AENSI Journals

**Australian Journal of Basic and Applied Sciences**

ISSN:1991-8178

Journal home page: www.ajbasweb.com

# Arabic Dictionary Application on Ubuntu

[1]Muhammad Nasir Ibrahim, [2]Siti Noormaya Bilmas, [3]Mariani Idroas

[1]*Universiti Teknologi Malaysia, Faculty of Electrical Engineering,81310 UTM Johor Bharu. Malaysia.*
[2]*Universiti Teknologi Malaysia, Faculty of Electrical Engineering,81310 UTM Johor Bharu. Malaysia .*
[3]*Universiti Teknologi Malaysia, Faculty of Petroleum and Renewable Energy,81310 UTM Johor Bharu. Malaysia*

**ARTICLE INFO**

***Article history:***
*Received 30 September 2014*
*Received in revised form*
*17 November 2014*
*Accepted 25 November 2014*
*Available online 6 December 2014*

***Keywords:***
*Ubuntu; Arabic dictionary; Masdar; language learner;*

**ABSTRACT**

Ubuntu is the popular open source operating system (OS) from Linux distribution. It is an easy platform to be manipulated for creating a new application. But there is small amount of Arabic dictionary application for this platform available. Therefore, this application is developed to make it easy to be used and accessed with compact content of information. Other than interactive interface, this dictionary database is designed with a new approach; the Arabic word is arranged in alphabetical order according to its masdar (root word). Masdar information is important in Arabic language for the construction pattern for each word to result derivation words. Besides, this database use actual dictionary approach which gives the full meaning of the word that will help in such a way on how to use the word. Besides meaning, form and example, this database also consist of grammatical information such as fi'il madi, fi'il mudhari' and fi'il ammar. This dictionary, available in Arabic to English and Arabic to Malay is suitable for student and language learner.



**To Cite This Article:** Muhammad Nasir Ibrahim, Siti Noormaya Bilmas, Mariani Idroas, Arabic Dictionary Application on Ubuntu. ***Aust. J. Basic & Appl. Sci.,*** *8(23): 219-224, 2014*

## INTRODUCTION

Previously, dictionary is only available in printed form. As the technology changes, electronic dictionaries are invented. Due to the easy access and portability, an electronic dictionary becomes highly demanded by the user rather than printed dictionary [Aladdin *et al*., 2004]; [Midlane, 2005].

Dictionaries can be classified into various categories and this can be called as typology. An example of electronic dictionary typology has been discussed by Lehr (1996: 315), which focuses on technical and Meta (lexicographic) evaluation. Based on technical evaluation, this author distinguishes between online or offline dictionaries. In 2003, Gilles propose a three step typology based on "WHO accesses WHAT WHERE?" question. Then Yu *et al.* (2010) made an improvement version of the typology based on information technologies changes. Here a new improvement of typology is proposed as the rapid technological changes as shown in Figure 1.

There are four main categories of users' devices; CD/DVD, PC (personal computer), mobile and portable electronic device (PED). These types of devices are portable and the accessibility of the application depends on the specification of the device. Basically, a dictionary producer will include CD/DVD that consists of material or dictionary application in their printed dictionary [Nesi, 1999]. PC is a device that will be used to run the CD/DVD or to download the application or to use any dictionary application online. Mobile device is referred to the hand-phone or smart-phone which can be used as device to run the dictionary application [Meurant, 2007]. Lastly, PED is referred to the stand alone machine that have only specific functions such as pocket dictionary [Chen, 2011].

There are three types of electronic database. First type is called digital version; the scan material of printed dictionary. It is usually in PDF file format, so that the user will look up the word – meaning in the same manner as printed dictionary. Second type is digitized database; the dictionary data is transferred into database management system software. The user can download the database but they need to look up the word – meaning in a way according to the database management system format. This means no interface is given for this type of application. The third type is called search engine; a complete application where a user just needs to type the word to search for the meaning [Yu *et al.*, 2010].

**Corresponding Author:** Muhammad Nasir Ibrahim, Universiti Teknologi Malaysia, Faculty of Electrical Engineering, 81310 UTM Johor Bahru, Malaysia.
Ph: (+607) 5535260. E-mail: mnasir@fke.utm.my

The important part of electronic dictionary is database. The database source comes from two sources. First, the database is based on printed dictionary content. Second, the database based on the data that are gathered by the developer from their own sources.

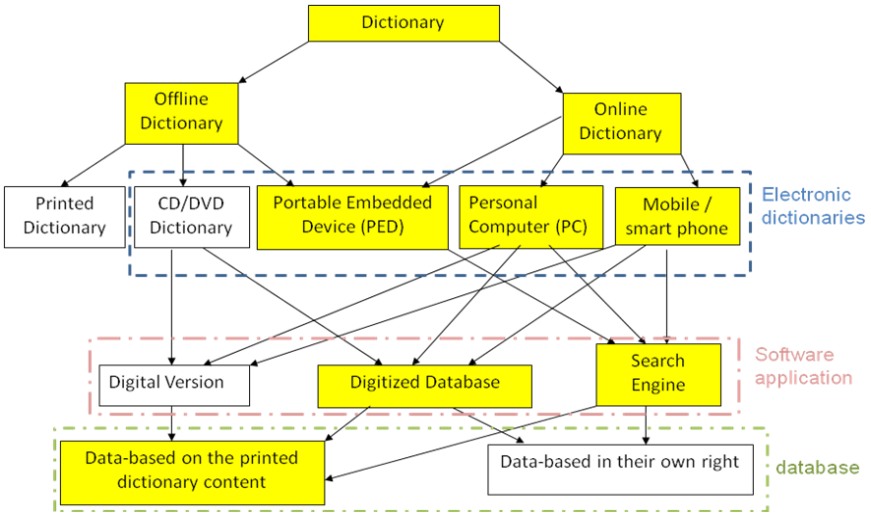


**Fig. 1:** Dictionary typology.

***Arabic Language:***

Generally, Arabic words classified into three main categories; noun (اسم – Ism), Verb (فعل – Fi'il) and Particle (حرف – harf). Noun is a word that indicates a meaning that is not associated with time. Next is verb, a word indicating a meaning that is associated with one of the 3 dimensions of time (past, present and future). The verb for past tense is called as fi'il madi, present tense is called as fi'il mudhari' and the future tense that called as fi'il mudhari' mustaqbal. The examples are He did…., He does…., He will do……. . From fi'il mudhari', fi'il ammar will be form. Fi'il ammar is command verb that will be varied according to the root word.

Arabic verb is based on root – pattern system [Wightwick *et al*.,2008]. Masdar or root word is usually consisting of three (3) consonants. Verb is expansion of the root word. Some expansions are lexical derivations which will result "new words". But others are variation of the verb's conjugation. Western scholars have assigned Roman numerals to the various patterns of derivation, which are called "forms". Also, the root is designated "Form I". A particular form does not have a consistent meaning across verbs, although it has a "usual" meaning. In addition, no root verb has all the derivations. The forms numbered beyond Roman numeral X are rare and obsolescent, for which reason many elementary grammars omit them. Then, for the derivation, it derives from the root word. The meaning of the derivation word has much similar meaning to the root word meaning. As for particle, it indicates a meaning in other than itself or can be called as preposition. Examples of particle are in, who, if and when.

***Operating System:***

Operating system (OS) is a software layer in a computer which schedules all the operations and manages all the computing resources. Basically OS will make sure the system has a certain initial setup during power on, makes appropriate access checks and requested facility or service is either mode available or denied, schedule the demand on system resources, records the sequence of events that lead to the error condition and supports operations and services that require communications over a network [Pramod, 2007].

Currently, there are four OS that commonly used for computer system; Windows, Linux, Mac OS and Ubuntu. Windows are developed by Microsoft which is designed with pleasant and convenient interface. It is also use of software components facilitated sharing objects among applications. The main advantage of Linux is its cluster mode of operation that can be use as servers and search engines. It is operate in multi-processor environment, free operating and distribution, secure because not many virus attack and good for basic function. Mac OS is designed especially for Macintosh computer that produced by Apple [Mark, 2008]. This OS is use graphical approach to do computing with the appearance.

Ubuntu is a part of Linux group project which is based on Debian GNU. It is a Linux desktop creation OS which is free and open source software so that it available to everybody on the same terms. Ubuntu has put a stronger emphasis on usability, accessibility, as well as expansion into other spheres like mobile. Due to these reason Ubuntu has become most popular OS to be download among Linux distribution product [www.ubuntu.com]. Other than that, Ubuntu does not pre-install with proprietary software but it allows the user to download some popular proprietary software such as Flash and Wi-Fi drivers with a simple installation

process. Ubuntu versions for both desktop and server are supported for 9 months while every two years, Ubuntu releases an LTS version (Long Term Support), which is guaranteed support for 5 years. Therefore, Ubuntu is one of the OS that usually choose to develop an open source project.

***Existing Arabic Dictionary Application:***

Dictionary application, especially Arabic electronic dictionary development on Ubuntu is merely in small amount. There is a ready dictionary is provided on the system called freedict package. User can find the dictionary in Application → Office → Dictionary. User need to install the dictionary server into the local Ubuntu machine. Then, the dictionary can be used offline. But the information offers are limited. Other than the ready package, there also application called aspell-ar. This application more to the spell checker application.

As for research study, there are several research that really involve the development of Arabic dictionary development. In 2010, Abd. Rahman *et. al*. introduced a project on online multimedia Malay – Arabic dictionary (OMMAr) for beginner. Besides the meaning of a word, this dictionary is completed with visual such as picture or video and pronunciation voice. This application used client server technology. A mobile device (Smart phone) is the client and a web server is facilitating to lookup into the database. Basically, this project is the improvement from the previous project. Therefore, the softwares that have been used are the same. The experiments have been done on Symbian (Nokia E63), iPhone (using Safari) and Windows mobile (mobile IE on HTC Touch) as the execution platform. There are positive results for the experimented platform, except for the iPhone where the pronunciation function is become a problem due to the Flash format is used. IPhone cannot support Flash application and this problem is still not resolved.

Then, Halperm in 2011 from CJKI (Chinese Japanese Korean Institute) has introduced a printed Arabic-English dictionary called CJKI Arabic Learner's Dictionary (CALD). This Arabic-English dictionary has been made into electronic version for the mobile format by name iCALD. iCALD is designed along with the Arabic verb conjugator (CAVE) application and phonemic transcription system name as CJKI Arabic Romanization System (CARS). This electronic Arabic dictionary is for the smart phones that use iOS and Android platform. CAVE provides exhaustive coverage of 170 conjugation paradigms for over 1600 common verbs. There are multiple search mode available for this system; Full Form Search allows the user to input any inflected form of the verb, Verb Search enables searching from the canonical form, Root Search allows the user to input the root to find the canonical forms of all the corresponding verbs and English Search allows the user to quickly find a verb from any of its English meaning, including partial and included matches. CARS is developed to helped learners to pronounce Arabic accurately and with ease. These features are designed to full fill learners need by providing richness of content and functional user interface that give the learner rapid access to detail information on all inflected form.

***Arabic Dictionary Application Design:***

Each dictionary has advantages and disadvantages. There are no perfect dictionaries but the main thing that is a must for a dictionary is their content and usage to the user [Varantola,2002]. Therefore, for this Arabic dictionary the database is designed using root word as the main element. Ordinarily in Arabic dictionary, the Arabic words will be arranged in alphabetically order. But for this dictionary the Arabic words is arranged according to their root words alphabetically. A RootID is created for each root word as the primary key to relate to their derivation words and form.

This dictionary offer Arabic to English and Arabic to Malay. Thus, the database is divided into three parts; Arabic data, English data and Malay data. Arabic data consist of the Arabic words. English and Malay parts generally consist of the meaning data and other related information.

The development of this dictionary is used XAMPP. XAMPP is Apache distribution open source software that consist MySQL, PHP and Perl. The database is organized using PHPMyAdmin underlying of MySQL. The interface is designed using PHP (Hypertext Pre-processor) scripting language.

For the interface, it is offer two search modes. Root word search method is offer for the user who has a basic knowledge of Arabic language. Then, for any word search method is for general user. The flow of search process is show in Figure 2.

***Result:***

The interface of this application is designed to be simple, attractive and user friendly. Other than that, the most important is the information offer by the application and the accessibility. For this application, the interface that serve to the user can be summarize as in Figure 3. Firstly users need to choose their preferred dictionary. Then they can start use the dictionary by choosing the next operation.

Basically, when user chooses to find by root word, they need to enter the root word. Once the 'SEARCH' button is click, the related information will be displayed. The frame with the example of the information display of root word searching mode is as shown in Figure 4.

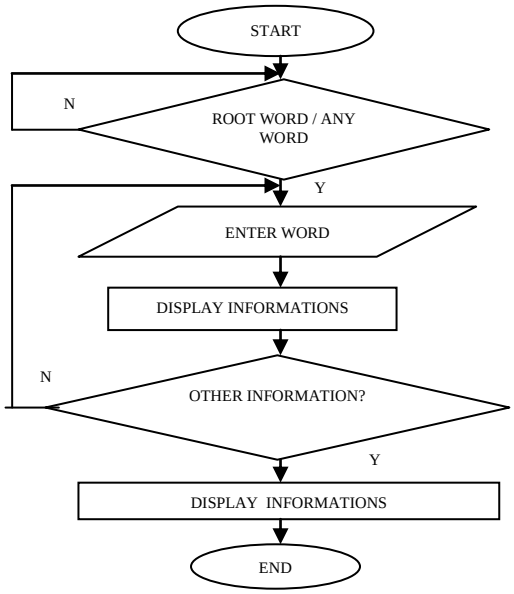


**Fig. 2:** Flowchart of searching word process.

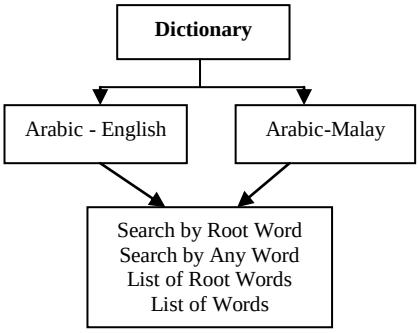


**Fig. 3:** Interface option.

| | |
|---|---|
| **Root Word** | كتب |
| **Fi'il Madi** | *kataba* |
| **Fi'il Mudhari'** | *Yaktubu* |
| **Fi'il Ammar** | *Uktub* |
| **Meaning** | *To write, pen, write down, put in writing, note down, inscribe, enter; record,book,register.* |

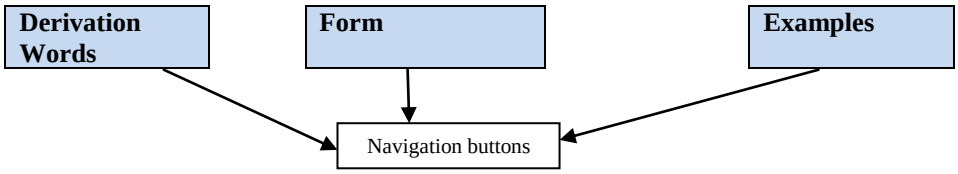


**Fig. 4:** Frame of information display for root word searching mode.

If one the navigation button is click, further information will be displayed. As the 'Derivation Words' button is clicked, all the derivation words will be displayed along with their information. It is same process for other navigation button. The frame of the display is shown in Figure 5.

It is similar with searching any word mode. Users just need to enter the word they need, then the information will be displayed. If the word that user enter belong to root word, the display format as in Figure 4 will be displayed. The interface for derivation words will be displayed as shown in Figure 6. The format is similar to root word but the information given are differs.

When navigation button of 'Root Word' is click the information display is same to Figure 4 except the navigation provided only left 'Form' button. Once 'Other Word' button is click, the display format will be same as Figure 5(A). Same to 'Example' button, the format will be same as Figure 5(C).

| (A) | | (B) | | (C) | |
|---|---|---|---|---|---|
| Derivation Words | | Form | | Examples | |
| | | | | | |
| **Word:** | كتاب | **Form:** | *II* | **Example:** | كتب كتابه |
| **Word Category:** | *noun* | **Meaning:** | *To accuse of heresy* | **Sound:** | *Kataba kitabahu* |
| **Sound:** | *kitab* | | | **Meaning:** | *To draw up the marriage contract* |
| **Female/Male:** | *male* | **Form:** | *III* | | |
| **Meaning:** | *Piece of writing, record, paper; letter, note, message; document; deed; contract; book;* | **Meaning:** | *To follow; to keep one's mind or eyes on; to agree, concur, be in agreement or conformity; to pursue, chase, follow up; to continue, go on* | **Example:** | كتاب تعليمى |
| | | | | **Sound:** | *Kitab ta'limi* |
| **Word:** | كتبي | . . . | | **Meaning:** | *Textbook* |
| **Word Category:** | *noun* | | | . . . | |
| **Sound:** | *kutubi* | | | | |
| **Female/Male:** | *male* | | | | |
| **Meaning:** | *Bookseller; bookdealer;* | | | | |
| . . . | | | | | |

**Fig. 5:** Frame of the information display.

| | |
|---|---|
| **Word** | مكتب |
| **Word category** | *noun* |
| **Sound** | *maktab* |
| **Female/Male** | *male* |
| **Meaning** | *Office; bereau; business office; study; school, elementary school; department, agency, office; desk;* |

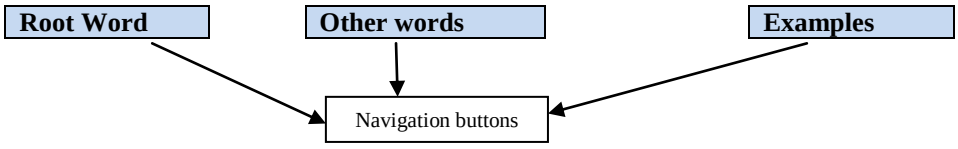


**Fig. 6:** Frame of information display for any word searching mode.

***Discussion and Conclusion:***

This Arabic dictionary application is developed for Ubuntu. It is an offline system which is the user just need one time download and it can be used. This dictionary offer two type of dictionary; Arabic to English and Arabic to Malay dictionary. The unique about this application is the database.

Database for this application is arranged according to the root word. It is important in order to know the pattern of each Arabic word structure. Besides, this database also keeps track the total entry of each root word which is a good reference to developer. This database structure is designed for multilingual target language. If there are more other languages need to be added into the database, they only need to add other language table with similar existing table format without affected the other old language data. This database is the first step to

other application. In future this database can be use to link to other application such as Al-Quran dictionary and language learning application. As the purpose of public sharing this database can also be a reference for other related Arabic research such as natural language and pattern recognition.

As for the interface, it is designed to be simple, interactive and user friendly. The information displayed in arrangement manner which is easy to understand and neat. This dictionary is suitable to be use for all language learners because of its compact information. Besides it is really help user to learn and do reference easily because it can be installed in any Ubuntu devices.

## REFERENCES

Abd. Rahman, K., A. Abd Aziz, S. Amirrudin, M.J. Ghazali, C.W. Shamsul Bahri, C.W. Ahmad and M.H. Husaini, 2011. The Design and Implementation of the Malay – Arabic Online Multimedia Dictionary on Smart Phones, *Proceeding of Regional Conference on Knowledge Integration in ICT 2010*, pp 179-188.

Aladdin, A., A. Hamat and M.S. Yusof, 2004. Penggunaan PBBK(Pembelajaran Bahasa Berbantukan Komputer) dalam pengajaran dan pembelajaran Bahasa Arab sebagai bahasa asing: Satu tinjauan awal. *GEMA Online Journal of Language Studies*, 4(1).

Chen, Y., 2010. Dictionary Use And EFL Learning. A Contrastive Study Of Pocket Electronic Dictionaries And Paper Dictionaries. *International Journal of Lexicography*, 23(3): 275-306. Oxford University Press.

Gilles-Maurice de Schryver, G.M., 2003. Lexicographers' Dreams in the Electronic Dictionary Age, *International Journal of Lexicography*, 16(2): 143-199.

Halpern, J., 2011. Pedagogical Lexicography Applied to Arabic Dictionaries and Smartphone Applications, *ALTIC -2011*, Eqypt.

I.M.A.M (Imam Mahdi Association of Marjaeya), 2009. *Arabic Grammar for the Holy Quran*. Unpublished Notes. Al-Qaem Institute.

Mark L. Chambers, 2008. *Mac OS X Leopard TM All- in-one Desk for Dummies References*. Wiley Publishing Inc., Canada, pp: 10.

Meurant, R.C., 2007. A Preliminary Survey of the Use of Cell Phones, Electronic Dictionaries, SMS, Email, Computers and the Internet by Korean College EFL Students with Respect to Patterns of L1:L2 Language Use and the Associated Language Learning Strategies Used in Accessing Online Resources, *2007 International Conference on Multimedia and Ubiquitous Engineering (MUE'07).*

Midlane, V., 2005. *Students' Use of Portable Electronic Dictionaries in the EFL/ESL Classsroom; A Survey of Teacher Attitudes*. Master Thesis. University of Manchester.

Nesi, H., 1999. A User Guide to Electronic Dictionaries for Language Learners, *International Journal of Lexicography*, 12(1): 55-66.

Pramod Chandra P. Bhatt, 2007. *An Introduction to Operating Systems Concepts & Practice*. 2nd Edition. Prentice Hall India, pp 6.

Varantola, K., 2002. Use and Usability of Dictionaries: Common Sense and Context Sensibility?. In M.H. Correard (ed.) *Lexicography and Natural LanguageProcessing: A Festschrift in Honour of B. T. S. Atkins*. Grenoble, France: EURALEX, 2002.

Wightwick, J. and M. Gaafar, 2008. *Arabic Verbs & Essentials of Grammar*. (2nd Ed). USA: McGraw Hill.

Yu, P.F. and J.L. DU, 2010. Towards the Electronic Dictionaries in Modern Commercial Context: A Comparative Perspective. *Third International Symposium on Electronic Commerce and Security*, 29-31 July.