Journal of Fundamental and Applied Sciences

ISSN 1112-9867

Available online at http://www.jfas.info

Research Article

Special Issue

TAGGING NARRATOR'S NAMES IN HADITH TEXT

N. A. Rahman[1], N. K. Ismail[1], Z. M. Nor[2], M. N. Alias[3], M. S. Kamis[4] and N. Alias[1,*]

[1]Faculty of Computer and Mathematical Sciences, UniversitiTeknologi MARA, 40450 Shah Alam, Selangor, Malaysia

[2]Faculty of Quran danSunnah Studies, UniversitiSains Islam Malaysia, 71800 Nilai, Negeri Sembilan, Malaysia

[3]Faculty of Islamic Studies, UniversitiKebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia

[4]PERMATApintar College, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia

Published online: 17 October 2017

ABSTRACT

Text document expresses enormous sort of information but it lacks the imposed structure of a traditional database. Unstructured data, particularly free running text data has to be transformed into a structured data. Extracting information from text is part of NLP process. The implementation of the NER algorithm for NLP is normally influenced by the domain of the studies. Besides, there is no existing system that is designed to detect the types of named entity in hadith text, develop POS tags and rule based extraction for narrator's name in Hadith Text in the Malay language. The POS tags were developed from 1000 hadith texts. The POS tags were created involving a total of 256 words which is part of narrator's names. The rule based was developed to determine five types of narrator's chain. Further research will determine the relationship between each narrator and the construction of narration's chain.

Keywords: tagging; hadith text; name.

Author Correspondence, e-mail: syahidah@uitm.edu.my

doi: http://dx.doi.org/10.4314/jfas.v9i5s.21

1. INTRODUCTION

In [1] stated that Natural Language Processing (NLP) is an artificial intelligence branch which has the ultimate goal to invent theories, discover techniques and build software that can understand, analyze and generate the nature of human languages in order to interface with computers; both in written and spoken contexts using natural human languages, so NLP gives computers the ability to understand the way humans learn and use language and it is the most challenge inherent in natural language processing. The NLP techniques parse linguistic input (word, sentence, text, dialogue) according to the rules (derivational rules, inflectional rules, grammatical rules, etc.) and resources (like lexicon, corpus and dictionary) of the target language.

Text document expresses enormous sort of information. Unfortunately, it lacks the imposed structure of the traditional database. Therefore, unstructured data, particularly free running text data has to be transformed into a structured data [2]. The NLP research in texts which based other languages usually accepts the predetermined NLP. However, this has resulted in setbacks because the existing NLP needs to be suitable for other languages for its different structure and syntax [3-4].

The NLP research for Malay language has been carried out in the form of [5-7], stemming [8], part of speech tagging [5] and parsing [9-11]. However, there were some NLP tasks for Malay language which is yet to be explored in research such as the language recognition, stop word removal, word sense disambiguation. All the aforementioned are using the rule-based approach to execute the NLP task. In addition, the NLP research was implemented to analyze the sentence at various level of complexity. This research aimed for the Malay language, which is on Morphologic [8], Lexical [3,12], Syntactic [13-14] except for semantics.

Extracting information from text is part of NLP process [15]. However, there was no existing system that is designed to detect types of named entity in the Malay language until it is done by their group of research for the news article. The implementation of the NER algorithm for NLP is normally influenced by the domain of the studies [16]. Moreover, there was no existing system that is designed to detect types of named entity in hadith text.The Table 1 shows the study which was conducted by [17] that differs a number of domain and hadith that have been researched, NLP tools or approaches, classification algorithm were used along with

the results produced. The number of hadith used is different between one another.

Table 1. Differentiating linguistic tools of approaches [17]

Reference	#Domains	#Hadiths	Linguistic Tools/Approaches	Classification Algorithm	Results
Harrag et al. (2008;2009)	14	453	Stop-word removal and rule-based morphological stemming	Decision trees, Bayesian, Entropy and Vector space models	F1-measure = 0.70 with decision trees
Harrag et al. (2011)			Three stemming approaches; rule-based; root based and light stemming	ANN vs. VSM	F1-measure = 0.5 with ANN + light or rule-based stemming
Alkhatib (2010)	8	1500	Removing chains of narrators, stop words and affixes	Rocchio, K-NN, Naive Bayes and SVM	Recall = 100% Precision = 63.36% (SVM) and 67.11% (Rocchio)
Al-Kabi and Al-Sinjilawi (2007)	12	80 (for testing)		Vector Space Models with several similarity measures	F1-measure: from 0.42 (Dice Factor) to 0.85 (Naïve Bayesian)
Jbara (2010)	13	1321	Removing chains of narrators, stop words and affixes Stem-based, Word	The cosine coefficient	49% and 37% of improvement in F-measure for hybrid

	based and hybrid representation	method compared to the word-based and the stem-based methods

This paper focuses only on tagging the narrator's names in hadith texts, to extract the names in order to form the narrator's chain. The hadith texts that we use are in the Malay language. The structure of hadith text will be discussed in the next section.Section 2 is the discussion on the development of tagging and entity recognition process.Section 3 discusses the methodology used for hadith texts. Section 4 is the conclusion.

2. RESULTS AND DISCUSSION

A total of 1000 hadith texts were used as data in this process. In the first step, the raw of text document is split into narrator's chain and content by identifying the last narrator before *Nabishallallahu 'alaihiwasallam*or*Rasulullahshallallahu 'alaihiwasallam*. This task is performed manually by hadith experts.

There are four types of narrator's chain part in hadith text as show in Table 2. The narrator's names were written in bold font.

Table 2. Types of narrator chain part in hadith text

Type	Narrator Chain Part in Hadith Text
1	ShahihBukhari 1: Telahmenceritakankepada kami **Al Humaidi Abdullah bin AzZubair**diaberkata, Telahmenceritakankepada kami **Sufyan** yang berkata, bahwaTelahmenceritakankepada kami **Yahya bin Sa'id AlAnshari**berkata, telahmengabarkankepada kami **Muhammad bin Ibrahim At Taimi**, bahwadiapernahmendengar**Alqamah bin Waqash Al Laitsi**berkata; sayapernahmendengar**Umar bin Al Khaththab** …
2	ShahihBukhari 131: Telahmenceritakankepada kami **Adam**berkata, telahmenceritakankepada kami **Ibnu Abu Dzi'b**dari**Nafi'**dari**Ibnu 'Umar** dariNabishallallahu 'alaihiwasallam, **dandari**AzZuhridari**Salim**dari**Ibnu**

'UmardariNabishallallahu 'alaihiwasallam, bahwa ..

3 ShahihBukhari 134: Telahmenceritakankepada kami **'Ali**berkata, telahmenceritakankepada kami **Sufyan**berkata, telahmenceritakankepada kami **AzZuhri**dari**Sa'id bin Al Musayyab**. (dalam**jalur lain**disebutkan) Telahmenceritakankepada kami **'Abbad bin Tamim**dari**Pamannya,**…

4 ShahihBukhari 908: Dan (masihdarijalurperiwayatan yang samadengan**haditssebelumnya)**dari**Jabir bin 'Abdullah**iaberkata,

If the narrator's names extracted from the Type 1 hadith texts, it will produce the narrator's chain as shown in Fig. 6. Type 2 hadith texts will produce the narrator's chain as shown in Fig. 7. Type 3 hadith texts will produce two narrator's chains for one hadith text. The narrator's chain for Type 4 hadith texts is a continuation of the previous hadith text. For example the narrator's chain in hadith text number 908 is a continuation from hadith text number 907. According to the domain expert, Type 2, 3 and 4 is featured by the underline sentence.

Abaidullah	Ady	Aqadi	Auf	Bukair	Fazari	Hanafi
Aban	Affan	Ar	Aun	Bunani	Fudaik	Hanafiah
Abas	Aflah	Arab	Auza'i	Bundar	Fudlail	Hanafiyah
Abayah	Aghar	A'raj	Awanah	Buraid	Fulaih	Hani
Abbad	Ahdab	Ar'arah	Awaqi	Buraidah	Ghailan	Hanzhalah
Abbas	Ahmad	Arat	Awza'i	Burdah	Ghalib	Hanzhali
Abdah	Ahnaf	Arim	Ayyasy	Busr	Gharair	Harami
Abdan	Ahwal	Arqam	Ayyub	Busyair	Ghasil	Harb
Abdu	Ahwash	Arubah	Az	Darawardi	Ghassan	Harits
Abdul	Aidzullah	As	Azdi	Darda	Ghifari	Haritsah
Abdullah	Aiman	Asad	Azib	Dastawa'I	Ghiyats	Harrani
Abdurrahim	Aisyah	Asadi	Aziz	Daud	Ghundar	Harun
Abdurrahman	Aizar	Ash	Badal	Dimsyqi	Habasyi	Hasan
Abdurrazaq	Akwa	Ashbagh	Bakali	Dinar	Habhab	Hassan
Abdurrazzaq	AL	Ashbahani	Bakar	Ditsar	Habib	Hasyim
Abdus	Al	Ashim	Bakr	Dlamrah	Habibah	Hatim
Abdush	Ala	Aslam	Bakrah	Dlamri	Had	Hatsmah
Abdushshamad	A'la	Aslami	Banani	Dluba'i	Haddad	Hawsyab
Abi	alaihi	Asma	Bani	Dluha	Hadi	Hayyan
Abidah	A'lam	AsSa'idi	Bara	Dukain	Hadrad	Hazim
Abis	Alaqah	Aswad	Barirah	Dzakwan	Hadza	Hazm
Abu	Ali	Asy	Barro	Dzar	Hadzdza	Hibban
AbuBakar	Aliyah	Asy'ari	Barzah	Dzarr	Hafsh	Hilal
Abulshaq	Alqamah	Asy'ats	Basyar	Dzarri	Hafshah	Hind
Abu'mar	A'masy	Asykari	Basyir	Dzi'b	Haiwah	Hindi
AbuMusa	Amir	Asyyab	Basysyar	Fadlal	Hajjaj	Hisam
AbuNadlr	Ammar	At	Bathin	Fadlalah	Hakam	Hishin
Abza	Amrah	Ath	Bazi	Fadll	Hakim	Hisyam
Abzaa	Amru	Atha	Bilal	Fadlol	Halhalah	HisyamBapaknya
Ad	An	Athiyah	bin	Faqir	Hamam	Hizami
Adam	AN	Athiyyah	binti	Faraj	Hamid	Hubaisy
Adi	Anas	Atho	Bisyir	Farqad	Hammad	Hubaiys
Adl	Anbasah	Atho'	Bisyr	Farsi	Hammam	Hudbah
Adullah	Anshari	Ats	Buhainah	Fatimah	Hamzah	Hudzaifah

Fig.1. Tagging narrator's name

The second step, we use regular expressions in Python statement number 1 is to tokenize the narrator's chain part into words. For further explanation please refer to [15]. However, before we apply the regular expressions; we remove "'" and "–" symbols that were used widely in

Malay hadith text to improve the POS tagging process. After the tokenization process, we remove the words consist of only "" symbol to improve the entity recognition process.

From the tags that we developed, it produces 671 words. For example **Al Humaidi Abdullah bin AzZubair** will be separated into 6 words which are **Al**,**Humaidi**, **Abdullah**, **bin**, **Az**and**Zubair**as shown in Fig. 1, 2 and 3. Fig. 4 shows narrators using relationship and additional information as a name.

Hudzifah	Jabal	Khair	Makhramah	Miswar	Muqaddam	Namir
Humaid	Jabar	Khaldah	Makki	Mu'adz	Muqaddami	Nashir
Humaidi	Jabir	Khali	Malih	Mu'adzah	Muqatil	Nashr
Humran	JabirSamurah	Khalid	Malik	Mu'afa	Muqotil	Nauf
Hunaif	Jabr	Khalil	Ma'mar	Mu'alim	Murrah	Naufal
Hunain	Ja'd	Khallad	Ma'n	Mu'alla	Musa	Nu'aim
Hurairah	Ja'di	Kharrabudz	M'an	Mu'allaa	Musadad	Numair
Hurmuz	Ja'far	Khaththab	Manaf	Mu'awanah	Musaddad	Nu'man
Hurru	Ja'fat	Khaththami	Manjufi	Mua'wiyah	Musafir	Qabishah
Husain	Jami	Khaulani	Manshur	Mu'awiyah	Musattab	Qais
Hushain	Jamrah	Khaza'i	Maqbari	Mu'awiyyah	Musayyab	Qalabah
Husyaim	Jarir	Khiyar	Maqburi	Mubarak	Mush'ab	Qa'nab
Huwairits	Jazari	Khubaib	Ma'rur	Mudlar	Mushir	Qa'qa
Huwirits	Jirasy	Khudri	Marwan	Mudrik	Muslim	Qari
Ibnu	Ju'aid	Khudru	Marwazi	Mufadlal	Muslimin	Qashim
Ibnulhab	Jubair	Khushaifah	Maryam	Mufadldlal	Musnadi	Qasim
Ibnul	Ju'fi	Kufi	Maslamah	Mughaffal	Mu'tamin	Qatadah
Ibrahim	Juhaifah	Kuraib	Masruq	Mughirah	Mutamir	Qaththan
Ibrahin	Juhaim	Laila	Mas'ud	Muhabbar	Mu'tamir	Qawariri
Idris	Juhaini	Lailatul	Masyruq	Muhajir	Muthahhar	Qaza'ah
Ikrimah	Juhani	Laits	Mathar	Muhammad	Muthalib	Qilabah
Ima'il	Junda'i	Laitsi	Mawali	Muhammadar	Mutharrif	Qotadah
Imran	Jundub	Ma'an	Mazani	Muharib	Muth'im	Qudamah
Irak	Juraij	Ma'bad	Mazini	Muharibi	Muththalib	Qurasyi
Isa	Jurairi	Madini	Mihran	Mujahid	Mutsanna	Qurrah
Ishaq	Juwairiah	Mahak	Mihshan	Mujmir	Muzani	Qutaibah
Iskandarani	Juwairiyah	Mahbub	Mijlaz	Mukhtar	Muzni	Rabbi
Isma'il	Juwairiyyah	Mahdi	Mikhwal	Mulaikah	Nabi	Rabbih
Israil	Ka'b	Mahmud	Minhal	Munabbih	Nadlar	Rabdzah
Isra'il	Kahmas	Maimun	Minqari	Munba'its	Nadlr	Rabi
Itban	Kaisan	Maimunah	Miqdad	Mundzir	Nafi	Rabi'ah
Iyadl	Katsir	Maisarah	Mis'ar	Munir	Nahdi	Rafi
Iyas	Khabbab	Majisyun	Mishri	Munkadir	Najasyi	Rahawaih
Iyats	Khadij	Makhlad	Miskin	Muntasyir	Najih	raj

Fig.2. Tagging narrator's name

And then, the rules based were developed to assign the tags to the narrator's chain part in the hadith text. Next, the narrator's names recognition process using the Python statement number 2 in Table 3 to identify the names.

Statement 2 in Table 3 is the patterns of narrator's chain exist in hadith text, as shown in Table 2. But, the narrator's name itself also has many patterns as shown in Table 4.

Raja	Shafiyah	Sya'bi	Tsaqafi	Wadlih	Zaid
Rasulullah	Shafiyyah	Syadad	Tsaur	Wahab	Zaidah
Rasyid	Shafwan	Syaddad	Tsauri	Wahb	Zai'dah
Rauh	Shalih	Syadzan	Tsumamah	Wahhab	Za'idah
Rawwad	shallallahu	Syaibah	Tumailah	Wahid	Zainab
Rib'i	Shalt	Syaiban	Ubadah	Wahsyiyyah	Zakaria
Rifa'ah	Shaltu	Syaibani	Ubaid	Wail	Zanad
Sa'ad	Shamad	Syaqiq	Ubaidah	Wa'il	Zaraqi
Sa'd	Shamit	Syarik	Ubaidillah	Waki	Zayadi
Safar	Sha'Sha'ah	Sya'unah	Ubaidullah	Walid	Zinad
Sahal	Sha'sha'ah	Syidad	Ubay	Waqash	Zinadbanwa
Sahl	Shiddiiq	Syihab	Ufair	Waqid	Ziyad
Sa'ib	Shiddiq	Syu'ah	Ulayyah	WaqidMuhammad	Ziyadi
Said	Shubaih	Syu'aib	Umair	Warits	Zubaid
Sa'id	Shuhaib	Syu'bah	Umais	Warqa	Zubaidi
Sa'idi	Shurad	Syuhab	Umamah	Warrad	Zubair
Saif	Simak	Syumail	Umar	wasallam	Zubairi
Sa'in	Siman	Syuraih	Umarah	Washil	Zuhair
Sakhtiyani	Sinan	Taghlib	Umayyah	Washithi	Zuhri
Salam	Sirin	Taim	Ummu	Wasi	Zuhrii
Salamah	Siyah	Taimi	Uqail	Wasithi	Zuhry
Salami	Sufyan	Taimillah	Uqbah	Wuhaib	Zur'ah
Salim	Suhail	Tamim	Urwah	Wuhhaib	Zurai
Salman	Sukain	Tamimah	Urwahl	Yaar	Zuraqi
Sam	Sulaim	Tayyah	Usamah	Ya'fur	Zurarah
Saman	Sulaiman	Thahman	Usmah	Yahya	
Samrah	Sumayya	Thalhah	Utbah	Ya'la	
Sarah	Sumayyah	Thalib	Utsman	Yaman	
Sayyar	Suqah	Thariq	Uwais	Ya'qub	
Sayyarah	Suraij	Thawil	Uwaisi	Yasar	
Shabah	Suramari	Thawus	Uwaisy	Yasir	
Shabbah	Suwaid	Thufail	Uyainah	Yazid	
Shadaqah	Suyan	Tiyah	Wadi'ah	Yunus	
Shafar	Syababah	Tsabit	Wadldlah	Yusuf	

Fig.3. Tagging narrator's name

Relationship	Additional information
bapakku	nama
Bapaknya	aslinya
Bapakku	adalah
Pamannya	dia
pamannya	yakni
bapaknya	adalah
Ibunya	yaitu
Kakeknya	Yaitu
Anak	namanya
anak	mantan
saudara	budak
bibiku	puteri
anak	isteri
ayahnya	kakek
saudaranya	anak
Saudaranya	saudara
Saudaraku	Bani
	bani
	Haritsah
	Nabi
	Ibu
	kaum
	Ummul
	Mu'minin
	Mukminin
	Muk'minin
	Mantan
	sahaya
	anaknya

Fig.4. The name using relationship and additional information

Table 3. Python statements

No	Python Statements
1	perPerkataan = re.findall(r"\w+(?:[-']\w+)*|'|[-.(]+|\S\w*", hadis)
2	grammer = r""" POPeriwayatan: {<NPW>+<PADD>*<NPW>+} {<NPRW>*} {<DUA><DUA>} {<PECAHAN><PECAHAN>} {<SAME><SAME>}

Table 4. Narrator's name pattern

Narrator's Name Pattern in Hadith	Explanation
Al Humaidi Abdullah bin AzZubair	Al Humaidi Abdullah: Name of narrator bin: son of AzZubair: Name of father's narrator
Pamannya	Mean his/her uncles of previous narrator
Sa'iddiaadalahanaknya Abu Sa'id	Sa'id: Name of narrator *diaadalahanaknya*: son of Abu Sa'id: Name of father's narrator
'AmruyaituIbnuMaimun	'Amru: Name of narrator *yaitu*: is IbnuMaimun: Name of narrator

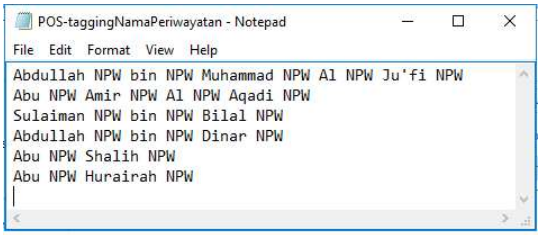
Sentence *"diaadalahanaknya"* means the son of and "*yaitu*" means is/are the additional information, which narrator used to differentiate between narrators that have the same name but different person. Fig. 5 shows hadith text after the POS tags were applied and the name entity extraction process.

Input 1:

ShahihBukhari 8: Telahmenceritakankepada kami Abdullah bin Muhammad Al Ju'fidiaberkata, Telahmenceritakankepada kami Abu 'Amir Al 'Aqadi yang berkata,

bahwaTelahmenceritakankepada kami Sulaiman bin Bilal dari Abdullah bin Dinar dari Abu Shalihdari Abu Hurairah ...

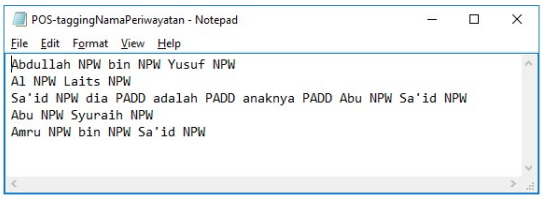
Output 1:



Input 2:

ShahihBukhari 101: Telahmenceritakankepada kami 'Abdullah bin Yusuf berkata,, telahmenceritakankepadasaya Al Laitsberkata, telahmenceritakankepadasayaSa'iddiaadalahanaknya Abu Sa'iddari Abu Syuraihbahwadiaberkatakepada 'Amru bin Sa'id ...

Output 2:



Input 3:

ShahihBukhari 134: Telahmenceritakankepada kami 'Ali berkata, telahmenceritakankepada kami Sufyanberkata, telahmenceritakankepada kami AzZuhridariSa'id bin Al Musayyab. (dalamjalur lain disebutkan) Telahmenceritakankepada kami 'Abbad bin TamimdariPamannya ..

Output 3:

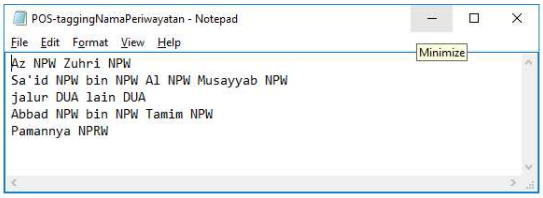


Fig.5. After POS tags and name entity extraction process

3. EXPERIMENTAL

This section will highlight on hadith text structure and the development design for tagging narrator's name in hadith text and extract them. Hadith text structures consist of sanaddanmatn [18] as shown in Fig. 6.

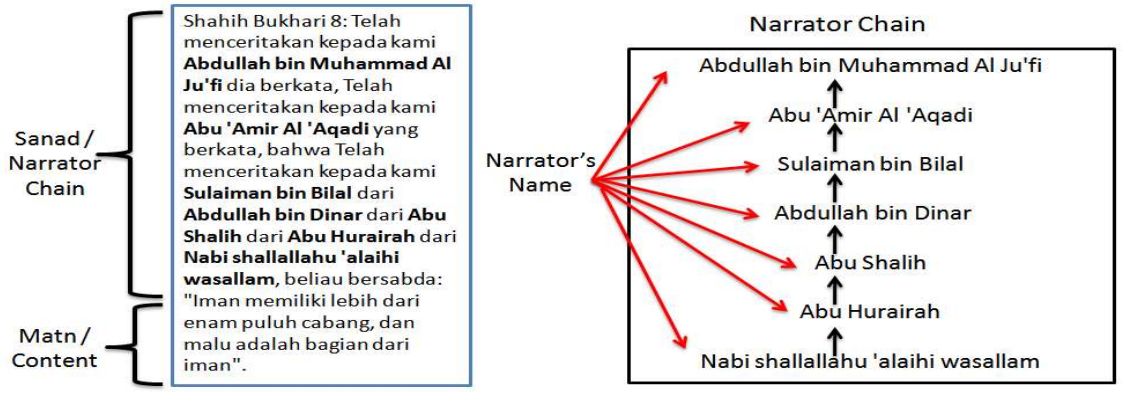


Fig.6. Hadith text in Malay text document

The figure shows the hadith text in ShahihBukhari book number 8, taken from [19]. The database that became the reference in the field of hadith research, education and public utilisation. Each Hadith is composed of two important component: the actual narrative text, known as Content/Matn and the chronological list of people who were the transmitters of the Content/Matn, also known as Narrator's Chain/Isnad [20-22]. Sanad/Narrator's Chain part in Fig. 6 is "*Telahmenceritakankepada kami Abdullah bin Muhammad Al Ju'fidiaberkata, Telahmenceritakankepada kami Abu 'Amir Al 'Aqadi yang berkata, bahwaTelahmenceritakankepada kami Sulaiman bin Bilal dari Abdullah bin Dinar dari Abu Shalihdari Abu HurairahdariNabishallallahu 'alaihiwasallam, beliaubersabda*:". Therefore, the narrator's chain existed in the aforementioned hadith were "Nabishallallahu 'alaihiwasallam → Abu Hurairah → Abu Shalih → Abdullah bin Dinar → Sulaiman bin Bilal → Abu 'Amir Al 'Aqadi → Abdullah bin Muhammad Al Ju'fi". Meanwhile, the content for the hadith in Fig. 6 is for the statement "*Imanmemilikilebihdarienampuluhcabang, danmaluadalahbagiandariiman*". The separation between the content and the narrator's chain in the diagram below are clear and marked with the symbol ":".

Fig. 7 shows that the hadith text has two narrator's chain in one hadith text. The first narrator's chain in the hadith is "*Rasulullahshallallahu 'alaihiwasallam→Ibnu 'Abbas →Ubaidullah bin Abdullah →AzZuhri→ Yunus→ Abdullah →Abdan*". The second narrator's

chain is "*Rasulullahshallallahu 'alaihiwasallam→Ibnu 'Abbas →Ubaidullah bin Abdullah →AzZuhri→ YunusdanMa'mar→ Abdullah→Bisyir bin Muhammad*". Meanwhile, the hadith content is "*Rasulullahshallallahu 'alaihiwasallamadalahmanusia yang paling lembutterutamapadabulanRamadlanketikamalaikatJibril 'Alaihis Salam menemuinya, danadalahJibril 'Alaihis Salam mendatanginyasetiapmalam di bulanRamadlan, dimanaJibril 'Alaihis Salam mengajarkan Al Qur'an. SungguhRasulullahshallallahu 'alaihiwasallamjauhlebihlembutdaripadaangin yang berhembus*". The separation between content and narrator's chain in the hadith was ambiguous and it has been referred to the hadith researchers.

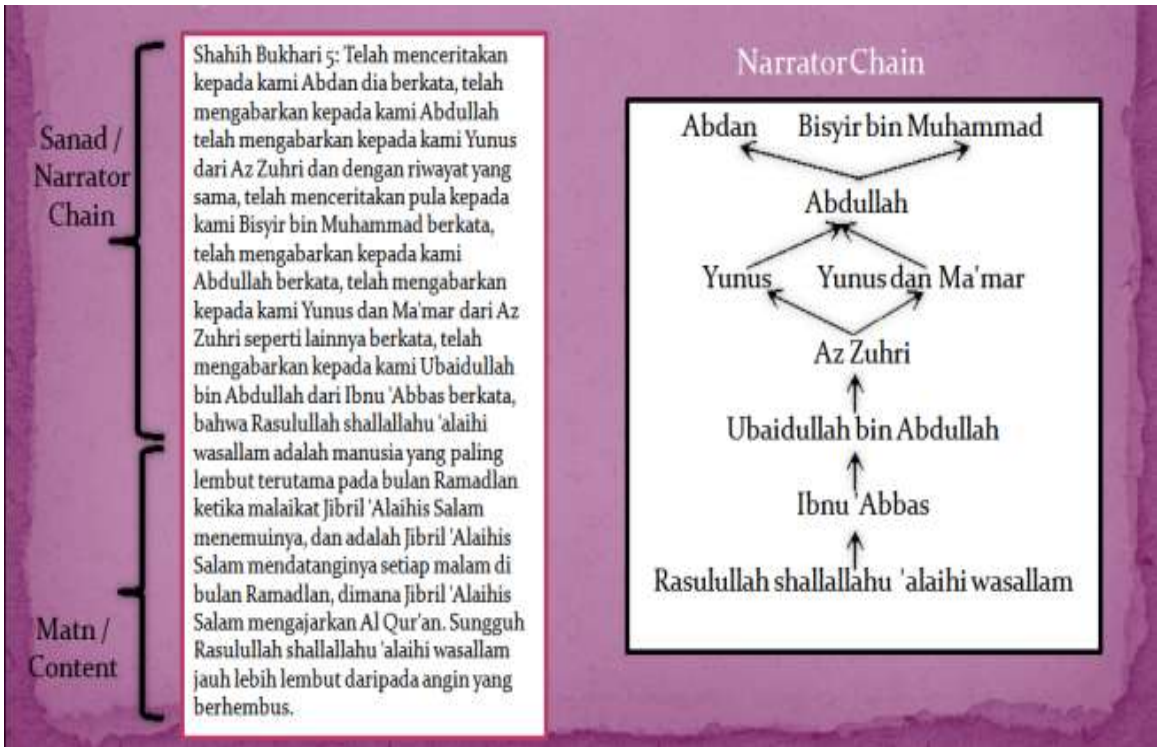


Fig.7. Hadith text with two narrator chain in Malay text document

Next, Fig. 8 shows the architecture for a simple information extraction system [15] that we referred to. First, the raw of the text document is split into narrator's chain and content by identifying the last narrator before *Nabishallallahu 'alaihiwasallam*or*Rasulullahshallallahu 'alaihiwasallam*. This process is conducted with domain expert who is an academician in Islamic Study specific in Hadith study. We processed only the narrator's chain part. The narrator's chain is further subdivided into words using tokenizer.

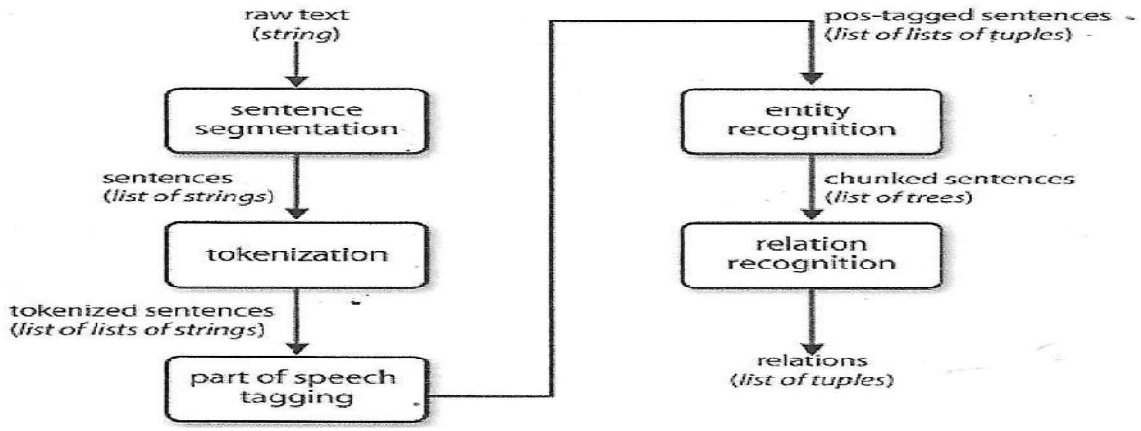


Fig.8. Simple information extraction architecture [15]

Then, the narrator's chain part is tagged with part-of-speech (POS) tags, which will prove to be helpful in the next step, narrator's name entity recognition. Due tothere are no existing POS tags for narrator's name in hadith text, we developed the POS tags. We also refer to a rule based Malay NER framework proposed by [16] to develop a rule based on identifying types of named entity in the Malay language for hadith text. For the final step, we reserved it for future research. The rule based was developed by using Python Programming language [23,15].

4. CONCLUSION

This research developed POS tags and rule based extractions for narrator's name in Hadith Text in the Malay language. The POS tags were developed from 1000 hadith texts. The POS tags were created involving a total of 256 words which were part of narrator's names. The rule based was developed to determine five types of narrator's chain. Further research will determine the relationship between each narrator and the construction of narration's chain.

5. ACKNOWLEDGEMENTS

This research was funded by the Malaysian Government under Fundamental Research Grant Scheme (FRGS) (FRGS/1/2015/ICT01/UITM/03/1) in UniversitiTeknologi MARA, Shah Alam.

6. REFERENCES

[1] Najeeb M M, Abdelkader A A, Al-Zghoul M B. Arabic natural language processing laboratory serving Islamic sciences. International Journal of Advanced Computer Science and Applications, 2014, 5(3):114-117

[2] Harish B S, Guru D S, Manjunath S. Representation and classification of text documents: A brief review. IJCA Special Issue on Recent Trends in Image Processing and Pattern Recognition, 2010, 110-119

[3] Hong J L. MalayWordNet-A novel lexical database for Malay language.In 10th International Conference on Fuzzy Systems and Knowledge Discovery, 2013, pp. 773-777

[4] Ab Rahman S, Omar N, Ab Aziz M J. A fundamental study on detecting head modifier noun phrases in Malay sentence.In International Conference on Semantic Technology and Information Retrieval, 2011, pp. 255-259

[5] Alfred R, Leong L C, On C K, Anthony P. Malay named entity recognition based on rule-based approach. International Journal of Machine Learning and Computing, 2014, 4(3):300-306

[6] Nazri M Z A, Shamsudin S M, Bakar A A. An exploratory study of the Malay text processing tools in ontology learning. In 8th International Conference on Intelligent Systems Design and Applications, 2013, pp. 1-8

[7] Ranaivo M B. Building a rule-based Malay text segmentation tool. In International Conference on Asian Language Processing, 2011,pp. 276-279

[8] Lee J, Othman R M, Mohamad N Z. Syllable-based Malay word stemmer. In IEEE Symposium on Computers and Informatics Syllable-based, 2013, pp. 7-11

[9] Mohamed H H, Abu B Z, Ismail M. Evaluation of Malay grammar on translation of Al-Quran sentences using Earley algorithm. In 5th International Conference on Information and Communication Technology for the Muslim World, 2013, pp. 1-4

[10] Ab Rahman S, Omar N, Che Hassan N B. Construction of compound nouns (CNs) for noun phrase in Malay sentence. In International Conference on Information Retrieval and Knowledge Management, 2012, pp. 22-25

[11] Yaman M B. MyParser: A Malay text categorization toolkit using inference rule. Master thesis, Johor: Universiti Tun Hussein Onn Malaysia, 2013

[12] Don Z M. Processing natural Malay texts: A data-driven approach. Trames, 2010, 14(1):90-103

[13] Kasbon R, Amran N A, Mazlan E M, Mahamad S. Malay language sentence checker. World Applied Sciences Journal, 2011, 12(Special Issue on Computer Applications and Knowledge Management):19-25

[14] Basri S B, Alfred R, On C K. Automatic spell checker for Malay blog. In IEEE International Conference on Control System, Computing and Engineering, 2013, pp. 506-510

[15] Bird S., Klein E., Loper E. Natural language processing with python: Analyzing text with the natural language toolkit. California: O'Reilly Media Inc, 2009

[16] Alfred R, Leong L C, On C K, Anthony P, Fun T S, Razali M N, Hijazi M H A.A rule-based named-entity recognition for Malay articles.In International Conference on Advanced Data Mining and Applications, 2013, pp. 288-299

[17] Faidi K, Ayed R, Bounhas I, Elayeb B. Comparing Arabic NLP tools for hadith classification. International Journal on Islamic Applications in Computer Science And Technology, 2015, 3(3):1-12

[18] Zaraket F, Makhlouta J. Arabic cross-document NLP for the hadith and biography literature. Procein 25th International Florida Artificial Intelligence Research Society Conference Arabic, 2013, pp. 256-261

[19] LidwaPusaka. Ensiklopediahaditskitab 9 imam. Jakarta Timur: LidwaPusaka andSaltanera, 2010

[20]Azmi A M, Badia N bin. e-Narrator-An application for creating an ontology of hadiths narration tree semantically and graphically. Arabian Journal for Science and Engineering, 2010, 35(2):51-68

[21] Al-MunajjidM S. Conditions of a Sahih (sound) hadith. Riyadh: Islamic Propagation Office in Rabwah, 2009

[22] Aldhaln K, Zeki A, Zeki A, Alreshidi H. improving knowledge extraction of hadith classifier using decision tree algorithm. In International Conference on Information Retrieval and Knowledge Management, 2012, pp. 148-152

[23] Zed A. S.Learn python the hard way: A very simple introduction to the terrifyingly beautiful world of computers and code. Boston: Addison-Wesley, 2014

How to cite this article:
Rahman N A, Ismail N K, Nor Z M, Alias M N, Kamis M S, Alias N. Tagging narrator's names in hadith text. J. Fundam. Appl. Sci., 2017, *9(5S), 295-309.*