A Webtool for Hadith Authorship and Verification with Scholarly Expertise

Amina Daoud
College of Science and Engineering
Hamad Bin Khalifa University
Doha, Qatar
e-mail: aminadaoud@hbku.edu.qa

Devi G. Kurup
College of Science and Engineering
Hamad Bin Khalifa University
Doha, Qatar
e-mail: dkurup@hbku.edu.qa

Hamada R. H. Al-Absi*
College of Science and Engineering
Hamad Bin Khalifa University
Doha, Qatar
e-mail: haalabsi@hbku.edu.qa

Jens Schneider
College of Science and Engineering
Hamad Bin Khalifa University
Doha, Qatar
e-mail: jeschneider@hbku.edu.qa

Wajdi Zaghouani
College of Humanities and Social Sciences
Hamad Bin Khalifa University
Doha, Qatar
e-mail: wzaghouani@hbku.edu.qa

Saeed Mohd H. M. Al Marri
College of Shariah and Islamic Studies
Qatar University
Doha, Qatar
e-mail: saealmarri@qu.edu.qa

Younss Ait Mou*
College of Science and Engineering
Hamad Bin Khalifa University
Doha, Qatar
e-mail: ymou@hbku.edu.qa

***Abstract*—Muslim scholars have collected Prophet Mohammed's teachings, sayings and actions several years after his death. These teachings have been collected into what is called "Hadith" in Islam, which is composed of several books. In recent years, digital databases have emerged based on the Hadith books; however, these databases lack validation by Muslim Hadith Scientists. This paper presents the development of a Webtool designed to validate Hadith collections, addressing the challenge of maintaining the integrity and accuracy of these essential Islamic texts in the digital era. By leveraging scholarly collaboration, the tool facilitates a thorough validation process for each Hadith and the chains of narrators. Feedback from Hadith Scientists has been instrumental in refining the tool's features, ensuring its academic accuracy and user-friendliness. The project signifies a crucial step towards preserving Islamic teachings and proposes a model for future technological advancements in the field of Islamic scholarship.**

***Index Terms*—Hadith; Hadith Scientist; Verification; Webtool.**

I. Introduction

A Hadith (which translates to speech or teaching) describes Prophet Mohammed's -Peace Be Upon Him (PBUH) - teachings, sayings, actions, or silent approvals [1]. The plural form of the word 'Hadith' is 'Ahadeeth,' or 'Hadiths' and these Ahadeeth have been observed/learned by the Prophet's companions throughout his life and were passed down through generations; however, it is essential to note that these Ahadeeth were only collected or documented and authenticated many years after the death of the Prophet PBUH [2]. These Ahadeeth are considered the second source of Islamic law; Muslims refer to them always besides the Holy Quran [3]. Each Hadith consists of two parts, the chain of narrators (Sanad/Isnad) which comes at the beginning and shows who have transmitted the Hadith, and the actual body of the Hadith (Matn/Matan), which is the actual teaching or saying or action of the Prophet PBUH. The proliferation of digitized Hadith collections online has brought concerns regarding their authenticity, accuracy, and reliability [2]. Many of these online databases lack scholarly validation, raising doubts among researchers and scholars [2]. For example, in this work, we used Hadith database of Sahih Bukhari from [4]. Hadith Scientists in our team have discovered and corrected multiple issues/errors in this database related to "Matan", "Sanad", the ordering of Hadiths, errors in chapters titles, etc. Table I shows some examples of these issues and how the Hadith Scientists corrected them.

Such issues or errors pose a significant challenge for those seeking to engage in rigorous academic or religious studies relying on these sources. To address these challenges and contribute to the field of Islamic scholarship, we aim to create a comprehensive and validated database of Hadith from the six major collections, including Sahih Bukhari, Sahih Muslim, Sunan Abu Dawood, Jami al-Tirmidhi, Sunan al-Sughra, and Sunan ibn Majah. With 114 chapters, 6,236 verses, and 157,935 words, the Quran provides teachings on daily matters, social dealings, and historical events. Advances in Natural Language Processing (NLP) enable computational techniques to facilitate Qur'anic research, offering new avenues for learning and understanding [5]. In our research, we employ NLP on Ahadeeth research, offering new techniques for learning and teaching. As an initial step for the research,

*Corresponding authors

 ISBN: 978-1-68558-147-3

TABLE I
EXAMPLES OF ISSUES/ERRORS FOUND IN THE ISLAMONLINE.NET HADITH DATABASE.

Type	Description
Matan	Hadith number 3535 was missing from the database.
Matan	Part of Hadith number 3757 was missing.
Sanad	In Hadith Number 3967, two of the narrators were substituted with well-known companions of the Prophet (Saad Bin Abi Waqas instead of Saad Bin Ibrahim and Abi Umama Al Bahili instead of Abu Umama Bin Sahl).
Chapters and Hadiths order	In the first few books, Hadiths were not ordered correctly (Note: this might be due to different editions of Sahih Bukhari) There were also errors in the ordering of chapters.
Books and Chapters	Some books have chapters with wrong title or without title. For example, in the Book of Marriage, there are four chapters, one without a title, one with the title: "Your prayers, your faith", another one with the title: "A Muslim is the one from whose hand Muslims are safe", and the fourth with the title "Food feeding". All titles were not related to marriage, yet the Hadiths enclosed in them are about marriage.

we create a comprehensive Hadith dataset through an expert authoring system. To ensure the accuracy and authenticity of each Hadith, we developed a Webtool which allows Hadith Scientists to manually validate the main text and the chain of narrators, correct any mistakes, and extract meaningful keywords for each Hadith. This tool allows several scholars to work in parallel. Ultimately, it provides a reliable and trustworthy resource for scholars, researchers, and the broader Muslim community. This contributes to the preservation and dissemination of authentic digital Hadiths for academic and religious purposes. At this stage, the Webtool has been developed and we started the validation of Sahih Al-Bukhari book. Noteworthy initiatives have paved the way for effective annotation processes. These include the development of interfaces like MADARi, a joint morphological annotation and spelling correction system for texts in Standard and Dialectal Arabic; and MANDIAC, a web-based annotation system designed for rapid manual diacritization of Standard Arabic text. These systems enable annotators to manage and organize the entire annotation process remotely and simultaneously [6]–[8]. To the best of our knowledge, there exists no previous work that has focused on validating Hadith digital collections, with Hadith Scientists going through each Hadith to ensure accuracy and integrity of the text.

The rest of this paper is organized as follows: Section II presents the interface design of the expert authoring and verification system; Section III presents a discussion on the web tool; and we present concluding remarks in Section IV.

II. EXPERT AUTHORING AND VERIFICATION SYSTEM INTERFACE DESIGN

A. User Requirements and Feedback

The primary users of this system are Hadith Scientists. These scientists are led by Professor Saeed Al-Marri from Qatar University, who is a principal investigator in this project and a renowned expert in Hadith science. Together with his team, they utilize their vast experience in Hadith science, and they use various books in the science of Hadith as well as specialized libraries, such as 'Al-Maktabah Al-Shamelah' – The Comprehensive Library. This library contains searchable books in different areas of the Islamic faith, including Hadith science. To facilitate the user interaction, an initial prototype was created to address their requirements which focused on:

- Verification of Hadith Authenticity – Hadith Scientists requested a feature to verify the displayed Hadith authenticity by entering comments on its "Matn" and its corresponding chain of narrators "Sanad".
- Correction of Hadith – The Hadith Scientists required a feature to correct any inaccuracies identified in the displayed Hadith.
- Displaying narrators' details – the users requested to include a functionality to access the details of each narrator in the chain, including information about their locations, number of Hadiths, number of students they had, etc.
- Additionally, the users also requested to save all the secondary names of narrators.

The prototype was used by the Hadith Scientists allowing them to interact with the system and to get their feedback which focused on the need to correct various discrepancies in Hadiths which included the following:

- Interface editability – Users requested for an editable interface, allowing them to correct the book names, chapter names, narrator names within the "Sanad", to enhance the accuracy of the verification process.
- Addition of collections, books, chapters, and Hadiths – Users expressed the interest of adding new collections, books, chapters and Hadith details instead of extracting data from online resources, which were not accurate.
- Refinement of Hadith Order – Users expressed the need of reordering the Hadiths through the interface for data accuracy.

These user requirements and feedback served as valuable insights, guiding the development and refinement of the verification system to meet the specific needs of Hadith Scientists.

 ISBN: 978-1-68558-147-3

B. System Architecture

The architecture of the interface for the expert authoring and verification system is designed to address the specific needs and feedback from Hadith Scientists. The system integrates various components to support data storage and retrieval of Hadith data in the form of user interaction and feedback incorporation. Data was initially sourced from [4], specifically Sahih Al-Bukhari data and is stored in MongoDB database. MongoDB is a popular NoSQL "Not only Structured Query Language" database, which uses a document-oriented data model that allows to store complex and unstructured data in a more natural way [9] [10]. During the initial phase of the data storage, we developed a robust structure to ensure an effective organization of data within the database. The database encompasses details about various elements such as Hadith collections, books, chapters, and specific details about each Hadith, including "Matn" and "Sanad". Other relevant information such as narrators' details, locations, etc. are stored effectively in their respective normalized collection within the database. As a first step of developing the interface, the first task was to create a login page to ensure the security of the system. This login page was designed to restrict access only to intended users, who are the Hadith Scientists. Creating a login page is an essential step in securing the application, guaranteeing that only authorized users have access. The login page is a simple and user-friendly interface that facilitates easy access for authorized users to interact and contribute to the verification system. Figure 1 shows the login page.

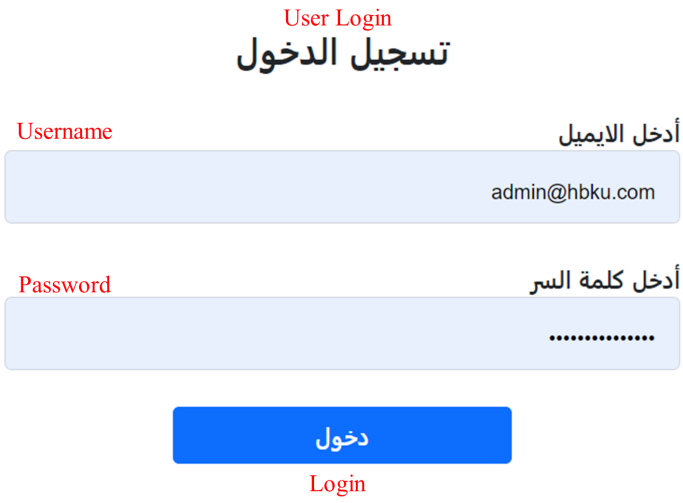

Fig. 1. Login page of the Webtool. English translation of all fields and functions are shown in red.

C. Core Features and Functionalities

The system's features and functionalities are comprehensively designed to cater to the needs of the primary users. A preliminary prototype was developed to enable the interaction between the primary users and the system. This prototype was developed based on the initial user requirements, focusing on the functionalities, such as Hadith authenticity verification, correction of inaccuracies in "Matns" and "Sanad", and the display of detailed narrator information. There is also a provision to add missing "Sanad" and "Moallakka" (*Moallakka Hadiths are those with incomplete initial chains of narrators, considered "suspended" in Islamic scholarship*) for the selected Hadith. By clicking on the displayed "Sanad" list, users can access corresponding detailed narrator information, including names, variations, locations, number of students, teachers, and associated Hadiths. Following user feedback, we incorporated additional features to the existing ones, encompassing the addition of new Hadith collections, books, chapters, and specific Hadith details. Moreover, the interface also includes the functionality to update Hadiths by employing a regular expression to separate "Sanad" and "Matn", enhancing the ease of extraction of those from Hadiths. Additionally, the system allows users to associate keywords with each Hadith, facilitating the information retrieval for research purposes. The system is designed in such a way that, for the initial use of the tool, it displays the first Hadith in the first chapter under the first book inside the first collection. As an update occurs, the subsequent openings of the tool will display the details of the last modified Hadith. Users can customize the display by selecting the collection followed by book, chapter and finally the desired Hadith for verification. The interface is thoughtfully organized into three blocks, aligned from right to left for Arabic letters. Each block can be maximized or minimized, enhancing user accessibility. The homepage provides functionalities for adding, updating, and deleting entries based on specific conditions. Additionally, the interface supports the reordering of Hadiths through a drag-and-drop method, allowing for the insertion of missing Hadiths while maintaining sequence numbers. To fortify security, the system incorporates measures like parameterized queries, to prevent injection attacks and ensuring the safety and validity of user inputs. The user interface is thoughtfully designed, featuring clear instructions for straightforward navigation. Figure 2 illustrates the main page of our Webtool application, which is segmented into three blocks:

- In the block on the right, there are three sections: collection, book, and chapter. Users have the option to select a collection, book, and chapter. Upon selection in the respective order, the Hadiths corresponding to the selections from these sections are displayed in their respective blocks. Additionally, each section provides an option to add new content, update existing content, and remove content if there are no subentries within their respective sections for collection, book, and chapter.
- The middle block displays all the Hadiths based on the selections made in the right block, with the first Hadith in the list highlighted, by default. Users have the option to choose which Hadiths to annotate, reorder Hadiths, change the chapters of Hadiths, add new Hadith, etc.
- The block on the left facilitates the users in editing the selected Hadith from the middle block. Users can rectify

 ISBN: 978-1-68558-147-3

Fig. 2. Webtool functionalities. English translation of all fields and functions are shown in red.

the mistakes in the selected Hadith (Matn), add and edit the chain of narrators (Sanad), Moallakka, including names and their order in the chain, add new narrator if they are not listed, and input comments related to the Sanad or the Matn. Furthermore, users can input relevant keywords for the selected Hadith.

The users can save the updates and navigate through Hadiths using the ”Next” and “Previous” buttons provided on the page without explicit user selection. Additionally, the users have the flexibility to expand and collapse each block according to their preference for enhanced accessibility. Figure 3 visually illustrates this functionality on the main page, with the third block expanded while the others collapsed, providing a clear navigation option for all users.

The system architecture supports an iterative development approach, allowing for continuous refinement based on user requirements and feedback. This comprehensive system architecture is tailored to meet the specific needs of Hadith Scientists, providing a robust and user-friendly environment for the verification and correction of Sahih Al-Bukhari data in this initial stage.

D. Tools and Technologies

The expert authoring and verification web interface system employs a comprehensive architecture, integrating Flask API [11]–[13] , MongoDB, HTML and jQuery to meet the specific needs of Hadith Scientists. The system’s structure is designed to facilitate data storage, retrieval, user interaction, and seamless incorporation of user feedback. The frontend of the web interface was developed with HTML and jQuery components, providing a user-friendly and interactive environment for Hadith Scientist to engage with the system. jQuery enables dynamic manipulation of the Document Object Model (DOM), enhancing the user experience during interactions with the interface. Figure 4 shows the architecture of the Webtool.

Flask is a lightweight and flexible web framework for Python and it is ideally suited for the development of web applications with simplicity and efficiency. Its modular design allows us to build a scalable and maintainable web application by integrating various extensions and libraries, as needed. Leveraging Flask built-in support for routing, request handling, and template rendering, the Webtool streamlines the development process while reducing boilerplate code. In our application, Flask API serves as an interface between the frontend and the database, handling HTTP requests from the frontend, processing data and orchestrating interactions with the MongoDB database. It communicates with the MongoDB database to retrieve, update and store data based on user interactions. MongoDB, a NoSQL database renowned for its flexibility, scalability, and performance in handling unstructured and semi-structured data, serves as the primary data storage solution. Its document-oriented model enables the storage of data in JSON-like documents, making it suitable for storing diverse and evolving data structures. MongoDB’s schemaless design eliminates the need for a predefined schema, facilitating easy modification of data models. MongoDB is used for data storage in our application, storing data in a document-oriented format. It is queried and updated by the Flask API to provide the required information to the frontend. By combining Flask, MongoDB, and HTML, we have developed a robust and feature-rich Webtool that leverage the strengths of each technology. Flask provides the backend infrastructure and routing capabilities, MongoDB offers a flexible and scalable data storage solution, and HTML

 ISBN: 978-1-68558-147-3

Fig. 3. Hadith block after expansion. English translation of all fields and functions are shown in red.

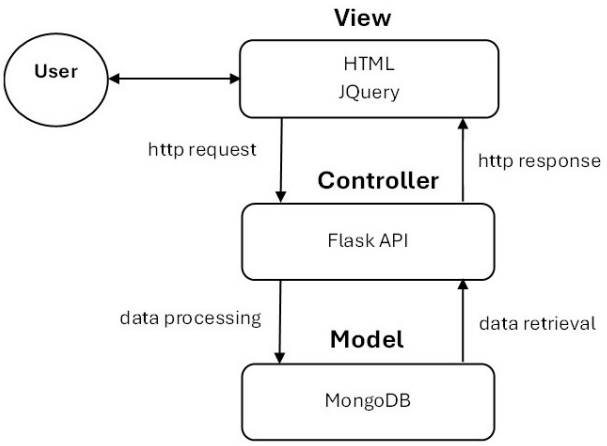

Fig. 4. Webtool Architecture.

enables the creation of intuitive and user-friendly interface. To further enhance the organization and maintainability of our codebase, we have adopted Flask Blueprints [13]. These blueprints serve as a powerful tool, allowing us to organize our application into separate components, each handling specific functionalities. This modular approach facilitates code reuse and scalability, making it particularly beneficial for large-scale applications where maintaining a well-organized codebase is crucial. Specifically, we utilized blueprints for functionalities such as user login, the main page, where the core functionalities are maintained, and the display of narrator details. This ensures that our code remains modular, scalable, and easily maintainable, contributing to the overall efficiency and robustness of our application. By combining these tools and technologies, the expert authoring and verification system is built to meet the specific needs of our primary users, providing a robust, secure, and user-friendly platform for the verification and correction of Hadiths from six collections.

III. Discussion

The development of a Webtool aimed at authenticating Hadith collections is a significant leap forward in the intersection of Islamic scholarship and digital technology. This work acknowledges the critical challenge of preserving the authenticity and accuracy of Hadith, an essential component of Islamic faith and jurisprudence, in the vast and often unregulated digital landscape. By enabling scholarly collaboration, the tool facilitates a thorough validation process, ensuring each Hadith's text and its chain of narrators are carefully examined. This collaborative approach not only streamlines the authentication process but also provides access to validated Hadiths, empowering scholars and the broader Muslim community with reliable resources. Moreover, the integration of feedback from Hadith scholars into the tool's design underscores a dynamic and responsive development process. This feedback is critical for refining the tool's functionalities, such as verification of authenticity, correction of inaccuracies, and the addition of comprehensive narrator details. Such features address the needs of Hadith scholars, ensuring the tool remains both user-friendly and academically accurate. This responsiveness to scholarly input highlights the project's commitment to serving the needs of Islamic academia and contributes to its potential as a cornerstone resource for Hadith studies.

 ISBN: 978-1-68558-147-3

IV. Conclusion

In this work, we presented a Webtool to authenticate and verify Hadiths. This work extends beyond the immediate benefits of improved access to authenticated Hadiths. It sets a precedent for future works in the digital preservation and study of Islamic texts, suggesting pathways for integrating technological solutions with traditional Islamic scholarship. The project's success could inspire further innovations in the field, including the development of similar tools for other Islamic texts and the adoption of advanced technologies for scholarly research. As such, this Webtool represents not just a solution to a current challenge but a steppingstone towards a future where technology and tradition blend to advance Islamic scholarship.

Acknowledgment

This work has been supported by Qatar National Research Fund (QNRF), now known as the Qatar Research Development and Innovation (QRDI) Council, under grant number (NPRP13S-0129-200194).

References

[1] A. Ramzy *et al.*, "Hadiths classification using a novel author-based hadith classification dataset (abcd)," *Big Data and Cognitive Computing*, vol. 7, no. 3, p. 141, 2023.

[2] S. Hakak *et al.*, "Digital hadith authentication: Recent advances, open challenges, and future directions," *Transactions on Emerging Telecommunications Technologies*, vol. 33, no. 6, p. e3977, 2022.

[3] F. Binbeshr, A. Kamsin, and M. Mohammed, "A systematic review on hadith authentication and classification methods," *Transactions on Asian and Low-Resource Language Information Processing*, vol. 20, no. 2, pp. 1–17, 2021.

[4] IslamOnline, https://islamonline.net/, accessed: 15/12/2023.

[5] M. H. Bashir *et al.*, "Arabic natural language processing for qur'anic research: a systematic review," *Artificial Intelligence Review*, vol. 56, no. 7, pp. 6801–6854, 2023.

[6] O. Obeid *et al.*, "Madari: A web interface for joint arabic morphological annotation and spelling correction," in *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, 2018, pp. 2616–2622.

[7] ——, "Mandiac: A web-based annotation system for manual arabic diacritization," in *The 2nd Workshop on Arabic Corpora and Processing Tools*, 2016, pp. 16–21.

[8] ——, "A web-based annotation framework for large-scale text correction," in *The Companion Volume of the Proceedings of IJCNLP 2013: System Demonstrations*, 2013, pp. 1–4.

[9] B. Jose and S. Abraham, "Exploring the merits of nosql: A study based on mongodb," in *2017 International Conference on Networks and Advances in Computational Technologies (NetACT)*. IEEE, 2017, pp. 266–271.

[10] A. Utka *et al.*, "Berri corpus manager: A corpus analysis tool using mongodb technology," in *Human Language Technologies–The Baltic Perspective: Proceedings of the Ninth International Conference Baltic HLT 2020*, vol. 328. IOS Press, 2020, pp. 166–173.

[11] P. Smyth, *Creating web APIs with Python and Flask. Programming Historian 7*. ProgHist Ltd, 2018.

[12] D. Gaspar and J. Stouffer, *Mastering Flask Web Development: Build Enterprise-grade, Scalable Python Web Applications*. Packt Publishing Ltd, 2018.

[13] J. Perras, *Flask Blueprints*. Packt Publishing Ltd, 2015.

 ISBN: 978-1-68558-147-3