



Deposited via The University of Leeds.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/89831/>

Version: Accepted Version

Proceedings Paper:

Abbas, N and Atwell, E (2013) Annotating the Arabic Quran with semantic web content tags. In: Proceedings of WACL'2 Second Workshop on Arabic Corpus Linguistics. WACL'2 Second Workshop on Arabic Corpus Linguistics, 22 Jul 2013, Lancaster University, UK.

This is an author produced version of a conference paper published in the Proceedings of WACL'2 Second Workshop on Arabic Corpus Linguistics

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

Annotating the Arabic Quran with semantic web content tags

Nora Abbas
Cumbria Police

noorhanabbas@yahoo
.co.uk

Eric Atwell
University of Leeds

e.s.atwell
@leeds.ac.uk

1 Introduction

Islamic tradition holds that Muslims should learn to read, understand, and appreciate the Quran in its original Classical Arabic, since any translations into other languages may propagate (mis)interpretations. The two billion Muslims worldwide who study the Classical Arabic Quran may want to examine and understand meanings of the text verse-by-verse, examining the Arabic verse alongside its translation into a known language, eg English. They may also want to search for **concepts** rather than specific words or letter-sequences, and find all the verses which involve this concept; the concept may be stated in English or Arabic, as long as the search leads to the relevant source Arabic verse(s).

Our semantic web search-by-concept approach, implemented in the Qurany website, demonstrates novel features to improve recall and precision when searching for concepts. The system accepts concept search terms in Arabic and/or English, and the underlying corpus includes original Arabic and eight English translations to cope with variant translations and transliterations, eg Jibril / Jibrail / Gabriel. The system can also lemmatize the concept search terms and offers the user a choice of synonyms from WordNet, to broaden their query and improve recall further. To allow search for abstract concepts, such as 'Juridical Regulations' or 'Women's Property', the Quran corpus is augmented with an ontology or index of key concepts, taken from a recognized expert source 'Mushaf Al Tajweed'. This contains a comprehensive hierarchical index or ontology of nearly 1200 concepts in the Quran. Scholars can use the online Qurany ontology browser to identify a precise concept and find the verses which allude to this concept, with higher precision. Alternatively, researchers can use the standard Google search engine on our website holding a tagged version of the Quran, where each verse is a separate web-document annotated with concept semantic tags; so finding documents to match the Google search terms delivers the verses containing the sought concepts. The concept semantic tags are concepts taken from a recognized expert Quran commentary; so should not be subjective or controversial.

The method can be extended to other Classical Arabic texts. For example, we have added a set of example Haddith verses (other Islamic religious texts), in Arabic and English translation, to our Quran search-by-concept online demonstrators.

2 Quran concept ontology browser

You can search for verses in the Arabic Quran containing a specified concept by using our online Quran concept ontology browser:

<http://xyzqurany.appspot.com/>

For example, if you choose from the tree of concepts on the left hand side, and select:

the concept "Pillars of Islam"

then, "The Prayers"

then "Performing the Prayers"

then "Friday Prayers"

...you will get Quran verses on this topic in the upper right frame, and Hadith verses on this topic in the lower right frame.

3 Google search for concepts in the Quran

The qurany.appspot.com implementation relies on search of an underlying Quran database on the website; this can be slow at times. We know that Google provides very efficient web-search, so it could be better to utilise the Google search engine. To achieve this, we transformed the Quran database including original Arabic, English translations and concept ontology into a large set of HTML web-pages, one per verse. Each Quran verse is stored as an HTML file:

Header: The University of Leeds Qurany Tool
Noorhan Abbas & Eric Atwell
Chapter Name in English
Verse Number in English
Chapter Name in Arabic
Verse Number in Arabic
Arabic text of verse
8 English translations of verse
List of Concepts/Themes Covered in Arabic
List of Concepts/Themes Covered in English
Navigation links: Home Prev Next
Footer: summary info, with links to.appspot.com

The set of HTML files were uploaded to a website <http://www.comp.leeds.ac.uk/nora/> which can be directly searched in Google using the "site:" constraint to restrict search to the website. For example, Google search for:

Friday prayers site:<http://www.comp.leeds.ac.uk/nora/>

Will return a set of verses about Friday Prayers.

Appendix: A sample search-by-concept

The screenshot displays the 'Quran Concepts' application window. On the left is a sidebar menu with a tree structure of topics. The main area on the right shows search results for the concept of 'Gardens of Eden'.

Search Results:

- YUSEFALI:-** To dwell therein. The promise of Allah is true: and He is Exalted in Power, Wise.
- سورة فاتر** (Surah Fatiha)
- جَنَّاتُ عَدْنٍ يَدْخُلُونَهَا يُحَلَّوْنَ فِيهَا مِنْ أَسَاوِرَ مِنْ ذَهَبٍ وَيُؤْتَوْنَ فِيهَا خَيْرٌ 33**
- Fatir**
- 33**
- KHAN:-** 'Adn (Eden) Paradise (everlasting Gardens) will they enter, therein will they be adorned with bracelets of gold and pearls, and their garments there will be of silk (i.e. in Paradise).
- MAULANA:-** Gardens of perpetuity, which they enter – they are made to wear therein bracelets of gold and pearls, and their dress therein is silk.
- PICKTHAL:-** Gardens of Eden! They enter them wearing armlets of gold and pearl and their raiment therein is silk.
- RASHAD:-** They will enter the gardens of Eden, where they will be adorned with bracelets of gold and pearls, and their garments in it will be made of silk.
- SAWAR:-** They will enter the gardens of Eden wherein they will be decked with bracelets of gold, pearls, and silk garments.
- SHAKIR:-** Gardens of perpetuity, they shall enter therein; they shad be made to wear therein bracelets of gold and pearls, and their dress therein shall be silk.
- SHERALI:-** Their reward will be Gardens of Eternity. They will enter them and will be adorned therein with bracelets of gold and pearl; and their garments therein will be of silk.
- YUSEFALI:-** Gardens of Eternity will they enter: therein will they be adorned with bracelets of gold and pearls; and their garments there will be of silk.

Left Sidebar Menu:

- The Holy Quran Search By Topic Tool
- Search Quran for a Word
- أركان الإسلام (Pillars of Islam)
- الإيمان (Faith)
- الأنبياء والرسل (The Prophets and the Messengers of Allah)
- الإيمان بالله (Belief In Allah)
- الغيب (The Unseen World)
- الأعراف (Al'Araf)
- الإيمان بالغيب (The Belief in the Unseen World)
- الجن (The Jinn)
- الجنة (The Paradise)
- أسمائها (It's Names)
- أصحابها (The Companions of the Garden)
- صفاتها (It's Characteristics)
- الخلود (The Eternity)
- السحر (Witching)
- الشيطان (Satan, the Devil)
- الغيب النفسي (The Psychological Secret)
- النار (Hell-Fire)
- الكتب السماوية الأخرى (The Other Divine Books)
- الله - جل جلاله (Allah, Glory to his Majesty)
- المؤمنون (The Believers)

Windows Taskbar:

- RADIO STATION - ...
- QuranSearchConc...
- *Python Shell*
- Quran Concepts
- temparabic - Micr...
- Microsoft PowerPo...
- EN
- 15:24