Résumé

Avec la pénétration d'Internet dans les pays arabes dépassant la pénétration moyenne mondiale et la croissance des utilisateurs d'Internet dépassant 2000% au cours de la dernière décennie, le nombre de documents arabes en ligne a connu une croissance exponentielle. Par conséquent, avec l'augmentation du nombre de publications en ligne en arabe. La présence de citations coraniques dans des articles en ligne devient de plus en plus notable.
Les documents en ligne utilisent des signes diacritiques complets et des symboles Tajweed pour tous les mots. Par conséquent, la vérification du texte cité du Coran, des signes diacritiques et des symboles Tajweed est essentielle pour le verset vérification de l'authenticité.

Il est important que les versets cités soient écrits sans faute et que les versets coraniques doivent être authentifié pour garantir que la parole d'Allah n'est pas altérée même un seul le déplacement ou l'omission d'une lettre ou d'un signe diacritique considère le texte comme invalide.
Dans ce travail on va étudier les méthodes et les techniques d'authentification du Saint Coran et mettre une application d'authentifier les versets Coraniques (Hafs) sur format de texte en utiliser tokenisation et traitement du langage naturel (NLP).
Les mots Clés : *NLP , Authentification, Tokenisation , Saint Coran (Hafs)*

Abstract

With Internet penetration in Arab countries exceeding average global penetration and Internet user growth exceeding 2000% in the last decade, the number of Arabic documents online has grown exponentially.
Therefore, with the increase in the number of online publications in Arabic.
The presence of Qur'anic quotes in online articles is becoming more and more noticeable.
Online documents use complete diacritics and Tajweed symbols for all words.
Therefore, the verification of the quoted text of the Qur'an, diacritic marks and Tajweed symbols is essential for the verse verifying authenticity.
It is important that the quoted verses are written without fail and that the Qur'anic verses must be authenticated to guarantee that the word of Allah is not altered even a single displacement or omission of a letter or a sign diacritic considers the text invalid.
In this work we will study methods and authentication techniques of the Holy Quran and put an application to authenticate the Quranic verses (Hafs) on text format using tokenization and natural language processing (NLP).

***Key words**: NLP, Authentication, Tokenization, Holy Quran (Hafs)*

ملخص

نظراً لتجاوز معدل انتشار الإنترنت في الدول العربية متوسط الانتشار العالمي ونمو مستخدمي الإنترنت الذي تجاوز 2000٪ في العقد الماضي ، فقد زاد عدد المستندات العربية على الإنترنت بشكل كبير.

لذلك ، مع زيادة عدد المنشورات على الإنترنت باللغة العربية.

أصبح وجود الاقتباسات القرآنية في المقالات على الإنترنت أكثر وضوحا.

تستخدم المستندات عبر الإنترنت علامات التشكيل الكاملة ورموز التجويد لجميع الكلمات.

لذلك ، فإن التحقق من النص المقتبس للقرآن وعلامات التشكيل والرموز التجويدية أمر ضروري بالنسبة للآية التي تتحقق من صحتها.

من المهم أن تتم كتابة الآيات المقتبسة دون إخفاق وأن يتم توثيق الآيات القرآنية لضمان عدم تغيير كلمة الله حتى إزاحة واحدة أو إهمال خطاب أو علامة التشكيل تعتبر النص غير صالح.

في هذا العمل ، سوف ندرس أساليب وتقنيات مصادقة القرآن الكريم ونقدم طلباً لمصادقة الآيات القرآنية (حفص) على تنسيق النص باستخدام الرمز المميز ومعالجة اللغة الطبيعية (*NLP*).

الكلمات المفتاحية :معالجة اللغة الطبيعية ، المصادقة ، الترقيم ، القرآن الكريم (حفص)