



Sadiqoon
صادقون
TECHNOLOGIES

بسم الله الرحمن الرحيم

Sadiqoon Technologies Ltd

صادقون للتقنية المحدودة

The Moheet Corpus: A Shift in the Architecture of Shi'i Islamic Knowledge

From the Digital Repository to the Semantic Knowledge Structure: A Reading of the Project and Its Place on the Path of Ijtihād

A research paper on the digitization of the Shi'i heritage, artificial intelligence, and the jurisprudence of state and society

Prepared by

Shadi Ali¹

September 2026

Sadiqoon Technologies Ltd · Registered in England & Wales, No. 17120720 · 167–169 Great Portland Street, London, England, W1W 5PF

¹ Researcher in Islamic thought; lecturer at the religious seminary (ḥawza); Director, Sadiqoon Technologies and AI Solutions.

Abstract

This paper offers an introductory reading of the Moheet Corpus (*Maknaz Muḥīt*) project, a knowledge system that has absorbed the complete Arabic and Persian corpus of the Ahl al-Bayt Library (peace be upon them) software—twenty-four gigabytes of texts in hadith, jurisprudence (*fiqh*), legal theory (*uṣūl al-fiqh*), Qur'ānic exegesis, doctrine, biographical evaluation (*rijāl*), and the remaining disciplines—and re-represented it in a semantic vector space and a relational knowledge graph, whose services are made available to researchers through a standard programmatic connector.

The paper argues that the transition from "storing texts and indexing them lexically" to "structuring knowledge and representing its relations" constitutes the third generation of the digitization of the Shi'i heritage, and that—contrary to what its technical surface suggests—it is an epistemological event in the history of engagement with the classical texts, whose anticipated effects are to be measured against those of the heritage's passage from manuscript to print.

The paper then delimits the disciplined function of artificial intelligence within this architecture: it remains outside the domain of issuing legal opinions (*iftā'*) and *ijtihād*, and its function is confined to awakening attention to the latent relations between juristic concepts and their sources—relations whose discovery once consumed entire scholarly lifetimes. It connects all of this to the possibility of serving the project of moving from a juristic engagement directed at individual rulings toward a jurisprudence of the state, society, and civilization, as charted by the directives of Imam Khamenei on the necessity that the seminary (*hawza*) take charge of the needs of the Islamic order and, at the same time, master artificial intelligence. It presents a set of foreseeable practical applications, including, by way of example rather than exhaustive enumeration, the documentation of lecture transcripts, machine-assisted textual criticism (*tahqīq*), and authorship through relational exploration across disciplines and specializations, and it presents the mass-audience model already in operation in the Imam Khamenei Istiftā'āt bot and its research connector.

The paper then draws a comparison with the achievements of Sunni institutions in the same arena—from al-Maktaba al-Shāmila to Qatar's Fanar system and the charter on *iftā'* and artificial intelligence.

Keywords: Moheet Corpus; vector databases; semantic search; knowledge graph; societal jurisprudence; governmental jurisprudence; artificial intelligence; Shi'i heritage; machine-assisted textual criticism; codification of *fiqh*.

Preface: The Question of the Paper

A student of the sciences of the Ahl al-Bayt (peace be upon them) spends decades of his life building in his mind what practitioners of the craft call *iḥāṭa*—comprehensive command: knowing where a question was raised, who raised it, what was built upon it, and what contradicts it in another chapter. It is this comprehensive command—not the memorization of texts alone—that makes a jurist into a *muḥaqqiq*, a critical scholar; for a juristic idea does not reside in a single book. Its evidence is scattered across the hadith compendia, its formulations shift between the earlier and the later

scholars, and its roots extend into the sciences of biographical evaluation, legal theory, language, and exegesis. What, then, happens to this comprehensive command when the entire heritage corpus—in its Arabic and its Persian—becomes addressable by meaning rather than by wording, and when the relations among its concepts are drawn into a computational structure that can be asked questions such as: What does this concept rest upon? Who criticized that opinion? Where do parallels of this idea appear in disciplines no one has searched?

This is the subject of the Moheet Corpus, and this paper is an introduction to it and a reading of its significance. Our claim is threefold. First, the project is not "a better search engine" but a shift in the architecture of knowledge, equivalent in its historical logic to the heritage's passage from manuscript to print. Second, the function of artificial intelligence within it is neither *iftā'* nor *ijtihād*—that is a station no machine can occupy—but rather the awakening of the jurist's and the researcher's attention to relations latent in the corpus, whose discovery once required lifetimes before a scholar attained the encyclopedic familiarity needed to reach them. Third, this knowledge infrastructure is a practical condition of possibility for the shift the seminary has awaited for decades: carrying the very same Jawāhirī method of inference (*istinbāṭ*) out of the arena of individual questions and into the arena of the state, society, and civilization, culminating in a jurisprudence amenable to codification, upon which systems of governance and the project of civilization may be built.

Part One: From Repository to Thesaurus—The Three Generations of Heritage Digitization

The digitization of the Shi'i heritage—like every digitization of the human heritage—has passed through two distinct generations, and it stands today on the threshold of a third that differs from them in kind, not merely in degree.

The First Generation: Preservation by Imaging

This is the practice that began by transferring the book into an image: the projects for imaging manuscripts and lithographs that saved thousands of titles from damage and loss. This was the generation of the "repository": it offered a new vessel for old content, preserving without processing. The digital image of a page does not know what it contains, and nothing can be asked of it.

The Second Generation: The Indexed Text and Lexical Search

Then came the generation of the "searchable text," through massive programs that entered the texts in textual form and made lexical search within them possible. Foremost among these in the school of the Ahl al-Bayt (peace be upon them) is the Ahl al-Bayt Library software, which in its third release gathered more than eighteen thousand Arabic and Persian titles², not to mention the pioneering efforts of the Computer Research Center of Islamic Sciences (Noor) in Qom³. This generation

² The Ahl al-Bayt Library software (peace be upon them), second release, ablibrary.net, comprising more than seven thousand Arabic and Persian titles across the various disciplines of Islamic knowledge.

³ See, as an example of its research publications: Computer Research Center of Islamic Sciences (Noor), "Dawr al-tiqniyyāt al-ḥadītha fī taṭwīr buḥūth al-rijāl" [The Role of Modern Technologies in the Development of Rijāl Research], *Majallat Risālat al-Nūr*, no. 15, Qom, 2008.

shortened researchers' path to the text enormously, and the question "Where does this word occur?" became one that the machine answers in seconds.

Yet the ceiling of this generation is the boundary of the "word": lexical search finds only what matches its letters, and the researcher finds only what he already knew how to ask for. Whoever searches for a term such as "preservation of order" (*hifẓ al-nizām*) will not come upon the passages in the juristic corpus that express the same meaning in phrases such as "the disruption of the affairs of the Muslims," "the paralysis of livelihood," or "chaos" (*al-haraj*). Whoever traces a later term such as "legal personality" will not find its roots scattered through the jurists' discussions of "the entity" (*al-jihā*), "the title" (*al-ʿunwān*), and "the public treasury" (*bayt al-māl*), because the newly coined word has no existence in the old text even though the meaning is at work within it. Here lexical search is "an index blind to meaning": it sees the letters and not what lies beneath them.

The Third Generation: Structuring Knowledge—When Meaning Becomes the Address

The shift that Moheet represents is the shift of this third generation: the entire corpus is re-represented so that meaning itself—not wording—becomes the address by which the text is reached, and so that the relation between ideas—not the adjacency of words—becomes a preserved, queryable structure. This representation rests on two pillars, which we shall set out in Part Two: vector embedding, which converts every textual segment into a point in a mathematical space where what is close in meaning lies adjacent even when distant in wording; and the knowledge graph, which preserves the relations attested in the texts between concepts, scholars, and books as labeled relations anchored to their literal evidence.

Our choice of the word "thesaurus" (*maknaz*) to describe the project is no linguistic indulgence. In information science a thesaurus is a network of concepts with controlled relations, not a list of words. The essential difference is that traditional library thesauri are built "from above": a committee lays down subject headings and then forces the texts into them. This thesaurus, by contrast, is built "from below": from the texts themselves the concepts and their relations are extracted, so that the architecture of knowledge appears as it actually stands in the corpus, not as an external classifier presumed it to be. In this respect Moheet is closer to a geological map of the strata of Shi'i knowledge than to a catalogue of its library.

This shift has a historical precedent that illuminates its significance. In-depth studies of the history of Arabic printing have shown that the passage of the Islamic heritage from manuscript to print in the nineteenth and twentieth centuries was no neutral change of vessel. It reshaped the "knowledge architecture" itself and redefined what is read, how it is read, and which books become foundational and which are forgotten—to the point that the efforts of the first editors and publishers fashioned for the heritage the face by which we know it today⁴. Semantic digitization repeats this logic at a deeper level: it does not re-export "which books are within reach of the hand," but "which relations are within reach of the mind." Just as printing opened the door to wide collation between copies,

⁴ El Shamsy, Ahmed. *Rediscovering the Islamic Classics: How Editors and Print Culture Transformed an Intellectual Tradition*. Princeton University Press, 2020.

semantic representation opens the door to what scholars of the digital humanities call "distant reading"⁵: seeing the aggregate patterns that cannot be seen from within a single book, however devotedly the researcher practices close reading. Pioneering academic projects have already cut this path in Islamic studies, such as the Open Islamicate Texts Initiative (OpenITI), which built a machine-readable corpus of the texts of Islamic civilization and laid the foundation for the computational study of intertextuality and text reuse within them⁶. Moheet, however, advances a qualitative step along this line: from a machine-readable corpus to a corpus structured semantically and relationally, and from a tool for historians of culture to a tool for the jurist and the critical scholar at the very heart of his inferential work.

Part Two: The Knowledge Architecture of the Moheet Corpus

The Absorbed Corpus and the Unit of Construction

The project rests on a complete—not selective—absorption of the corpus of the Ahl al-Bayt Library (peace be upon them) software: all of the Arabic and Persian works, amounting to nearly twenty-four gigabytes of critically edited and imaged texts, distributed across fourteen disciplines: hadith, jurisprudence, legal theory, exegesis, doctrine and theology, biographical evaluation, ethics, history, supplication, lexicons, scholarly commentaries, philosophy and logic, language, grammar, and rhetoric, and literature⁷. This bilingual absorption has a methodological significance beyond quantity. A weighty portion of the output of the school of the Ahl al-Bayt in recent centuries—lecture notes (*taqrīrāt*), commentaries, and the literature of religious thought—was written in Persian and, owing to the language barrier, has remained almost cut off from the Arabic-reading researcher. Semantic representation pierces this barrier from the side of meaning, placing the two heritages in a single space that can be questioned in both languages.

The unit of construction in the thesaurus is neither "the book" nor "the page" but "the documented segment": a textual passage of limited length that carries with it its full card of provenance—the book, the author and his year of death, the volume, the page, and the chain of headings under which it falls in the structure of the book. The editors' footnotes are merged in their places within the text with a special tag that distinguishes them from the main text; imaged books are tagged with their origin and the archival images of their pages are preserved, so that the image of the original page can be summoned for visual verification of any retrieved text⁸. By this design it is structurally impossible for an unattributed, free-floating text to appear in the system; documentation is not a subsequent layer added to the results but a constituent of the building material itself.

The First Layer: The Vector Space—The Geometry of Meaning

⁵ Moretti, Franco. *Distant Reading*. Verso, 2013.

⁶ Miller, Matthew Thomas, Maxim G. Romanov, and Sarah Bowen Savant. "Digitizing the Textual Heritage of the Premodern Islamicate World: Principles and Plans." *International Journal of Middle East Studies*, vol. 50, no. 1, 2018, pp. 103–109.

⁷ The Moheet Corpus (*Maknaz Muḥīt*), official website moheet-ablibrary.net, and the introductory documentation of the research connector (the identity card and the tool descriptions), 2026.

⁸ Ibid. (identity card: the structure of the "documented segment"; the tags for the editors' footnotes and comments; and the tagging of imaged books and the preservation of the archival images of their pages).

The founding technology of the third generation is "vector embedding": a language model trained by Sadiqoon Technologies Ltd reads a textual segment and represents it as a numerical vector—in Moheet's case, a vector of one thousand five hundred and thirty-six dimensions produced by a multilingual, multimodal model—so that the entire corpus becomes a "nebula" of points in a mathematical space of very high dimensionality, in which segments close in meaning lie adjacent even when they share not a single word. The phrase "it is not permissible to harm others" and the phrase "whoever destroys his brother's property is liable" inhabit a single neighborhood in this space, because the model has learned from billions of examples that the two meanings attract one another⁹. The literature of information retrieval has established that this "dense retrieval" captures by a wide margin what lexical matching misses, especially in open conceptual questions¹⁰.

The thesaurus does not, however, dispense with wording when wording is what is sought—as in tracing a specific text to its source. It therefore adopts a hybrid retrieval that blends dense search (by meaning) and sparse search (by wording) through the Qdrant vector database, merging their results by the reciprocal rank fusion (RRF) algorithm, whose experimental superiority in combining heterogeneous retrieval lists has been demonstrated¹¹, and then passing the candidates to a reranking model (rerank-v4.0) that reweights them with greater precision in light of the question. Behind this chain run the high-performance approximate nearest neighbor algorithms—of the family of hierarchical navigable small world graphs (HNSW)—that make searching millions of segments a matter of fractions of a second¹². These technical details do not concern the religious scholar in themselves, but they concern him in their result: a juristic question phrased in the language of today finds its way to a text written a thousand years ago in the language of its age, without a human intermediary who knows both languages.

The Second Layer: The Knowledge Graph—The Geometry of Relation

Above the vector space stands a deeper layer: a knowledge graph built on the Neo4j database, whose nodes are concepts, scholars, and books, and whose edges are relations labeled by type: "rests upon," "contradicts," "criticized," "branches from," "commented upon," and the like. The governing design constraint here is the purest expression of the spirit of critical scholarship in the project: no relation is recorded in the graph without its literal evidence from the text and its location by book and page¹³. Through the graph the thesaurus does not "see" opinions but interrelated texts; when it answers the question "What is the relation between concept A and concept B?" it answers in the words of the authors themselves, with attribution. In the last decade knowledge graphs have

⁹ On the reference technique of sentence embedding see: Reimers, Nils, and Iryna Gurevych. "Sentence-BERT: Sentence Embeddings Using Siamese BERT-Networks." *Proceedings of EMNLP-IJCNLP 2019, ACL*, 2019, pp. 3982–3992.

¹⁰ Karpukhin, Vladimir, et al. "Dense Passage Retrieval for Open-Domain Question Answering." *Proceedings of EMNLP 2020, ACL*, 2020, pp. 6769–6781.

¹¹ Cormack, Gordon V., Charles L. A. Clarke, and Stefan Büttcher. "Reciprocal Rank Fusion Outperforms Condorcet and Individual Rank Learning Methods." *Proceedings of SIGIR 2009, ACM*, 2009, pp. 758–759.

¹² Malkov, Yu. A., and D. A. Yashunin. "Efficient and Robust Approximate Nearest Neighbor Search Using Hierarchical Navigable Small World Graphs." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 42, no. 4, 2020, pp. 824–836.

¹³ The description of the graph-exploration tool in the connector documentation: "It returns documented relations between concepts, scholars, and books (rests_upon, contradicts, criticized, branches_from, commented_upon, ...) together with the literal evidence, its page, and its source."

become the standard structure for representing interconnected knowledge in the scientific industry¹⁴, and they have entered Islamic studies through peer-reviewed academic projects. The most prominent are the SemanticHadith graph, which represented the Sunni hadith corpus, its chains of transmission, and its transmitters in an ontological representation published in the peer-reviewed *Journal of Web Semantics*¹⁵, and the graph of hadith transmitters published in the *Journal of King Saud University – Computer and Information Sciences*¹⁶. These projects, however, have remained within the bounds of limited corpora (usually the Six Canonical Collections), whereas Moheet extends the same logic across an encyclopedic, multidisciplinary corpus in both Arabic and Persian, and binds it to the constraint of literal evidence with a rigor that most academic graphs do not know.

The Third Layer: Thematic Clusters

Above the graph lies the layer of "community detection": algorithms—of the standard Leiden family¹⁷—that examine the structure of connectivity in the graph and discover "clusters," crystallizing islands of concepts with dense internal interconnection, and then generate for each cluster a thematic summary. This is the layer that answers the explorer's question before the critical scholar's: "What are the major axes connected to this conceptual domain in the corpus?" It is the mode of exploration that recent research on graph-augmented retrieval (GraphRAG) has shown to open, for aggregate global questions, what direct segment retrieval cannot¹⁸. By a close analogy, it is the "panoramic view" through which the researcher sees the topography of the corpus before descending into a particular valley of knowledge.

The Vision Layer: An Eye on the Imaged Books

The architecture is completed by an artificial vision agent that processes the imaged books never entered as text: on demand, it machine-reads the images of the retrieved pages and provides signed links to the images of the original pages¹⁹—thereby joining the end of the digital chain to its material origin and making "recourse to the original," the pillar of the craft of critical scholarship, a procedure embedded in every query.

The Tools: From Search to the Automated Craft of Tahqīq

These layers are made available to the researcher through ten integrated tools covering the cycle of scholarly work: general semantic search with fine-grained control parameters, relational exploration in the graph, a panoramic cluster survey, a bibliographic card for each book, source-tracing (*takhrīj*)

¹⁴ Hogan, Aidan, et al. "Knowledge Graphs." *ACM Computing Surveys*, vol. 54, no. 4, 2021, Article 71.

¹⁵ Kamran, Amna Binte, Bakhtawar Abro, and Amna Basharat. "SemanticHadith: An Ontology-Driven Knowledge Graph for the Hadith Corpus." *Journal of Web Semantics*, vol. 78, 2023, Article 100797.

¹⁶ Safe, M., et al. "Constructing a Knowledge Graph for Hadith Narrators." *Journal of King Saud University – Computer and Information Sciences*, vol. 34, no. 6, 2022, pp. 3385–3396.

¹⁷ Traag, V. A., L. Waltman, and N. J. van Eck. "From Louvain to Leiden: Guaranteeing Well-Connected Communities." *Scientific Reports*, vol. 9, 2019, Article 5233.

¹⁸ Edge, Darren, et al. "From Local to Global: A Graph RAG Approach to Query-Focused Summarization." arXiv:2404.16130, Microsoft Research, 2024.

¹⁹ The Moheet Corpus, identity card: "a vision agent for imaged books," and the tool for displaying the pages of imaged books through temporary signed links together with the machine-read excerpt and full documentation.

of reports, verification of hadith texts, a Qur'ānic interpreter, scholarly explanation of questions, sequential reading of pages, and display of images of the original pages²⁰. Three features among them deserve particular attention, because they reflect an awareness of the principles of the seminary craft and not mere programming skill:

First, temporal restriction by authors' dates of death. The researcher may confine his query to authors who died before or after a given Hijri year—asking, for instance, how a given question was treated by those who died before the age of al-Muḥaqqiq al-Ḥillī. This is a gateway to studies of the development of concepts across the ages (diachronic tracing) that once required the researcher to build for himself a library arranged by generational strata.

Second, preference for the earliest sources in source-tracing. The report-tracing tool accepts a request to order the locations by the priority of the author's death—an explicit automation of the rule of the craft of *taḥqīq*: attribute to the earliest, then the next earliest—with a guard that prevents a weakly matching source from taking precedence merely because of its antiquity²¹.

Third, an explicit verdict in verification. The hadith-verification tool does not content itself with displaying similar texts; it issues a classified verdict—"found," "something similar is reported," or "not found in this wording"—with a degree of match for each candidate and a guard that detects discrepancies in numbers between the two texts, the most common site of hidden error in the collation of reports²².

The Architecture of Honesty: When Acknowledging Ignorance Is a Design Feature

The gravest danger facing the deployment of large language models in the religious sciences is the phenomenon of "hallucination": the model's generation of texts and attributions that do not exist, delivered with complete linguistic confidence—a phenomenon that has become the subject of an extensive scholarly literature²³. Moheet's answer to it is not an improvement of the model but an architectural constraint upon it, following the method of retrieval-augmented generation (RAG), which shackles generation to retrieved, documented texts instead of relying on the model's memory²⁴; it then goes further by making the output of every tool subject to explicit matching thresholds. Whenever the match falls short of the threshold, the system returns the answer "no relevant results" or "no documented relations," and it treats this empty answer as "an honest result, not a malfunction." We have witnessed this behavior at first hand in documented testing sessions of

²⁰ See the tool descriptions in the research connector documentation: `moheet_search_corpus`, `moheet_graph_explore`, `moheet_get_communities`, `moheet_book_card`, `moheet_hadith_lookup`, `moheet_hadith_verification`, `moheet_quran_interpreter`, `moheet_scientific_explanation`, `moheet_read_range`, `moheet_view_page`.

²¹ Description of the `moheet_hadith_lookup` tool: "source_preference: earliest (giving priority to the earliest sources—those whose author died earliest, in accordance with the principles of the craft of *taḥqīq*)... the chronological reordering takes place exclusively within the range of strong matches—so that a weakly matching source does not take precedence merely because of its antiquity."

²² Description of the `moheet_hadith_verification` tool: "It returns an explicit verdict (found / something similar is reported—with a guard that detects discrepancies in numbers between the two texts / not found in this wording) together with the two closest matches and the degree of match of each, and full documentation of every match."

²³ Ji, Ziwei, et al. "Survey of Hallucination in Natural Language Generation." *ACM Computing Surveys*, vol. 55, no. 12, 2023, pp. 1–38.

²⁴ Lewis, Patrick, et al. "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks." *Advances in Neural Information Processing Systems*, vol. 33, 2020, pp. 9459–9474.

the Moheet Corpus: in response to a complex relational query, the system returned an explicit answer that the graph holds no relation with literal evidence matching it, rather than fabricating a plausible-looking relation. In sciences in which the text is an object of devotion, the machine's acknowledgment of ignorance is no engineering luxury; it is the condition for the permissibility of its use.

This architecture is completed by its openness to standards. The thesaurus is not presented as a closed website but as a programmatic connector on the Model Context Protocol (MCP)—the emerging open standard for connecting language models to knowledge sources—so that any research system or artificial agent can address its tools and compose upon them within a single workflow. It may thus become an early prototype of what could grow into a "network of thesauri" for Islamic knowledge, whose services layer upon and complement one another.

Part Three: The Function of the Machine—Alerting, Not Issuing Legal Opinions

Delimiting the Point at Issue

The first thing to settle in any discussion of "artificial intelligence and *ijtihād*" is the delimitation of the point at issue, for most legitimate anxiety—and most naive enthusiasm as well—arises from confusing two stations. Neither in this project nor in its founding logic is there any claim that the machine "infers," "issues legal opinions," or takes the place of the jurist in any of this. *Iftā'* is a report about the ruling of God, by which probative authority (*ḥujja*) is established over the legally responsible person; and the establishment of that authority depends upon the qualification of the reporter, his *ijtihād*, and his moral integrity (*ʿadāla*)—attributes of responsible, accountable persons, not of mathematical functions. The machine has no moral liability and no conscience, and it does not stand before God answerable for what it has reported. Hence the exclusion of the station of *iftā'* and *ijtihād* from its function is an objective exclusion by the nature of the thing, not a provisional modesty awaiting the advance of technology.

Its function, rather—and it is a noble function, sufficient to make this generation of projects proud—is "the alerting of minds": to place before the jurist and the critical scholar, in seconds, what once took decades of the lives of learned collectors to gather. It shows him the latent relations between concepts and their sources; the distant parallels across chapters; the hidden echoes in unexpected places; the dormant contradiction between two chapters that no single gaze had brought together before; and the locations of an opinion across the generations of authors. After that, the judgment is his own: he scrutinizes, argues, discards, and infers. The machine brings forward the indicators; the jurist establishes the proof.

The Precedent of the Indexes: The Thesaurus as a Lexicon of Relations

This function is no innovation in the history of the Islamic sciences; it is the newest link in a venerable chain of "instruments of attention" that scholars fashioned for themselves. The later hadith compendia—such as *Biḥār al-anwār* and *Wasā'il al-Shi'a*—were in essence projects of "restructuring" the transmitted reports: they gathered what was scattered and arranged it under the

chapters of jurisprudence, and thereby changed in practice the method of inference without claiming *ijtihād* for themselves. Indeed, the later jurist came to infer from the *Wasā'il* directly—and the *Wasā'il* is nothing but a paper "database" that al-Ḥurr al-ʿĀmilī indexed well. Then came the indexed concordances of the words of the Qur'ān and hadith, which made access by word possible, and the systematically arranged juristic encyclopedias, which made access by question possible. The semantic thesaurus inherits this lineage and raises it a qualitative degree: it is a lexicon of meanings and relations, not of words and chapters—a *Wasā'il* in which the reports are not arranged under predetermined chapters but in which the maps of interconnection are drawn as they actually stand in the textual reality. And just as no one has said that al-Ḥurr al-ʿĀmilī "issued legal opinions on behalf of the jurists" by his compilation, it cannot be said that the thesaurus "practices *ijtihād* on behalf of the jurist" by its disclosure.

Operational Evidence: Latent Relations Seen with the Eye

Lest the discussion remain in the realm of abstract possibility, we present four pieces of evidence from a documented research session conducted on the system during the preparation of this paper:

The First Case: Source-Tracing to the Earliest Sources in Seconds

The source-tracing tool was asked about the hadith "No harm and no harming" (*lā ḍarar wa-lā ḍirār*), with a request to give priority to the earliest sources. It returned the locations from *al-Kāfi*—with page numbers, hadith numbers, and contexts, among them the context of the well-known case of Samura ibn Jundab and the Prophet's (peace be upon him and his family) judgments on preemption (*shuf'a*) and surplus water—accompanied by the data of the critical edition and printing²⁵. This is a task whose manual cost anyone who has practiced source-tracing knows; its skeleton was completed in seconds, and to the critical scholar remained the station of review and ratification.

The Second Case: One Concept across Diverse Disciplines

The knowledge graph was asked about the concept of "harm" (*ḍarar*). It returned twenty documented relations extending from the books of hadith commentary to jurisprudence to ethics to the commentaries on *al-Ṣaḥīfa al-Sajjādiyya*: harm subsumed under the blameworthy traits in the ethical commentaries, branching into harm to body, property, and offspring in the literature of supplication, and connected to the rulings on recitation in jurisprudence—each relation with its literal evidence and its page²⁶. This "cross-sectional panel" of a single concept across disciplines not ordinarily brought together in a single study circle is precisely what we call the awakening of attention: the machine voiced no opinion, but it showed the researcher an entire field which, without it, he would have had to walk step by step over many years.

The Third Case: A Famous Maxim under the Lens of Verification

²⁵ Documented research session on the system (August 2026). The tool returned the locations from: *Guzīda-yi Kāfi* (Selections from al-Kāfi), edited by Muḥammad Bāqir al-Bihbūdī, Markaz-i Intishārāt-i 'Ilmī va Farhangī, 1st ed., vol. 4, pp. 468–471, "The Maxim of No Harm and No Harming," nos. 2857–2858; the original being in: al-Kulaynī, Muḥammad ibn Ya'qūb, *al-Kāfi*, Kitāb al-ma'īsha, Bāb al-ḍirār.

²⁶ Documented research session: a moheet_graph_explore query on the concept of "harm" (*ḍarar*) (August 2026); among the sources of the twenty relations returned: *Majmū'a-yi rasā'il dar sharḥ-i aḥādīthī az Kāfi*, p. 180; *Tafsīr va sharḥ-i Ṣaḥīfa-i Sajjādiyya*, p. 91; *Daryā dar Qur'ān, ḥadīth va fiqh-i islāmī*, pp. 80 and 107; *Nahj al-faṣāḥa*, p. 473.

The verification tool was presented with the maxim "People have authority over their property" (*al-nās musallaṭūn 'alā amwālihim*)—the Prophetic report that runs on the tongues of jurists in the chapters on transactions as an accepted axiom. The system's verdict was: "something similar is reported—a partial match or a match in meaning; collate the two wordings before relying on it," and it displayed the reports closest in content, with their degrees of match and their sources²⁷. This cautious machine verdict converges with what the critical scholars have established: that the maxim in this wording is transmitted only without a chain (*mursal*) in works such as *'Awālī al-la'ālī*, and that reliance upon it requires scrutiny of its chain of transmission. Observe how the "awakening" worked here in the opposite direction: the machine did not confirm the well-known, but sent the researcher back to the question of critical verification—without itself passing any judgment on the maxim.

The Fourth Case: The Virtue of Saying "I Do Not Know"

When the graph was presented with a query about a compound relation for which it holds no literal evidence, it did not strain to fabricate a plausible-looking answer; it declared explicitly that it holds no documented relation. This is a behavior that has become rare in generative AI tools in general, and its value in the domain of the religious sciences is beyond price, for it is from such conduct that the trust of the learned in a tool is built. "I do not know" is half of knowledge in the machine, as it is half of knowledge in the scholar.

The Uṣūlī Grounding: A Factory of Indicators, Not the Station of Proof

For those who wish to place this function in its proper category within the epistemology of legal theory, its nearest chapter is that of indicators (*qarā'in*) and the accumulation of probabilities. The martyred Sayyid al-Ṣadr showed in his inductive theory how objective certainty is generated from the accumulation of probabilistic indicators in an accumulation governed by a calculable mathematical logic²⁸; and contemporary seminary research has established the relevance of the calculus of probabilities to the science of biographical evaluation in particular, where reliability is a matter of graded degrees rather than a single rank²⁹—a point we set out in detail in an earlier study that presented the modeling of biographical authentication through probabilistic networks and the mechanisms of belief propagation within them³⁰. If it is true, then, that many of the premises of inference—widespread acceptance (*shuhra*), the jurists' turning away from a report (*i'rāḍ*), the tracing of opinions, and the verification of manuscript copies—are in fact the accumulation of

²⁷ Documented research session: a moheet_hadith_verification query on the text "al-nās musallaṭūn 'alā amwālihim" (August 2026)—verdict returned: "something similar is reported," with a highest degree of match of 0.72. The maxim in this wording is transmitted by Ibn Abī Jumhūr al-Aḥsā'ī without a chain (*mursal*) in: *'Awālī al-la'ālī al-'azīziyya*, edited by Muṭabā al-'Irāqī, Mu'assasat Sayyid al-Shuhadā' (peace be upon him), Qom, AH 1403 [1983], vol. 1, p. 222; and from him: al-Majlisī, *Biḥār al-anwār*, vol. 2, p. 272. Its lack of a chain has been excused on the ground that it is compensated by the practice of the early jurists and their reliance upon it across the various chapters of jurisprudence.

²⁸ al-Ṣadr, Muḥammad Bāqir. *al-Usus al-manṭiqiyya li-l-istiqrā'*. 4th ed., Dār al-Ta'āruf li-l-Maṭbū'āt, Beirut, 1982, pp. 134–137.

²⁹ al-'Āmilī, Aḥmad Abū Zayd. *Madkhaliyyat ḥisāb al-iḥtimālāt fī 'ilm al-rijāl: dirāsa fī al-manhaj*. 1st ed., Dār al-Walā' li-l-Ṭibā'a wa-l-Nashr wa-l-Tawzī', Beirut, 2011, p. 88: "Reliability is not a single rank... rather, it is graded... and the calculus of probabilities allows us to assign a numerical value to the degree of reliability of each transmitter."

³⁰ Ali, Shadi. "Taḥaddiyāt al-huwiyya al-islāmiyya fī al-'aṣr al-raqamī: al-ta'lim al-dīnī al-ḥawzawī namūdhajan" [Challenges of Islamic Identity in the Digital Age: Seminary Religious Education as a Case Study]. Previously published study by the author, version 1.1, Part Three (Reverse Engineering the Heritage: Computational Applications); and see therein the reference to: Pearl, Judea. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann, 1988, pp. 175–180.

indicators, then the thesaurus is precisely "the factory and organizer of indicators": it surveys what the individual is unable to survey and presents it ordered and weighed. Probative authority (*hujjiyya*)—that is, the conversion of these indicators into a ruling for which one is answerable—remains a purely human, *ijtihādī* station that the machine does not tread upon and should not.

Part Four: The Knowledge Infrastructure for a Jurisprudence of State and Society

Clarifying the Terminology: The Jawāhirī Method and the Arena of Its Application

Before proceeding, a common confusion must be dispelled. When it is said that we must "move from individual Jawāhirī jurisprudence to a civilizational jurisprudence of the state," what is meant—and what may legitimately be meant—is not the abandonment of the Jawāhirī method of inference. This method, the method of the author of *al-Jawāhir* in the rigorous craft of demonstrative reasoning, is the very one to which Imam Khomeini (may Allah sanctify his soul) declared his adherence even at the height of his transformative project: "I believe in traditional jurisprudence and Jawāhirī *ijtihād*, and I do not consider departing from it permissible; *ijtihād* in that manner is correct. But this does not mean that the jurisprudence of Islam is not vital and renewing; time and place are two fundamental elements in *ijtihād*"³¹. Imam Khamenei affirmed this meaning in his reading of his teacher's school: "What the Imam did in this area of jurisprudence was entirely methodical and legitimate—that is, in accordance with what the Imam himself called Jawāhirī jurisprudence; that is to say, there is no innovation in jurisprudence, but rather the correct use of the juristic criteria customary among the jurists"³².

The problem, then, lies not in the method but in the arena of its application: centuries of engagement directed this inferential apparatus at the questions of the individual—his ritual purity, his prayer, his personal transactions—and left the domains of the state, society, and civilization at the margins of attention. The martyred Sayyid al-Ṣadr diagnosed this phenomenon early and penetratingly when he distinguished between the orientation of the movement of *ijtihād* toward the "individual objective"—the application of the Islamic theory in the life of the individual—and its orientation toward the "social objective"—its application in the construction of society and its systems—and showed that the first orientation had historically monopolized the effort of *ijtihād* to the point of shaping the very manner in which questions were formulated³³. In *Iqtisādunā* he laid the foundation for the idea of the "discretionary zone" (*minṭaqat al-farāgh*), whose filling he assigned to the holder of authority

³¹ Khomeini, Ruhollah. *Ṣaḥīfat al-Imām*, vol. 21, Mu'assasat Tanzīm wa-Nashr Turāth al-Imām al-Khumaynī, Tehran, p. 289 (from his message to the religious scholars, the sources of emulation, the teachers, and the students of the seminaries, 3 Esfand 1367 SH / 21 February 1989); the rendering from the Persian is the author's.

³² Khamenei, Ali. "Speech of Imam Khamenei on the Thirty-First Anniversary of the Passing of Imam Khomeini." Official website of the Office for the Preservation and Publication of the Works of Grand Ayatollah Imam Khamenei, 3 June 2020.

³³ al-Ṣadr, Muḥammad Bāqir. "al-Ittijāhāt al-mustaqbaliyya li-ḥarakat al-ijtihād." Lecture delivered in Najaf in AH 1397 [1977], published among the publications of the Center for Specialized Research and Studies on the Martyr al-Ṣadr, Qom.

(*walī al-amr*) according to the public interest within the framework of the fixed principles—thereby raising juristic theorization from the level of the particular ruling to the level of theory and system³⁴.

It was Imam Khomeini who gave this transformation its highest formulation: "Government, in the view of the true *mujtahid*, is the practical philosophy of the whole of jurisprudence in all the corners of human life; government is the practical dimension of jurisprudence in confronting all social, political, military, and cultural problems; jurisprudence is the real and complete theory for the administration of the human being from the cradle to the grave"³⁵.

The Khameneian Demand: A Jurisprudence on the Scale of the Order

Along this line falls the body of Imam Khamenei's directives to the seminary, which—when gathered from their various places—compose an integrated demand, not scattered exhortations. He first establishes the principle of capacity and responsibility: "This seminary is capable of meeting all the needs of any Islamic order and any Islamic society,"³⁶ recalling the example of Sayyid al-Ṣadr: if the governor of the central bank or the head of a university asks about the Islamic administration of his institution, to whom does he turn? He second establishes the functional assignment: "The seminaries and the religious scholars are the pillars, and they are charged with extracting the Islamic theories from the heart of the divine texts, bringing them to light, and placing them within reach in the process of planning"³⁷. He third names the required intellectual instrument by its term: "In Qom there must be strong demonstrative lessons devoted to governmental jurisprudence at the level of the advanced seminars (*baḥth al-khārij*), so that the new governmental issues and the challenges confronting the government may be identified and examined from a juristic standpoint"³⁸. He fourth places jurisprudence in its structural rank: "Some imagine jurisprudence to be among the secondary matters because its heading is the 'branches of religion'; no, jurisprudence is the framework, and in reality the backbone of social life"³⁹; "jurisprudence means the administration of life and the elucidation of the system of social and political life, and if we pay no attention to many of its chapters, this is a deficiency on our part"⁴⁰—reaching the decisive declaration in weighing jurisprudence against other forms of knowledge: "What can administer society at present is our jurisprudence"⁴¹. He fifth warns of the cost of inaction: "Were it not for this *ijtihād*, we would have become a stagnant swamp... All these questions being raised in the world await the seminary's reply, and it must neither absent itself nor merely react,"⁴² acknowledging that "the jurisprudence that

³⁴ al-Ṣadr, Muḥammad Bāqir. *Iqtisādunā*. Dār al-Ta'ārif li-l-Maṭbū'āt, Beirut, 1987, the section "The Discretionary Zone in Islamic Economic Legislation."

³⁵ Khomeini, *Ṣaḥīfat al-Imām*, op. cit., vol. 21, p. 289.

³⁶ Khamenei, Ali. *Khiṭāb al-walī 2019* (the annual collection of Imam Khamenei's speeches). Markaz Nūn li-l-Ta'līf wa-l-Tarjama, Jam'iyyat al-Ma'ārif al-Islāmiyya al-Thaqāfiyya, Beirut, p. 125. (Quotations from the *Khiṭāb al-walī* series in this study are documented by book and page as retrieved by the SAKH research system from the approved corpus edition.)

³⁷ Khamenei, Ali. *Khiṭāb al-walī 2010*. Markaz Nūn li-l-Ta'līf wa-l-Tarjama, Jam'iyyat al-Ma'ārif al-Islāmiyya al-Thaqāfiyya, Beirut, p. 399.

³⁸ Khamenei, Ali. *Khiṭāb al-walī 2011*. Markaz Nūn li-l-Ta'līf wa-l-Tarjama, Jam'iyyat al-Ma'ārif al-Islāmiyya al-Thaqāfiyya, Beirut, p. 377.

³⁹ Khamenei, *Khiṭāb al-walī 2019*, op. cit., p. 290.

⁴⁰ Ibid., p. 291.

⁴¹ Khamenei, *Khiṭāb al-walī 2019*, op. cit., p. 125.

⁴² Khamenei, *Khiṭāb al-walī 2010*, op. cit., p. 402.

constitutes our fundamental concern has not yet developed and expanded to encompass the newly arising domains,"⁴³ and enjoining, in his message to the centenary of the Qom seminary: "Our jurisprudence must enumerate the newly arising questions of this civilization and determine their rulings"⁴⁴.

Parallel to this juristic line runs, in his thought, an explicit technological line concerning artificial intelligence in particular. He proposes including it "as one of the principal issues," because "this issue will play a role in the administration of the world in the future, and we must certainly be among the top ten countries in the world in artificial intelligence"⁴⁵. He warns against contenting oneself with the surface of the technology: "This technology has deep levels that must be mastered... If you are unable to understand the deep and varied levels of this technology and to acquire them, they will establish an agency for artificial intelligence resembling the International Atomic Energy Agency, and they are laying the foundations for that at present"⁴⁶. And he links it to the seminary's function of religious outreach (*tablīgh*): "These new things that have appeared—artificial intelligence and the like... here outreach acquires a redoubled importance"⁴⁷. Nor is this orientation confined to a single religious authority; Shaykh Nāṣir Makārim Shīrāzī has explicitly called upon the seminary to make use of artificial intelligence, holding that leaving this arena to others is impermissible; rather, "we ourselves must do this with a godly intention"⁴⁸.

At the point where these two lines intersect—a jurisprudence called upon to expand to the scale of civilization, and a technology whose depths one is called upon to master—the Moheet project finds its precise place: it is, to the best of our knowledge, the first practical realization of the meeting of the two counsels in a single structure that serves the very core of seminary work.

Why Can Societal Jurisprudence Not Rise without This Infrastructure?

It might be objected: what need has the desired juristic shift of a thesaurus and a knowledge graph? Does the resolve of the jurists not suffice? The answer emerges from an analysis of the epistemic nature of the questions of the state, and it contains four points:

First, cross-sectionality. The question of the individual usually resides in a single chapter; the question of the state is a cross-section that sweeps across chapters and disciplines. Take the question of "digital currency" as an example: delimiting the question touches the chapters on sale, currency exchange, usury, liability, *khums*, the rulings of the public treasury, and authority over currency, and it summons from outside jurisprudence the discussions of economic subject matter,

⁴³ Khamenei, Ali. *Khiṭāb al-walī 2014*. Markaz Nūn li-l-Ta'ālīf wa-l-Tarjama, Jam'īyyat al-Ma'ārif al-Islāmiyya al-Thaqāfiyya, Beirut, p. 12.

⁴⁴ Khamenei, Ali. "Message of Imam Khamenei to the International Conference on the Centenary of the Re-establishment of the Qom Seminary." Official website, 2022; cited from: Ali, Shadi. *Falsafat al-tārīkh 'inda al-Sayyid al-Khāmānā'ī* [The Philosophy of History in the Thought of Sayyid Khamenei], pp. 6–7 and 335.

⁴⁵ Khamenei, Ali. "Statements at a Meeting with Outstanding Scholarly Elites and Talents." Official website of the Office for the Preservation and Publication of the Works of Grand Ayatollah Imam Khamenei, 17 November 2021.

⁴⁶ Khamenei, Ali. "Speech of Imam Khamenei at the First Meeting with the Fourteenth Government." Official website of the Office for the Preservation and Publication of the Works of Imam Khamenei, 27 August 2024.

⁴⁷ Khamenei, Ali. "Speech of Imam Khamenei at a Meeting with Directors of the Seminaries and Preachers." Official website of the Office for the Preservation and Publication of the Works of Imam Khamenei, 12 July 2023.

⁴⁸ Makārim Shīrāzī, Nāṣir. "Statement of His Eminence Grand Ayatollah Shaykh Nāṣir Makārim Shīrāzī to the International Conference on Artificial Intelligence and the Islamic Sciences." Hawzah News Agency, Qom, 25 October 2020.

the lexicographers' discussions of "property" (*māl*) and "monetary value" (*thamaniyya*), and the history of coinage. Tracing this cross-section linearly across seven thousand titles exceeds individual capacity beyond all measure; in a semantic-relational space it is a structured query completed in a single session.

Second, the mass of opinions. Inference in the major questions weighs widespread acceptance, consensus, and the development of the legal opinion across the generations, and these weights can be established only upon a near-complete survey of the locations of an opinion—not upon a sample selected according to what happens to be present in the researcher's mind or on his bookshelf. A complete survey is by its nature the work of a machine.

Third, the jurisprudence of theory. The Ṣadrian school taught us that behind the particular rulings lies a "doctrine" and a system that must be discovered—as *Iqtīṣādunā* discovered the economic doctrine from the folds of the rulings. The discovery of the overall structure from scattered particulars is literally what the cluster layer in the knowledge graph does (that is, it reveals the framework within which thousands of particulars are organized); for the "jurisprudence of the system" first requires "seeing the system" in the corpus, and this seeing is the thesaurus's foremost commodity.

Fourth, codification and traceability. The transition from legal opinion to law—the procedural passage for any "jurisprudence of the state"—entails purely relational operations: enumerating the opinions and their sources for every article, measuring the consistency of the maxims across chapters, uncovering contradictions before the courts uncover them, and establishing a reverse chain of attribution from the legal article to its inferential evidence—what systems engineers call "traceability." The historical experiments in codification—from the Ottoman *Mecelle* to the codification of Ja'farī jurisprudence in the Iranian Civil Code—accomplished this by hand, through the efforts of generations and at the cost of severing the living connection between the article and its sources. The thesaurus promises to make this connection a permanent structure: a law in which every article is connected to its reports, opinions, and maxims by a living link that can be updated whenever *ijtihād* is renewed.

By this fourth point the thesaurus project is entitled to have its say in a prominent contemporary intellectual debate. Wael Hallaq's thesis of "the impossible state" holds that the sharia is a pluralistic epistemic system, social in its formation, which the monism of the modern state and its bureaucracy necessarily crush when they codify it⁴⁹. The structure of the thesaurus presents—at the epistemic level at least—a third path that the thesis did not contemplate: a codification that does not sever the legal text from the space of *ijtihād* from which it was born, but keeps it suspended from a structure that preserves the plurality of opinions and the sources of each, so that the law remains a "façade" for a living jurisprudence behind it, not a "substitute" that dries it up. Since Hallaq's "impossibility" rests—in one of its pillars—on the impossibility of the state apparatus achieving comprehensive

⁴⁹ Hallaq, Wael B. *The Impossible State: Islam, Politics, and Modernity's Moral Predicament*. New York: Columbia University Press, 2013 [Arabic translation: *al-Dawla al-mustaḥīla*, trans. 'Amr 'Uthmān, 1st ed., Arab Center for Research and Policy Studies, 2014].

command of the living *ijtihādī* heritage, we contend that comprehensive command is precisely what has now become possible.

Part Five: Foreseeable Applications—From the Daily Lesson to the Civilizational Project

The fruits hoped for from this structure are not all of one horizon; some are ready for harvest today, others are the project of a generation. We present them in ascending order, from what is already in the laboratory to what looms on the horizon:

First: Authorship through Relational Exploration—The Atlas of a Concept

The deepest promise the thesaurus holds for the researcher-author is a new mode of work that we call the "atlas of a concept": the researcher summons a concept—"justice," "the implicit condition," "preservation of order"—and sees before him its locations across the fourteen disciplines, its documented relations in the graph, its thematic cluster and its neighbors, and its temporal distribution through the restriction by authors' dates of death. Authorship is thereby transformed from a journey of tracing that consumes half a scholarly lifetime into a structured exploration that begins from the complete map and spends effort where it ought to be spent: on reflection, analysis, and argument.

Second: Seminary Education—From Deciphering the Expression to Grasping the Ideas

The martyred Sayyid al-Ṣadr complained that the student "was compelled to spend a large portion of his mind and energy on understanding the expression and deciphering its codes, instead of directing his whole mind to grasping the idea,"⁵⁰ and Shaykh al-Muẓaffar wrote his *Uṣūl* to free the student from spending "the bulk of his days in solving lexical puzzles"⁵¹. Semantic access places this modernizing objective on a technical footing: it lets the student ask by meaning and reach the text, summon the context of the page and what precedes and follows it, and see the commentaries on the expression from the books of commentary in ordered form. Moreover, the cluster layer offers curriculum designers thematic maps extracted from the corpus itself, suitable as a basis for conceptually organized courses. The literature of Qur'ānic and hadith computing has accumulated peer-reviewed experiments in question-answering systems bound to the sacred text, demonstrating the maturity of this direction and its amenability to control⁵².

Third: Outreach and Documented Scholarly Dialogue

⁵⁰ al-Ṣadr, Muḥammad Bāqir. *al-Ma'ālīm al-jadīda li-l-uṣūl*. 2nd ed., Dār al-Ta'āruf li-l-Maṭbū'āt, Beirut, 1981, pp. 44–46.

⁵¹ al-Muẓaffar, Muḥammad Riḍā. *Uṣūl al-fiqh*. 4th ed., Mu'assasat al-Nashr al-Islāmī, Qom, AH 1410 [1990], vol. 1, pp. 12–14 (introduction).

⁵² See: Malhas, Rana, and Tamer Elsayed. "Arabic Machine Reading Comprehension on the Holy Qur'an Using CL-AraBERT." *Information Processing and Management*, vol. 59, no. 6, 2022, Article 103068; and for a comprehensive systematic review: Bashir, Muhammad Huzaifa, et al. "Arabic Natural Language Processing for Qur'anic Research: A Systematic Review." *Artificial Intelligence Review*, vol. 56, 2023, pp. 6801–6854.

In a digital space dominated by unsupported answers, the thesaurus gives the school's preacher and its researcher the weapon of instant documented reference: an answer attributed by book, volume, and page, with the image of the page on request. This is a language that comparative scholarship across schools and religions understands, and it raises dialogue from an exchange of impressions to a confrontation of texts.

Fourth: A Foundation Layer for a Network of Research Agents

In the longer term, the most important thing about Moheet may not be its present tools but the fact that it was built from the first day as an open-standard "foundation layer" upon which specialized systems can be stacked: agents for textual criticism, teaching, and indexing, all communicating through the same standard protocol, MCP. Then it will be proper to speak of a "knowledge infrastructure for the Umma": a network of thesauri accumulating one upon another, as the generations of compendia and lexicons accumulated one upon another in our scholarly history.

Part Six: The Parallel Sunni Experience—A Map of Achievement and an Invitation to Compete

A reading of Moheet's place in the Islamic knowledge landscape is incomplete unless it is set in its wider arena. The Sunni school—with its institutions, states, and universities—has accumulated real expertise in deploying artificial intelligence and semantic, vector, and graph tools in the service of its heritage, expertise that must be known and studied rather than overlooked. Presenting it here is neither a digression nor a courtesy; it is intended for its own sake, on the principle: *وَفِي ذَلِكَ فَلْيَتَنَافَسِ* *الْمُتَنَافِسُونَ* "for such let the viers vie" (Qur'ān 83:26).

First: The Textual Infrastructure—A Culture of Open Access

The cornerstone of the Sunni experience is al-Maktaba al-Shāmila: the project that grew from the middle of the first decade of this century until it became the de facto standard of the Sunni digital corpus. According to the count on its official website, it today comprises some seven million pages in some eight thousand books by some three thousand authors, transcribed as text and freely searchable, copyable, and downloadable⁵³. Upon this foundation an intermediate layer of community tools has arisen—such as the Turāth and al-Bāḥith al-'Ilmī projects, which have represented the Shāmila texts through modern interfaces and fast search engines⁵⁴.

The first methodological lesson of this experience is the culture of open data. The Shāmila texts—through their free availability—became the great nucleus upon which the international academic OpenITI corpus was built⁵⁵, and thereby the Sunni corpus became, in practice, the material on which most of the world's computational research on Arabic Islamic texts operates. Along the same line, the sunnah.com encyclopedia has made the hadith collections available with an open data structure,

⁵³ Al-Maktaba al-Shāmila, official website shamela.ws (published count, August 2026): "about 7 million pages, 8 thousand books, and 3 thousand authors."

⁵⁴ Nuqayah Foundation, the Turāth project (a library and applications built over the Shāmila texts) and al-Bāḥith al-'Ilmī (a search engine for the heritage corpus), nuqayah.com.

⁵⁵ Miller, Romanov, and Savant, op. cit., pp. 104–106.

multiple translations, and programming interfaces relied upon by a great many applications and studies⁵⁶; the Tarteel application has turned Qur'ānic recitation practice into an artificial intelligence service that detects recitation errors through voice recognition for millions of users, and has opened its Qur'ānic data to the public through the Quranic Universal Library (QUL)⁵⁷. The upshot is a clear network effect: openness brought the world's academia to work on their corpus, academic work generated new tools upon it, and the wheel turned.

Second: The Academic Arm—From Indexing to the Graph

On the peer-reviewed research side, the Sunni corpus dominates the literature of computing in the Islamic sciences with a manifest quantitative dominance. A systematic review published in the journal *Education and Information Technologies* surveyed dozens of studies deploying machine learning in the study of hadith alone—in classification, authentication, and automated verification⁵⁸. We have already noted that the two most mature peer-reviewed knowledge graphs in this field—the SemanticHadith graph and the hadith-transmitters graph—were built on the Six Books and Sunni corpora⁵⁹; and Qatar University launched the Qur'ān question-answering shared tasks, which have become an international arena in which research teams compete annually on unified benchmarks⁶⁰. None of this is surprising; the demographic, institutional, and funding mass is larger—but acknowledging reality is the beginning of knowledge.

Third: Institutional Generative AI—Qatar's Fanar Model

The furthest point the Sunni experience has reached in the generative generation is the Qatari model: the state built its national language model "Fanar" at the Qatar Computing Research Institute of Hamad Bin Khalifa University⁶¹, and placed at its core a source-"grounded" religious question-answering system named Fanar-Sadiq, whose architecture was published at the world's foremost conference in computational linguistics: an intent router that classifies the question into nine categories (juristic ruling, Qur'ānic retrieval, calculation of *zakāt* and inheritance shares, prayer times, Hijri calendar, and so on), over more than half a million documents indexed down to the level of the verse, with a checker that prevents any non-literal Qur'ānic quotation, and deterministic (not generative) calculators for *zakāt* and inheritance. Its mean accuracy reached 76.5% across five benchmarks, and its actual use exceeded 1.9 million accesses in less than a year⁶². The significance of this model for our subject is great: "textual grounding of the religious answer"—the very essence

⁵⁶ The sunnah.com encyclopedia: the hadith collections in the Arabic text with multiple translations and open programming interfaces.

⁵⁷ Tarteel AI (tarteel.ai); and its open Qur'ānic data through the Quranic Universal Library (QUL).

⁵⁸ Sulistio, Bambang, et al. "The Utilization of Machine Learning on Studying Hadith in Islam: A Systematic Literature Review." *Education and Information Technologies*, vol. 29, no. 5, 2024, pp. 5381–5419.

⁵⁹ Kamran, Abro, and Basharat, op. cit.; Safe et al., op. cit.

⁶⁰ Malhas and Elsayed, op. cit.; and see the two Qur'an QA shared tasks (2022 and 2023) organized by the Qatar University team on the QRCD corpus.

⁶¹ Qatar Computing Research Institute (Fanar Team). "Fanar: An Arabic-Centric Multimodal Generative AI Platform." arXiv:2501.13944, 2025.

⁶² Abbas, Ummar, et al. "Fanar-Sadiq: A Multi-Agent Architecture for Grounded Islamic QA." *Proceedings of the 64th Annual Meeting of the Association for Computational Linguistics (Industry Track)*, ACL, 2026.

of the thesaurus philosophy we have set out—has become a declared industrial and state concern, not a research luxury.

Fourth: The Fatwā Institution—From the Index to the Charter to the Robot

At the level of official fatwā institutions, Egypt's Dār al-Iftā' offers the clearest institutional model. It established the Global Fatwa Index, an apparatus for automated monitoring and analysis of the fatwā space worldwide in dozens of languages; among its most recent reports is that 85% of the fatwās of official fatwā institutions have settled on the legitimacy of artificial intelligence so long as it involves nothing religiously prohibited⁶³. It held its ninth international conference (Cairo, July 2024) with the participation of delegations from 104 countries, among whose most prominent recommendations was the drafting of an ethical charter for the use of artificial intelligence in the field of fatwā⁶⁴; it then devoted its following conference (2025) to the theme "Forming the Judicious Mufti in the Age of Artificial Intelligence"⁶⁵. In Saudi Arabia, technological fatwā has descended into the densest mass arena on earth: multilingual smart fatwā and guidance robots have served the visitors of the Two Holy Sanctuaries in the courtyard of the Sacred Mosque since 2022, with their versions renewed year after year⁶⁶.

Fifth: An Evaluative Reading—Where Do We Stand in the Race?

Fairness requires that both sides be recorded together:

The strengths of the Sunni experience are four: the culture of open data and its network effect, which brought the world's academia to their corpus; the academic mass publishing in English in the major peer-reviewed venues, reaching the world's foremost publication platforms; explicit state sponsorship from Qatar, Egypt, and Saudi Arabia, which secures the continuity of funding and dissemination; and mass deployment in the field, from the façade of the Sanctuary to the phones of millions.

By contrast, the observable gap in their experience—within the limits of the survey we conducted—lies at the heart of what this paper claims for the Moheet project. Their achievement is dominated by the logic of "access, retrieval, and conversation"—open texts, improved search, and retrieval-grounded bots. But "the structuring of knowledge" in the sense we have set out—an encyclopedic graph crossing disciplines and languages, every relation in it bound to its literal evidence and page, over a complete absorption of a school's corpus, with classified layers of authority distinguishing the fatwā from the lecture transcript from the speech from the school's literature, with architectural controls for honesty, and with openness to the standard of research composition (MCP)—this

⁶³ The Global Fatwa Index, Dār al-Iftā' al-Miṣriyya (dar-alifta.org); on the report "Fatwā and Artificial Intelligence" see: *al-Yawm al-Sābi'*, Cairo, 13 April 2025.

⁶⁴ "Conclusion of the Iftā' Conference: Drafting an Ethical Charter for the Use of Artificial Intelligence in the Field of Fatwā." *al-Dustūr* newspaper, Cairo, July 2024 (the Ninth International Iftā' Conference, with the participation of delegations from 104 countries).

⁶⁵ "The Grand Mufti of the Republic: The Conference on Forming the Judicious Mufti Seeks to Keep Pace with Artificial Intelligence." Egyptian National Media Authority, 26 July 2025.

⁶⁶ "A Smart Robot in the Sacred Mosque: Religious Fatwās by Technological Means." *Al Arabiya.net*, 14 May 2026; the General Presidency for the Affairs of the Two Holy Mosques had launched multilingual smart fatwā and guidance robots as early as May 2022.

complete combination has as yet no counterpart in their experience. Their peer-reviewed graphs are confined in corpus (mostly the Six Books), their generative models retrieve but do not structure, and their mass platforms serve but do not classify the level of authority.

The position, then, is that of a compound race: they are ahead in sponsorship and dissemination, while the experience described in this paper is ahead in the depth of structuring, the method of documentation, and the layers of authority. Each side may learn from the other. Let us learn from them the culture of openness, peer-reviewed publication in the global language of scholarship, and the mobilization of a sponsoring state; and let them find in the model of the evidence-bound graph and the layers of authority something with which to develop their platforms. Indeed, the great commonalities between the two corpora—the Noble Qur'ān, the Arabic tongue, centuries of shared history, and the books of language and the lexicons—are, by virtue of open standards, candidates to become one day the foundation of cross-school networks of thesauri from which the student of knowledge may benefit whatever his school. Nor is this rhetorical wishful thinking; the technical infrastructure of both sides has become—as we have seen—mutually addressable through the same standards. وَفِي ذَلِكَ فَلْيَتَنَافَسِ الْمُتَنَافِسُونَ "for such let the vriers vie."

Part Seven: Limits and Safeguards—Lest the Promise Turn into Illusion

Rigor requires that the limits of this shift be stated with the same clarity as its promises. They are limits at three levels:

Technical Limits

Vector representation is a statistical approximation of meaning, not a comprehension of it; the model may draw together what is distant or drive apart what is close, especially in precise technical terminology whose meaning differs between disciplines—"cause" (*'illa*) for the philosopher is not what it is for the biographical critic, nor what it is for the physician. The accuracy of the thesaurus is bounded by the ceiling of the quality of the digitization beneath it: if a text enters the corpus corrupted or inadequately edited, this is reflected in its outputs. And although the relations automatically extracted into the graph are bound to literal evidence, the classification of their type (criticism or commentary? contradiction or specification?) remains a fallible machine judgment requiring cumulative human scrutiny. The most dangerous of the limits is the most hidden: "the illusion of comprehensiveness." The corpus—however vast—is not everything that was written, and not finding something in it does not equal its nonexistence in reality; the system itself states this limitation in the definitions of its tools⁶⁷. But the human mind is inclined to treat an apparently comprehensive database as an exhaustive enumeration, and this inclination must be resisted through education.

Methodological and Religious Safeguards

⁶⁷ From the description of the verification tool in the connector documentation: "Non-existence does not amount to a judgment on the report—coverage is bounded by the books absorbed."

The golden rule that ought to be inscribed over the door of every scholarly use of these systems is this: the outputs of the machine are stimuli for the treasures buried in minds, not proofs in themselves; the retrieved segments are an indication of the existence of a text whose source must be consulted; a drawn relation is a claim whose evidence must be examined; and a high degree of match is a statistical preponderance that does not take the place of verification.

Hence human verification of every output upon which a scholarly or practical consequence is built is no research luxury but a condition of the validity of the work. Under no circumstances may a fatwā be attributed to the machine, may its statistical retrieval be taken as the basis of a religious ruling, or may the layperson resort to it as one resorts to a source of emulation (*taqlīd*); the station of *iftā'* and *taqlīd* lies entirely outside this architecture, as we set out in detail in Part Three.

A Civilizational Safeguard: Sovereignty over the Knowledge Infrastructure

In our earlier study we showed how Orientalism moved from the study of texts to the engineering of data, and how biased algorithms became an intermediary reshaping the Muslim's awareness of himself and his heritage⁶⁸. That analysis carries over here directly: whoever owns the structure through which the heritage is questioned owns—in practice—the power to shape what is seen of the heritage and what is hidden. If thesauri of the Islamic heritage will inevitably be built—the technology is available and the corpora are digitized—then the question is not "Will they be built?" but "Who will build them, and on what values of documentation and trustworthiness?" Building them by trustworthy hands from within the school's own scholarly milieu—as in this project—is one instance of what we called in the earlier study "mastering the technology" as a collective civilizational obligation (*farḍ kifāya*); for the alternative is not a vacuum but a compulsory foreign intermediary between us and our heritage.

Conclusion

The upshot of the paper in three sentences. The Moheet Corpus—by absorbing the complete Arabic and Persian corpus of the Ahl al-Bayt Library (peace be upon them) and re-representing it semantically and relationally—carries the digitization of the Shi'i heritage from the generation of the repository and the index to the generation of knowledge structuring, an epistemological shift equivalent in its logic to the heritage's passage into print. The legitimate, disciplined function of this architecture is to awaken the attention of the jurist and the critical scholar to the relations latent in the corpus—bringing forward the indicators and surveying them—not *iftā'* or *ijtihād*, which remain a human station of religious responsibility that the machine does not tread upon. And this knowledge infrastructure is a practical condition of possibility for the seminary to rise to what the eminent figures of the school—from Imam Khomeini to Imam Khamenei—have demanded of it: a jurisprudence that expands to encompass the state, society, and civilization, and that admits of codification without losing its living connection to its inferential sources.

⁶⁸ Ali, "Taḥaddiyāt al-huwiyya al-islāmiyya fī al-'aṣr al-raḡamī," op. cit., Part One (The New Orientalism: From the Study of Texts to the Engineering of Data).

If we may sum up the project's place in the knowledge landscape in a single image, we would say: the thesaurus is a nervous system for the heritage, not a thinking mind in its stead. It connects the extremities to the centers, carries signals from the farthest reaches of the corpus to the point of attention, and makes the whole scholarly body feel each part of itself—while the mind that weighs, judges, and infers remains where it has always been: the fully qualified jurist, heir to a method of argument refined over a thousand years. Technology here is no foreign newcomer demanding a station for itself, but a new servant in the lineage of the compendia, indexes, and lexicons that our scholars have never ceased to fashion in order to perfect their command of what lies before them. The servant of this age, however, reads meaning and draws relation, and for that reason the transformation it brings is qualitative, not quantitative.

It remains to be said that the shift is not completed by the erection of the structure but by its people taking charge of it: that the seminaries, the centers of textual criticism, and the codification committees adopt it as material for their daily work; that a generation of students of knowledge graduate who combine the craft of inference with the culture of these tools and an awareness of their limits; and that on its basis a new specialization worthy of the name "Islamic knowledge engineering" be opened. Then it may rightly be said that the Umma did not receive the age of artificial intelligence as a passive consumer, but brought its most venerable intellectual heritage into this age on its own terms: the text preserved, the attribution exact, probative authority in its proper station, and attention—the most precious thing a scholar possesses—multiplied by a trustworthy machine.

References

First: Arabic and Persian Sources

Ibn Abī Jumhūr al-Aḥsā'ī, Muḥammad ibn 'Alī. *'Awālī al-la'ālī al-'azīziyya fī al-aḥādīth al-dīniyya*. Edited by Muḥtabā al-'Irāqī. Qom: Mu'assasat Sayyid al-Shuhadā' (peace be upon him), AH 1403 [1983].

Khamenei, Ali. *Khiṭāb al-walī 2010* (The Complete Collection of Imam Khamenei's Speeches). Beirut: Markaz Nūn li-l-Ta'līf wa-l-Tarjama, Jam'iyyat al-Ma'ārif al-Islāmiyya al-Thaqāfiyya.

Khamenei, Ali. *Khiṭāb al-walī 2011*. Beirut: Markaz Nūn li-l-Ta'līf wa-l-Tarjama, Jam'iyyat al-Ma'ārif al-Islāmiyya al-Thaqāfiyya.

Khamenei, Ali. *Khiṭāb al-walī 2014*. Beirut: Markaz Nūn li-l-Ta'līf wa-l-Tarjama, Jam'iyyat al-Ma'ārif al-Islāmiyya al-Thaqāfiyya.

Khamenei, Ali. *Khiṭāb al-walī 2019*. Beirut: Markaz Nūn li-l-Ta'līf wa-l-Tarjama, Jam'iyyat al-Ma'ārif al-Islāmiyya al-Thaqāfiyya.

Khamenei, Ali. "Statements at a Meeting with Outstanding Scholarly Elites and Talents." Official website of the Office for the Preservation and Publication of the Works of Grand Ayatollah Imam Khamenei, 17 November 2021.

Khamenei, Ali. "Speech of Imam Khamenei on the Thirty-First Anniversary of the Passing of Imam Khomeini." Official website of the Office for the Preservation and Publication of the Works of Imam Khamenei, 3 June 2020.

- Khamenei, Ali. "Speech of Imam Khamenei at a Meeting with Directors of the Seminaries and Preachers." Official website of the Office for the Preservation and Publication of the Works of Imam Khamenei, 12 July 2023.
- Khamenei, Ali. "Speech of Imam Khamenei at the First Meeting with the Fourteenth Government." Official website of the Office for the Preservation and Publication of the Works of Imam Khamenei, 27 August 2024.
- Khamenei, Ali. "Message of Imam Khamenei to the International Conference on the Centenary of the Re-establishment of the Qom Seminary." Official website of the Office for the Preservation and Publication of the Works of Imam Khamenei, 2022.
- Khomeini, Ruhollah. *Ṣaḥīfat al-Imām*. Tehran: Mu'assasat Tanzīm wa-Nashr Turāth al-Imām al-Khumaynī (may Allah sanctify his soul), vol. 21.
- Hallaq, Wael B. *The Impossible State: Islam, Politics, and Modernity's Moral Predicament*. New York: Columbia University Press, 2013. [Arabic translation: *al-Dawla al-mustaḥīla*, trans. 'Amr 'Uthmān, 1st ed., Doha/Beirut: Arab Center for Research and Policy Studies, 2014.]
- Makārim Shīrāzī, Nāṣir. "Statement of His Eminence Grand Ayatollah Shaykh Nāṣir Makārim Shīrāzī to the International Conference on Artificial Intelligence and the Islamic Sciences." Hawzah News Agency, Qom, 25 October 2020.
- al-Ṣadr, Muḥammad Bāqir. *al-Usus al-mantiqiyya li-l-istiqrā'*. 4th ed. Beirut: Dār al-Ta'āruf li-l-Maṭbū'āt, 1982.
- al-Ṣadr, Muḥammad Bāqir. "al-Ittijāhāt al-mustaqbaliyya li-ḥarakat al-ijtihād." Lecture delivered in Najaf in AH 1397 [1977], published among the publications of the Center for Specialized Research and Studies on the Martyr al-Ṣadr, Qom.
- al-Ṣadr, Muḥammad Bāqir. *Iqtisādunā*. Beirut: Dār al-Ta'āruf li-l-Maṭbū'āt, 1987.
- al-Ṣadr, Muḥammad Bāqir. *al-Ma'ālim al-jadīda li-l-uṣūl*. 2nd ed. Beirut: Dār al-Ta'āruf li-l-Maṭbū'āt, 1981.
- al-'Āmilī, Aḥmad Abū Zayd. *Madkhaliyyat ḥisāb al-iḥtimālāt fī 'ilm al-rijāl: dirāsa fī al-manhaj*. 1st ed. Beirut: Dār al-Walā' li-l-Ṭibā'a wa-l-Nashr wa-l-Tawzī', 2011.
- Ali, Shadi. "Taḥaddiyāt al-huwiyya al-islāmiyya fī al-'aṣr al-raqamī: al-ta'līm al-dīnī al-ḥawzawī namūdhan" [Challenges of Islamic Identity in the Digital Age: Seminary Religious Education as a Case Study]. Previously published study by the author, version 1.1.
- Ali, Shadi. *Falsafat al-tārīkh 'inda al-Sayyid al-Khāmānā'ī* [The Philosophy of History in the Thought of Sayyid Khamenei]. (Within the extended corpus of the SAKH research system.)
- al-Majlisī, Muḥammad Bāqir. *Biḥār al-anwār al-jāmi'a li-durar akhbār al-a'imma al-aṭḥār*. 3rd ed. Beirut: Dār Iḥyā' al-Turāth al-'Arabī, AH 1403 [1983], vol. 2.
- Computer Research Center of Islamic Sciences (Noor). "Dawr al-tiqniyyāt al-ḥadītha fī taṭwīr buḥūth al-rijāl" [The Role of Modern Technologies in the Development of Rijāl Research]. *Majallat Risālat al-Nūr*, no. 15, Qom, 2008.
- al-Muẓaffar, Muḥammad Riḍā. *Uṣūl al-fiqh*. 4th ed. Qom: Mu'assasat al-Nashr al-Islāmī, AH 1410 [1990].

The Imam Khamenei Istiftā'āt Project. Official website (pages: the Service, Methodology, the Research Connector), khamenei-online.org, 2026.

The Moheet Corpus (*Maknaz Muḥīt*). Official website and the introductory documentation of the research connector, moheet-ablibrary.net, 2026.

al-Maktaba al-Shāmila. Official website of the al-Maktaba al-Shāmila project, shamela.ws.

Nuqayah Foundation. The Turāth and al-Bāḥith al-'Ilmī projects, nuqayah.com.

Dār al-Iftā' al-Miṣriyya. The Global Fatwa Index, dar-alifta.org.

"Conclusion of the Iftā' Conference: Drafting an Ethical Charter for the Use of Artificial Intelligence in the Field of Fatwā." *al-Dustūr* newspaper, Cairo, July 2024.

"The Grand Mufti of the Republic: The Conference on Forming the Judicious Mufti Seeks to Keep Pace with Artificial Intelligence." Egyptian National Media Authority, 26 July 2025.

"Fatwā and Artificial Intelligence: The Latest Reports of the Global Fatwa Index." *al-Yawm al-Sābi'*, Cairo, 13 April 2025.

"A Smart Robot in the Sacred Mosque: Religious Fatwās by Technological Means." Al Arabiya.net, 14 May 2026.

The Ahl al-Bayt Library Software (peace be upon them). Second release, ablibrary.net.

Second: Foreign-Language Sources

Abbas, Ummar, et al. "Fonar-Sadiq: A Multi-Agent Architecture for Grounded Islamic QA." Proceedings of the 64th Annual Meeting of the Association for Computational Linguistics (Industry Track), ACL, 2026.

Bashir, Muhammad Huzaifa, et al. "Arabic Natural Language Processing for Qur'anic Research: A Systematic Review." *Artificial Intelligence Review*, vol. 56, 2023, pp. 6801–6854.

Cormack, Gordon V., Charles L. A. Clarke, and Stefan Büttcher. "Reciprocal Rank Fusion Outperforms Condorcet and Individual Rank Learning Methods." Proceedings of the 32nd International ACM SIGIR Conference on Research and Development in Information Retrieval, ACM, 2009, pp. 758–759.

Edge, Darren, et al. "From Local to Global: A Graph RAG Approach to Query-Focused Summarization." arXiv:2404.16130, Microsoft Research, 2024.

El Shamsy, Ahmed. *Rediscovering the Islamic Classics: How Editors and Print Culture Transformed an Intellectual Tradition*. Princeton University Press, 2020.

Harrag, Fouzi. "AI-based Hadith Authenticity Checking and Classification." *International Journal of Islamic Applications in Computer Science and Technology*, vol. 2, no. 1, 2014, pp. 1–9.

Hogan, Aidan, et al. "Knowledge Graphs." *ACM Computing Surveys*, vol. 54, no. 4, 2021, Article 71.

Ji, Ziwei, et al. "Survey of Hallucination in Natural Language Generation." *ACM Computing Surveys*, vol. 55, no. 12, 2023, pp. 1–38.

Kamran, Amna Binte, Bakhtawar Abro, and Amna Basharat. "SemanticHadith: An Ontology-Driven Knowledge Graph for the Hadith Corpus." *Journal of Web Semantics*, vol. 78, 2023, Article 100797.

- Karpukhin, Vladimir, et al. "Dense Passage Retrieval for Open-Domain Question Answering." Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP), ACL, 2020, pp. 6769–6781.
- Lewis, Patrick, et al. "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks." Advances in Neural Information Processing Systems, vol. 33, 2020, pp. 9459–9474.
- Malhas, Rana, and Tamer Elsayed. "Arabic Machine Reading Comprehension on the Holy Qur'an Using CL-AraBERT." Information Processing and Management, vol. 59, no. 6, 2022, Article 103068.
- Malkov, Yu. A., and D. A. Yashunin. "Efficient and Robust Approximate Nearest Neighbor Search Using Hierarchical Navigable Small World Graphs." IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 42, no. 4, 2020, pp. 824–836.
- Miller, Matthew Thomas, Maxim G. Romanov, and Sarah Bowen Savant. "Digitizing the Textual Heritage of the Premodern Islamicate World: Principles and Plans." International Journal of Middle East Studies, vol. 50, no. 1, 2018, pp. 103–109.
- Moretti, Franco. Distant Reading. Verso, 2013.
- Pearl, Judea. Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan Kaufmann, 1988.
- Qatar Computing Research Institute (Fanar Team). "Fanar: An Arabic-Centric Multimodal Generative AI Platform." arXiv:2501.13944, 2025.
- Reimers, Nils, and Iryna Gurevych. "Sentence-BERT: Sentence Embeddings Using Siamese BERT-Networks." Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing (EMNLP-IJCNLP), ACL, 2019, pp. 3982–3992.
- Safe, M., et al. "Constructing a Knowledge Graph for Hadith Narrators." Journal of King Saud University – Computer and Information Sciences, vol. 34, no. 6, 2022, pp. 3385–3396.
- Sulistio, Bambang, et al. "The Utilization of Machine Learning on Studying Hadith in Islam: A Systematic Literature Review." Education and Information Technologies, vol. 29, no. 5, 2024, pp. 5381–5419.
- Sunnah.com. The Hadith of the Prophet Muhammad at Your Fingertips, sunnah.com.
- Tarteel AI. AI-Powered Qur'an Memorization and Recitation, tarteel.ai.
- Traag, V. A., L. Waltman, and N. J. van Eck. "From Louvain to Leiden: Guaranteeing Well-Connected Communities." Scientific Reports, vol. 9, 2019, Article 5233.

Sadiqoon Technologies Ltd

Shadi Ali — Director

shady.ali@sadiqoon.uk — sadiqoon.uk

Sadiqoon Technologies Ltd · Registered in England & Wales, No. 17120720 · 167–169 Great Portland Street, London, England, W1W 5PF