

The text undergoes tokenization and lemmatization through Farasa or CAMEL tools, transforming classical inflections into Modern Standard Arabic for wider tool compatibility (Taha et al., 2021). With the appropriate tools, Farasa or CAMEL, the text undergoes tokenization and lemmatization. This process converts classical inflections into Modern Standard Arabic for broader tool compatibility (Taha et al., 2021; Abdelali et al., 2021; Obeid, Khalifa & Habash, 2022).

Applying NLP to jurisprudential texts like Al-Mughnī requires identifying and extracting named entities and evidentiary references. The NLP model needs to be fine-tuned to detect the legal entities, such as the names of prominent jurists—Abū Ḥanīfa, Mālik, al-Shāfi‘ī, Aḥmad ibn Ḥanbal.

The model should recognize jurisprudential terminology, which includes *qiyās* (analogical reasoning) and *ijmā‘* (scholarly consensus) and *ḥadīth ṣaḥīḥ* (authentic Prophetic tradition), classification terms such as *mu’tamad* (relied-upon opinion) and *marjūḥ* (weaker view). In addition to identifying such entities, the model should also recognize scriptural references, notably citations of Qur’anic verses and Hadiths, often presented in various forms and chains of transmission.

Last but not least, the detection of argumentation markers, including “*wa qāla*” (and he said), “*dhākara*” (he mentioned), “*radda*” (he refuted), and “*dalīluhum*” (their evidence), is equally important because these markers are used to structure inter-school debates that support or challenge legal positions. The components play a crucial role in understanding the legal

reasoning structure both within the school and between madhāhib. This task's management becomes possible through a fine-tuned Arabic BERT-based model such as AraBERT or CAMELBERT, which includes custom entity recognition layers trained on annotated Fiqh literature samples (Khalifa & Belinkov, 2021; Hassan & El-Haj, 2022; AlKhamissi & Shaalan, 2023).

Methodology

The research centers on the Hanbali school of law through the analysis of Ibn Qudāmah's classical works *Al-Mughni*, *Al-Muqni'*, and *Al-Kāfi*, and additional scanned manuscripts and printed editions, and digital versions when available.

(i) Data Collection and Preprocessing

The NLP-based legal analysis will draw from many text categories, such as classical Hanbali legal texts that include canonical works by Ibn Qudāmah al-Maqdisī (d. 620 AH), such as *Al-Mughni*, *Al-Kāfi*, *Al-Muqni'*, and *Al-Umdah*. These texts present varying levels of legal abstraction, intra-madhhab differences, and inter-school comparisons.

The original texts of scanned manuscripts and early printed editions serve as essential resources for maintaining pre-modern scholars' orthographic and linguistic characteristics. The sources exist exclusively in archival libraries and online repositories such as Al-Maktaba al-Shamela and King Saud University Manuscript Collection. The inclusion of modern digital books in Unicode or XML formats will be considered when available. The machine-readable versions of these edited and annotated texts prepared by contemporary scholars decrease the time needed for preprocessing.

The NLP pipeline will directly incorporate digital texts that already exist in Unicode format. The sources exist in online repositories such as Al-Maktaba al-Shāmilah or Dār al-Manṭūrah and require only basic preprocessing for parsing. The texts contain defined chapter headings and standardized Arabic spelling and typographical diacritics, which improve morphological analysis.

The team will use Optical Character Recognition (OCR) to process non-machine-readable documents, including scanned PDFs and manuscripts, and early printed editions. The open-source Tesseract OCR engine with Arabic language capabilities will serve as the main tool for this project.

The recognition quality will be enhanced by using a Tesseract language model that was fine-tuned with classical Arabic script samples to handle pre-modern orthography and character shapes. Images will be converted to binary format, straightened and cleaned up before applying OCR to improve character recognition. Outputs will be reviewed manually or corrected using dictionary-based spell-checkers and n-gram language models to fix common OCR errors such as the confusion between “ى” and “ي” or diacritic misreadings.

The steps taken have not eliminated all challenges because Arabic script is cursive and diacritics are used inconsistently, and classical texts feature ornamental calligraphy. The challenges with OCR extraction require manual verification of specific sections before they can be corrected. The following stage of text preprocessing converts digital book or OCR raw textual data into a format that NLP can analyze syntactically and semantically. The preprocessing of Arabic text requires specialized tools and techniques because of its morphological complexity and inflectional characteristics.

The Arabic language presents difficulties during tokenization because of its agglutinative structure, which includes clitics such as “ف”, “ب”, “و” that attach to words. The tokenization process will employ Farasa Tokenizer and CAMeL Tools Tokenizer because these tools specialize in handling Modern and Classical Arabic texts (Taha et al., 2021).

Normalization requires text standardization to minimize morphological variations. The following steps will be taken: The removal of diacritics (tashkīl) will be performed unless they are necessary for disambiguation. The process standardizes letter forms through the conversion of “ى” to “ي” and the normalization of different “Hamza” forms. The process removes punctuation irregularities that occur during OCR scanning or in contemporary editions.

The process of Arabic stemming involves shortening words to their core elements, whereas lemmatization identifies the base dictionary form of words. These techniques serve two critical purposes: they help cluster semantically related terms, and they help identify legal terms that exist in different forms. The research will employ Farasa Stemmer for light stemming, which preserves semantic information. CAMeL Tools Lemmatizer performs both morphological disambiguation and part-of-speech tagging functions.

The verification process of annotations will occur through manual review or expert validation to confirm theological and legal accuracy when identifying disputed evidentiary uses. The preprocessed corpus will receive semi-automatic annotations through the following steps:

- Named Entity Recognition (NER) will identify jurists and schools, and locations by using pre-trained AraBERT models that have been fine-tuned on Fiqh texts.
- Evidence Markers will be used to label instances where Qur’anic verses, Hadiths, or juristic statements are introduced (e.g., “qala al-Nabī”, “istadalla bihi”, “wa fi al-hadīth”).

ii) Semantic and Contextual Analysis

The section explains the main computational methods used to analyze the corpus by combining word embeddings with semantic role labeling (SRL) and topic modeling and argument mining, and ontology-based and corpus-linguistic frameworks. The methods allow researchers to conduct an in-depth, structured analysis of the sophisticated reasoning found in classical Arabic legal discourse.

The analysis of Arabic legal corpora relies on word embedding as a fundamental technique that transforms words into dense vector representations within high-dimensional spaces. The research uses AraBERT as a BERT-based language model, which receives training on extensive Arabic corpora before fine-tuning it with jurisprudential texts to understand domain-specific terms including *qiyās* (analogical reasoning), *ḥadīth ṣaḥīḥ* (authentic tradition), *mu‘tamad* (relied-upon opinion), and *marjūḥ* (less preferred opinion).

The main difference between AraBERT and traditional bag-of-words models is its ability to understand semantic relationships in context, which enables it to distinguish between *ijmā‘ al-ummah* (consensus of the entire Muslim community) and *ijmā‘ al-ṣaḥābah* (consensus of the Companions). This method reveals both intra-school agreement and disagreement in terminology through its analysis of similar legal terms based on their usage patterns. The embedding technique enables researchers to study how different schools employ common terms in distinct ways, such as the legal definition of *‘illah* (legal cause), which differs between Hanbali and Hanafi jurisprudence (Hassan & El-Haj, 2022).

Semantic Role Labeling (SRL) functions to identify the functional roles of words and phrases in sentences by determining who performs actions and what receives actions, and under

what circumstances. The dense and elliptical argumentation in Hanbali Fiqh texts makes SRL particularly useful because it helps structure compound reasoning. An example of the SRL process is mentioned here:

"قال الإمام أحمد بعدم جواز بيع السلعة قبل قبضها واستدل بحديث زيد بن ثابت رضي الله عنه----

The system determines “Imām Aḥmad” as the legal subject and “البيع قبل القبض” (sale before possession) as the legal aspect, and “حديث زيد بن ثابت” (the hadith of Zayd) as the evidence to back up his claim. The model uses SRL to analyze the entire corpus, which helps it understand how legal arguments are constructed and defended and how they relate to each other (Al-Saadi, 2022).

The process of argument segmentation requires this step because Al-Mughnī presents multiple opinions that follow each other in succession. The analysis of legal reasoning requires SRL as its fundamental step, according to Yousef et al. (2023).

- The analysis of thematic patterns in the corpus depends on Latent Dirichlet Allocation (LDA) to perform topic modeling. The unsupervised method detects word clusters that appear together in documents to reveal concealed thematic patterns, including:
- The three main categories of religious acts include prayer and purification, and fasting, which fall under ‘Ibādāt.
- The three main categories of transactions include contracts and partnerships and loans.
- The judicial system includes three main components, which are testimony and evidence, and oaths.
- The fundamental principles of Islamic jurisprudence consist of analogy and consensus, and textual ambiguity.

Through this method, scholars can determine the frequency and specific emphasis of legal themes in Hanbali works relative to other schools of thought. The technique demonstrates how a single school modifies its priorities through time by showing commercial contracts receive more focus in later Hanbali works.

The analysis of Fiqh texts depends heavily on argument mining because legal rulings in these texts always require detailed explanations. The process requires the identification and extraction of four main elements:

- Claim: The legal ruling or fatwa.
- Evidence: Qur’anic verses, Hadiths, *ijmā’*, *qiyās*.
- Arguments: Other schools’ opinions
- Rebuttals: Refutations and clarifications.

The system used both methods (rule-based and machine learning) to determine argument markers such as *dalīluhum*, *wa qāla al-Maālīk*, and *radd ‘alayhim*. Consequently, the model divides the text into its argumentative segments.

The analysis of corpora through collocation patterns and n-gram frequency, and concordance lines reveals the repeated rhetorical strategies and emphasized terms (e.g., *fīhi ikhtilāf*, *ḥukmuḥu al-wājib*, *lā yajūz*).

The combination of ontology with corpus analysis enables researchers to ask questions like:

- Which legal issues most frequently cite analogy?
- When does consensus function as the main argument instead of serving as a supporting point?

- Which Hadiths appear in multiple domains, which shows their broad applicability?

References

1. Sungkar, Z.H., 2024. Foundations and Principles: An Introduction to Islamic Jurisprudence & A Brief Comparison to American Jurisprudence. *Student Works*, Seton Hall University. Available at: https://scholarship.shu.edu/student_scholarship/1451 [Accessed 4 July 2025].
2. Faizi, H.F.S. and Ali, D.H.S., 2024. The core principles of Islamic jurisprudence within legal theory: A comprehensive analysis. *Online Journal of Research in Islamic Studies*, 11(2), pp.57–72. Available at: <https://journals.iium.edu.my/al-fiqh/index.php/al-fiqh/article/view/282> [Accessed 3 Jul. 2025].
3. Ahmad, S. and Ijaz, M., 2022. Compilation and Evolution of Hanbali Jurisprudence. *Al-Āfāq Islamic Research Journal*, 1(2), pp.11–19.
4. Al-Qurishi, M., AlQaseemi, S. and Soussi, R., 2022. AraLegal-BERT: A pretrained language model for Arabic legal text. In: *Proceedings of the Natural Legal Language Processing Workshop 2022*, pp. 338–344. Elm Company, Research Department, Riyadh, Saudi Arabia.
5. Antoun, W., Baly, F. and Hajj, H., 2020. AraBERT: Transformer-based model for Arabic language understanding. In: *Proceedings of the 4th Workshop on Open-Source Arabic Corpora and Processing Tools with a Shared Task on Offensive Language Detection, Language Resources and Evaluation Conference (LREC 2020)*, Marseille, 11–16 May 2020. European Language Resources Association (ELRA), pp. 9–15.
6. Al-Saadi, M., 2022. *Hanbali Legal Reasoning and the Textual Tradition: A Study of al-Muqni' and al-Kāfi*. Riyadh: Dār al-Bashā'ir.

7. Azzam, M., 2023. Legal methodologies in Sunni jurisprudence: The case of Ibn Qudamah's *al-Mughni*. *Journal of Islamic Legal Studies*, 45(2), pp.112–138.
8. Saeed, A. and Ijaz, M., 2022. فقہ حنبلی کی ترویج و تطور: ایک تحقیقی جائزہ [Development of Hanbali Jurisprudence: A Critical Review]. *Al-Mu'arif*, 40(1), pp.45–78.
9. Khalifa, S., & Belinkov, Y. (2021). Named Entity Recognition and Beyond: Arabic NLP for Religious Texts. In *ACL Anthology*.
10. Hassan, S. and El-Haj, M., 2022. Towards semantic modeling of Arabic legal discourse: Applications to Islamic jurisprudence. *Language Resources and Evaluation*, 56(2), pp.319–343.
11. AlKhamissi, A., & Shaalan, K. (2023). Towards Understanding Islamic Jurisprudence through NLP: Detecting Legal Argument Structures in Arabic Texts. In *Proceedings of the Arabic NLP Workshop at EMNLP 2023*, pp. 112–121.
12. Sakr, R., Shaalan, K. and Omar, N., 2022. A survey of Arabic OCR: Challenges and techniques. *ACM Transactions on Asian and Low-Resource Language Information Processing*, 21(2), pp.1–29.
13. Taha, M., Mansour, H. and Khalifa, S., 2021. Linguistic challenges in NLP for Classical Arabic: A case study on legal texts. *Computational Linguistics for the Middle East*, 2(1), pp.45–67.
14. Abdelali, A., Darwish, K., Durrani, N. & Mubarak, H., 2021. *Farasa: A fast and furious segmenter for Arabic*. *Journal of Computational Linguistics and Arabic Language Processing*, 3(1), pp.15–32.
15. Obeid, O., Khalifa, S. & Habash, N., 2022. *CAMEL Tools: A toolkit for Arabic natural language processing*. In: *Proceedings of the Language Resources and Evaluation*

Conference (LREC 2022). Marseille: European Language Resources Association, pp.7011–7020.

16. Yousef, M., Al-Rashidi, A. and Al-Rubaye, A., 2023. Extracting Fiqh rulings from scanned Islamic texts: A hybrid OCR-NLP framework. *Journal of Islamic AI Research*, 1(1), pp.23–40.
17. Ibn Qudamah, A.A., n.d. *Al-Mughni* [online]. Available at: <https://shamela.ws/book/8463> [Accessed 4 July 2025].