

## Program for Developing the Novel Quran and Hadith Authentication System

Amirrudin Kamsin<sup>1</sup>, Abdullah Gani<sup>1</sup>, Ishak Suliaman<sup>2</sup>, Salinah Jaafar<sup>3</sup>, Aznul Qalid Md Sabri<sup>1</sup>, Mohd Yamani Idna Idris<sup>1</sup>, Zaidi Razak<sup>1</sup>

<sup>1</sup>Faculty of Computer Science and Information Technology;

<sup>2</sup>Academy of Islamic Studies, <sup>3</sup>Academy of Malay Studies;

University of Malaya, 50603 Kuala Lumpur, Malaysia.

[1amir, 1abdullah, 2ishaks, 3b1salina, 1aznul, 1yamani, 1zaidi}@um.edu.my](mailto:{1amir, 1abdullah, 2ishaks, 3b1salina, 1aznul, 1yamani, 1zaidi}@um.edu.my)

### ABSTRACT

It is a challenge to identify who actually holds the valid copy of the holy Quran, or whether one digital copy is tampered or not. In fact, previous literature has shown that most of people were not aware of the distribution of fake copies of Quran online. Majority of them have raised the importance of having a central Islamic body to control and determine the authenticity of the holy Quran. We therefore, propose to develop and evaluate the Quran authentication system. The aims are to provide reliable and intuitive system to assist both the central body and end users to assess the authenticity of the digital Quran applications, before using them. It can be used as a tool/mechanism to improve the digital Quran publishing regulations and users' confidence towards digital Quran applications.

**Keywords:** Quran, Hadith, Malay translation, Authentication System, Framework, Prototype

### 1. Introduction

The holy Quran and Hadith are the two main sources for Muslims to guiding their life. Allah sent the holy Quran, as the absolute revelation from Him to all human beings. Hadith signifies particularly the words and the life of the Prophet Muhammad (peace be upon him) as a premier source of Islam after the Holy al-Quran. More than a billion of Muslims used the Quran, as the first criterion to guide and unite them in order to live in peace. It is intolerable to even do the slightest modifications to the content of the holy Quran. With the advancement of technology, people have now not only used the physical Quran (i.e. Mushaf) but also the digital Quran on the websites or phone applications which are easily accessible. Unfortunately, little attention has been given to determine and classify the authenticity of these digital Quran applications. Such ignorance has led to other serious problems – distributions of illegal or distorted digital copies/applications of the holy Quran worldwide. This has led in to proposing a research program to address the issue. In this paper, we divide the paper into four sections. Section 2 discusses the issue in detail and the current state of the art. Section 3 explains the proposed methodology to undertake this research project. Section 4 describes the initial prototype of the Quran and Hadith authentication to be developed.

## 2. Background

The Quran is the holy religious book for more than one billion Muslims around the world (Tabrizi and Mahmud, 2013; Alshareef, and El Saddik, 2012). It is the first criterion to refer and apply when there are differences or social or religious problems among people (Alshareef and El Saddik, 2012; Khan, 2010). It is the absolute speech of Allah, explaining what is right and what is wrong, distinguishing between truth and falsehood (Al-Jabari, 2008). The Quran represents Allah's authority which can also be used to verify the contents of hadith (Al-Jabari, 2008). There have been various types of applications developed and abundance of websites about the Quran published online, providing a quick and easy way for users to recite and learn the Quran (Adhoni et al., 2013).

However, due to the lack of research effort and controlling authority to address this issue, there have been some fake versions of the Quran exist and attempts to create unauthentic Quran to undermine Muslims around the world (Khan and Alginahi, 2013; Kurniawan et al., 2013). It is a challenge for readers to verify whether a certain verse is true or fake, and to determine the accuracy of a verse due to unintended typo errors (Kurniawan et al., 2013). This may contribute to the loss of confidence towards e-content among users, and thus it is essential to control and improve the credibility of the content distributed through Internet or digital applications (Alshareef, 2013). The Quran which is universal and not meant to specific region or people, and therefore, attention must be paid to facilitate non-Arabic speakers to understand the word of their Creator, Allah, using authentic resources (Elobaid et al., 2014). The developers who provide existing applications for distributions, which have no information on the certification and authentication, may not have the Quranic scholars to approve their applications [6]. Without the body to regulate the development and distribution of these existing applications, there is an issue whether the applications on users' mobile are authentic or not (Khan and Alginahi, 2013). The ignorance of the issues has led to other problems of illegal copying and distributing of digital information, making it difficult to control who has invalid copies of the digital Quran (Laouamer and Tayan, 2013). To date, there is a lack of complete applications with reliable and authentic data.

To address this, it is important to improve publishing regulations laws, and to develop an efficient, a better and comprehensive framework or an effective and efficient system to determine the authenticity of the Quran in order to satisfy the end user's needs (Alshareef and El Saddik, 2012; Alginahi, et al., 2013; Alshareef, 2013). This is to confirm that the Quran has no mistake and not even a single letter or symbol of it has been tampered (Alginahi, et al., 2013). Previous study identified that almost 77% of their respondents require an International Islamic body to determine the authenticity of the digital copies of the Quran (Khan and Alginahi, 2013). The same study also found that most of their respondents were not aware of the existence of fake copies of the Quran available online. It is a challenge for regular/lay people to distinguish whether a particular surah is tampered or not (Alshareef, 2013).

It is important to ensure that the digital holy Quran distributed on the Internet, mobile phones, etc. are authentic as it first revealed about 1400 years ago. The main goal of the Holy Quran is to achieve and establish peace and harmony, resolve any conflicts with justice, bringing humanity out from darkness into light, from ignorance into knowledge, from disease

and malady into health, from straying into guidance, from seclusion and exclusion into inclusion and togetherness, from war and adversity into peace and friendship, and from distancing away from one another into getting to know one another (Jassem and Jassem, 2001). It is a huge responsibility for all Muslims to strive for it and look after the holy Quran and combat any malicious or unintentional attempts that can destroy the authenticity of the Quran.

Hence, this project aims to develop an authentication system to validate verses from the digital Quran (applications or websites) with the original and standard digital mushaf. It is hoped that this can improve the efficiency and accuracy in checking the authentication of Quran and Hadith which being practiced in Malaysia currently. The holy Quran Islamic religion was transcribed into its original text in Arabic and in lithographic forms which known as the Mushaf, and the Mushaf of Al-Uthmani has become the Quranic standard lithograph (Shamsudin and Farooq, 2000). The system will verify the match and detect any discrepancies, even to the slightest found between them. This will enable a central Islamic body or users to verify the authenticity the digital Quran applications distributed over the Internet, before using them (Alshareef, 2013). Furthermore, the central body can use the authentication analysis provided by the system in order to produce a more reliable and up-to-date list of authentic Quran published on the Internet and available applications for download (Alginahi, et al., 2013). At present, there is still no standard used to verify and show the authenticity of the Quran and Hadith in Malaysia (for example, please refer to JAKIM's website (JAKIM, 2014). The outcome from this research will be beneficial for Malaysians, in particular, for having a system where they can easily and quickly check whether a particular Quranic or Hadith material is authentic or not.

### **3. Methodology**

This section explains the objectives of the research project (or main program) as a whole as well as its underlying research phases, activities, and key components involved.

#### **3.1 Objectives**

The objectives of this program are as follows:

1. To develop criteria for evaluating the usability of authentication systems/tools.
2. To design and develop the digital Quran and Hadith authentication system:
  - 2.1 To validate whether a copy of digital Quran is authentic or fake.
  - 2.2 To determine the types of differences/distortions found between a copy of digital Quran and the standard digital Quran (i.e. Mushaf Uthmani).
  - 2.3 To determine the accuracy of a copy of digital Quran as compared with the standard digital Quran and highlight the differences found between them.
3. To evaluate the usability of the authentication:
  - 3.1 To analyze the differences/distortions/temperaments found between two copies of digital Quran at different types/levels of comparisons (by surah, ayah, huruf, diacritics, tajweed and other recitation symbols).
  - 3.2 To evaluate how efficient/strong the chosen algorithm.
  - 3.3 To provide and test algorithm for verification and authentication of the fundamental text of the digital Quran as compared to the standard Uthmani mushaf.
4. To establish Hadith database that consists of authenticated texts of Hadith from authority

references, translated texts of Hadith and elaborated lessons, injunctions and conclusions that derive from the Hadith.

5. To preserve and validate the authority of texts of Hadith Nabi s.a.w. that appears on the website of E-Hadith Malaysia.

### 3.2 Research phases

The program is planned for 3 years and involves the following 3 main phases which refer to (see Table 1):

Table 1: Research phases and activities

Year/ Phases	Sub-activities			
<b>I_a</b>	Identify requirements for digital Quran and Hadith authentication system	Identify guidelines for determining the authenticity of Quran and Hadith	Identify techniques to implement the Digital Quran and Hadith authentication system	Identify the Digital Quran and Hadith authentication system specification or architecture
<b>I_b</b>	Constructing algorithm for the digital Quran and Hadith authentication system	Develop system architectural framework and high-level design for the digital Quran and Hadith authentication system	Design GUI/mockup for the digital Quran and Hadith authentication system	Conduct focus groups/workshops to determine requirements for the digital Quran and Hadith authentication system
<b>II</b>	Develop coding for the digital Quran and Hadith authentication system	Build and test database the digital Quran and Hadith authentication system	Develop security mechanism for the digital Quran and Hadith authentication system	Integrate components/modules for the digital Quran and Hadith authentication system
<b>III</b>	Deploy the digital Quran and Hadith authentication system	Conduct expert/user evaluation of the digital Quran and Hadith authentication system	Prepare documents for the digital Quran and Hadith authentication system (e.g. report, manual)	Provide system consultation/support

## 4. Results

### 4.1 Framework

In order to run this program, we have proposed a framework that identifies key components involved in the project (see Table 2) as well as their respective sub-programs.

Table 2: Framework for developing Quran and Hadith authentication system

<b>Services</b>	JAKIM (Department of Islamic Development Malaysia)	KDN (Malaysia Ministry of Home Affairs)	MAYBANK, etc.	Focus
<b>Applications</b>	Data mining	Search engine	Meaning based query	Sub program 1
<b>Repository</b>	Database Semantic query Knowledge representation			Sub program 3
<b>Enablers</b>	Knowledge extraction Binary coding Tagging			Sub programs 1 and 2
<b>Infrastructure Device</b>	Cloud computing Mobile			Sub program 1

The programs involved in the project are as follows:

*Sub program 1: Quran and Hadith Authentication System, a Unicode Centric Approach*

*Sub program 2: Developing Authentication Policies for a Repository of Authenticated Al Quran and Hadith*

*Sub program 3: Al Hadith Semantic Validation System*

The focus of each program will be summarized below.

#### 4.1.1 Sub program 1

The focus of sub program 1 will be on the developing a novel authentication system for the Quran and Hadith. The existing procedure for checking the authenticity is based on a manual approach. This manual approach of authentication is both time consuming and prone to human errors. Therefore, it is important to implement an automated approach in order to maximize the efficiency of the authentication procedure. To automate this process, we

propose to develop the Quran and Hadith authentication systems using a Unicode centric string matching approach. Firstly, an authenticated repository will be established. Next, test materials will be extracted and converted into standard Unicode data. Both the authenticated data and the test data will be compared using a proposed string matching method. Specific machine learning techniques will be used to classify the authenticity of the Quran and Hadith. In the final stage, the algorithm will be evaluated for their performance, efficiency and effectiveness. The expected outcome from this research will be a novel algorithm for determining the authenticity of Quran and Hadith.

#### **4.1.2 Sub program 2**

The main issues that the program will focus are as follows:

- The issue of establishing the authentic texts of the digital Hadith becomes a crucial part for this research. What are the main sources of authentic texts of Hadith in Islam that will be derived as the authentic texts for the digital Hadith? For the first phase of research, the textual Hadith books of al-Kutub al-Sittah that consists of 6 main books of Hadith i.e. Sahih al-Bukhari, Sahih Muslim, Sunan Abu Dawud, Sunan al-Tirmidhi, Sunan Nasa'ie and Sunan Ibn Majah will be derived as the authentic texts for the digital Hadith in this research. The research will investigate the authentic aspects of Arabic letters, the Arabic terms and its signing of textual Hadith that derive from those 6 main books of Hadith. Bearing in mind there are 34,961 texts of Hadith derive from the al-Kutub al-Sittah. Therefore, indeed it is a comprehensive research to be undertook under this research.
- The issue of verifying the quality of texts of Hadith (takhrij) is the critical issue to be investigated in this research. In the science of Hadith studies, there are 3 main standard of qualities for the texts of Hadith i.e. Sahih (authentic), Hasan (good), and Da'if (weak). Only the texts of Hadith that derive in Sahih Bukhari and Sahih Muslim are verified by Muslim scholars as Sahih. However, the texts of Hadith in Sunan Abu Dawud, Sunan al-Tirmidhi, Sunan Nasa'ie and Sunan Ibn Majah consist of 3 qualities of Hadith i.e. Sahih (authentic), Hasan (good) and Da'if (weak). Therefore, the challenge for this research is to determine and verify all texts of Hadith of those 4 Sunan books of Hadith into the 3 standard of qualities i.e. Sahih (authentic), Hasan (good) and Da'if (weak).
- The issue of establishing the standard translation of the texts of Hadith that derives from the textual Hadith books of al-Kutub al-Sittah is a vital issue for this research. This part of research requires the expertise of Arabic language as well as Malay language. Apart from languages expertise, the expertise in reading the book of Syuruh al-Hadith (explanation of textual Hadith) is much required in order to establish the standard translation of the texts of Hadith.
- The issue of establishing the repository of the digital Hadith for the outcomes of the previous process is the essential part for this research. This issue is a technical issue of which the expertise of information technology is required in order to establish and design the repository of the digital Hadith. This issue is a climate part of the research in which the issue of innovative and development of the research outcomes is applied.

### 4.1.3 Sub program 3

The sub program 3 focuses on determining the accuracy of the semantic of Quran and Hadith. The program will develop a checklist for determining the meaning and the most suitable Malay words to be used in the translations of Quran and Hadith. The program will also process from the old Malay and Indonesia languages into a standard Malay language. The program will check the accuracy the meaning of Quran and Hadith that have been translated.

### 4.2 Initial prototype

We have designed an initial prototype in our attempt to develop the first Quran authentication system. In this section, we summarize the implementation of the prototype and provide important figures to illustrate the key processes involved in it (i.e. a flow of the Quran authentication system that we aim to develop). First, a Quran page captured from the Internet and the authenticity of the page will be checked (see Figure 1). The application will scan each line for errors and differences. Each line will be compared to a Quran page that has already been authenticated by the Department of Islamic Development Malaysia (JAKIM) stored in a database. The application will continue scanning all lines for errors until the end of page (see Figure 2). If there are no errors, the line will show up green. If there are any errors or differences, the line will show up red. If the page has errors, the application will display the percentage of errors or differences once the page has finished being scanned (see Figure 3). In the event if the page that has successfully scanned and no errors were found, the successful message showing its authenticity will be displayed (see Figure 4). In the event that an error is detected, the application will highlight the area that contains the error and displays a warning.

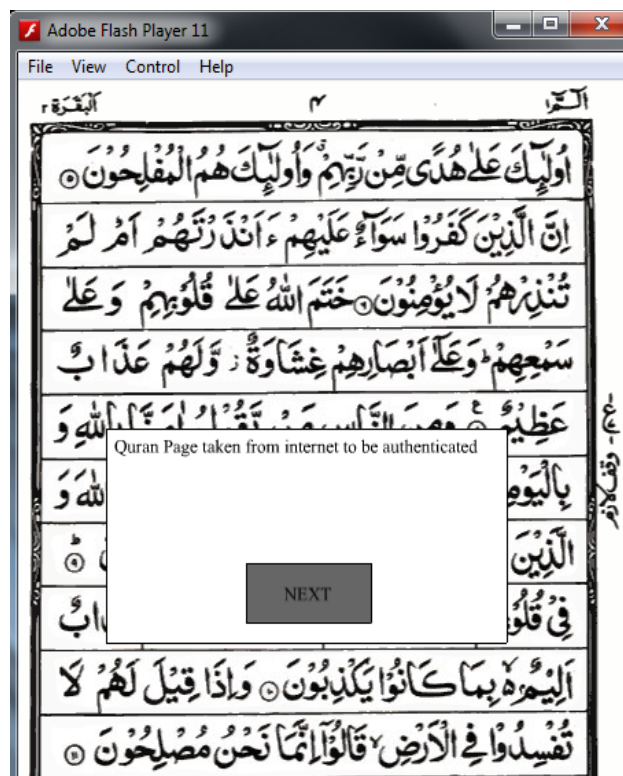


Figure 1: Example of a page of Quran captured form the Internet



Figure 2: The page of Quran will be scanned to check for its authenticity

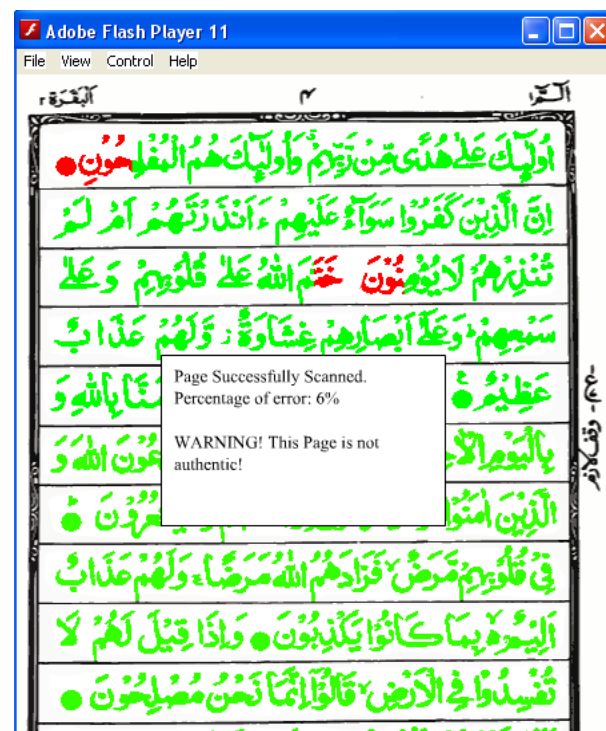


Figure 3: Errors found message



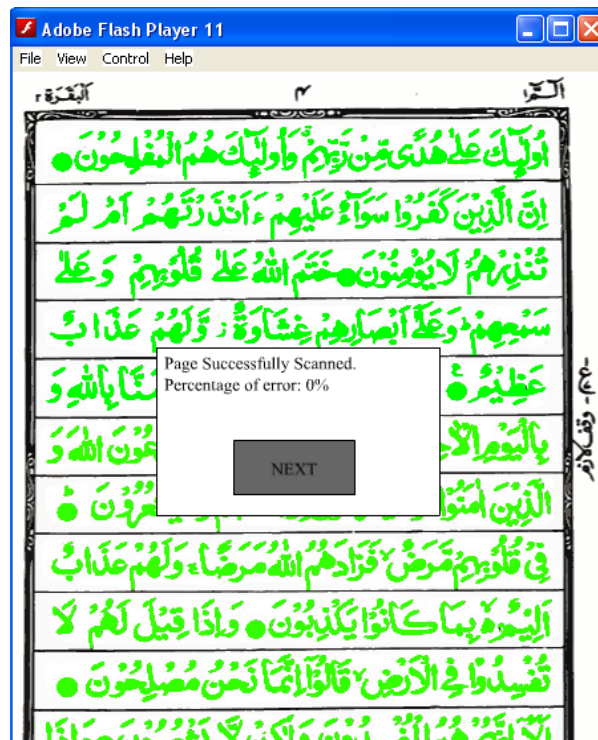


Figure 4: A message to show that the page of Quran is authentic

Similar to the Quran authentication system described above, the Hadith authentication system also apply the similar approach used in the Quran authentication system (described previously). Additionally, the Hadith authentication application will also scan for its translation in order to determine the authenticity of the Hadith. The application will scan each line of the hadith (captured from the Internet) as well as its translation for errors. If there is no error, the scanned line will display the green colour. This means that the hadith and its translation are authentic. If errors are found, the area will be marked with red colour. This means that the hadith is not authentic.

## 5. Conclusion

Upon completion this project, we aim to have developed the first repository (globally) that is endorsed by the Malaysian government through JAKIM, containing authentic Quran and Hadiths (translations and interpretations) from primary sources. We also aim to have built the novel & accurate authentication system for digital Quran & Hadith. The system can validate them from semantic & linguistic aspects. Such accuracy is vital in order to ensure that the usage of primary sources of references is authentic and acknowledged by Muslim scholars.

## References

Tabrizi, A. and Mahmud, R. (2013). Issues of Coherence Analysis on English Translations of Quran. 1st International Conference on Communications, Signal Processing, and their Applications, 1-6, Sharjah.

- Alshareef, A. and El Saddik, A. (2012). A Quranic Quote Verification Algorithm for Verses Authentication. International Conference on Innovations in Information Technology (IIT). 339-343.
- Khan, I.A. (2010). Authentication of Hadith. Redefining the Criteria. MPG Books Group, UK.
- Al-Jabari, R. (2008). Reasons for the Possible Incomprehensibility of Some Verses of Three Translations of the Meaning of the Holy Quran into English. PhD Thesis. University Salford, UK.
- Adhoni, Z.A and Al Hamad, H, Siddiqi, A.A, Parvez, M, Adhoni, Z.A. (2013). Cloud-based Online Portal and Mobile Friendly Application for the Holy Qur'an. Life Science Journal;10(12s).
- Khan, M.K and Alginahi, Y.M. (2013). The Holy Quran Digitization: Challenges and Concerns. Life Science Journal 2013; 10(2). 156-164.
- Laouamer, L and Tayan, O. (2013). An Enhanced SVD Technique for Authentication and Protection of Text-Images using a Case Study on Digital Quran Content with Sensitivity Constraints. Life Science Journal 2013;10(2), 2591-2597.
- Shamsudin, F. and Farooq, A. (2000). "AI natural language in metasynthetics of Al-Qur'an," 2000 TENCON Proceedings. Intelligent Systems and Technologies for the New Millennium (Cat.No.00CH37119), pp. 464-467.
- Alginahi, Y.M., Tayan, O and N. Kabir, M. (2013). Verification of Qur'anic Quotations Embedded in Online Arabic and Islamic Websites. International Journal on Islamic Applications in Computer Science And Technology, Vol. 1, Issue 2, September 2013, 41-47.
- Kurniawan, F., Khalil, M.S., Khan, M.K and Yasser M. Alginahi, Y.M. (2013). Authentication and Tamper Detection of Digital Holy Quran Images. International Symposium on Biometrics and Security Technologies. IEEE, Computer Society. 291-296.
- Alshareef, A. (2013). Design and Development of a Quote Validation Tool for Arabic Scripts. Master thesis. School of Electrical Engineering and Computer Science University of Ottawa
- Elobaid, M., Hameed, K, Eldow. M.E.Y. (2014). Toward Designing and Modeling of Quran Learning Applications for Android Devices. Life Science Journal 2014;11(1), 160-171.
- Jassem, Z. A. and Jassem, A.J. (2001). Abdullah Yusuf Ali's Translation of the Quran: An Evaluation. Issues in Education, Volume 24, 2001.
- JAKIM. 2014. Senarai Hadith. URL: <http://www.islam.gov.my/en/e-hadith>. (Last accessed: 18th J