A Profession Identification On Narrative Scenario Document

Reza Fauzan[1], Daniel Oranova Siahaan[2] and Nurul Fajrin Ariani[2]
{reza.fauzan@poliban.ac.id, daniel@if.its.ac.id, nurulfajrin@if.its.ac.id }

[1]Department of Informatics, Politeknik Negeri Banjarmasin, Banjarmasin, Indonesia
[2]Department of Informatics, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia

Abstract: A clear determination of the profession of the stakeholders in the software requirement becomes important in order to get a good requirement. One way of identifying stakeholders is by reading a narrative scenario document. In narrative scenario documents, the system is required to read sentence patterns that can only be understood by humans while the profession of the stakeholders is in the document. Narrative scenario document as formal portrayals are not inconsequential, and it requires a solid learning of formal displaying. However, the system needs to be known all stakeholders in a software product. Stakeholders in software requirement can be represented as a profession. In this paper, we do a search for sentences that identify the profession of the stakeholders. Then we change it into sentence patterns that can be understood by the system. The results of sentence patterns that have been established indicate there is a substantial agreement with an expert.

Keywords: Narrative Scenario, Profession Identification, Software Engineering, Software Requirement

1. Introduction

Software development that meets the requirements and expectations of stakeholders is the final objective of software providers. To accomplish this, we should viably and precisely deal with the stakeholder's software requirements. These requirements incorporate the highlights, capacities, and characteristics in their software (Davis, 1993; Choi, Park and Sugumaran, 2012; Curcio *et al.*, 2018). Determining the requirement for good software is very important. However, the process of determining good requirements is not easy. The determination of software requirements has many problems. Some of the problems are the difficulty of determining the stakeholders, the conflict of interest, the existence of inconsistencies of interests, and so forth (Siahaan, 2012; Alshazly, Elfatatry and Abougabal, 2014; Sabatucci *et al.*, 2015).

One of the many things that can help the user in problem-solving of the software requirement's problems is the scenario document. The creation of a good and correct scenario document can improve the quality of the software requirements generated (Alspaugh and Anto, 2008). The use of scenario documents can help to understand a specific situation and prioritize requirements (J. C. S. P. Leite, G. Hadad, J. Doorn, 2000; Sarmiento *et al.*, 2016).

Scenario documents have several patterns in it, namely: setting up the storyline, expounding things that change, recognizing specialists and their associations, and disentangling the objective and its subgoals (Wang *et al.*, 2013). Scenario documents can help the target of the story to understand the reasons why the problems need to be solved by the system, why the system is important to be built, why the system needs to be built right away, and what potential benefits of the target may have. Document scenarios may be narrative,

ICASI 2019, July 18-19, Banda Aceh, Indonesia

DOI 10.4108/eai.18-7-2019.2288595

pictures or videos. The narrative scenario document has several components. They are objectives, scope, stakeholder point of view, visualization, brief, recursive, decomposition and refinement (Siahaan, 2012).

Components contained in software requirements must be in synergy with each other. This makes it possible to avoid inconsistencies in software requirements (Georg *et al.*, 2015; Condori-Fernandez and Lago, 2018; Irshad, Petersen and Poulding, 2018). So in the initial phase, we must determine the stakeholders correctly. Stakeholders can be represented in the form of the profession of the target.

From some of the previous explanations, we can conclude that the identification of the profession appropriately becomes very important. The narrative scenario document is not formal. Therefore, it is very difficult for the system to identify the actual profession on the document. Natural Language Processing (NLP) can be used to retrieve information from text-based documents. Some research using natural language processing can help program a more precise word recognition system (Demner-Fushman, Chapman and McDonald, 2009; Yue *et al.*, 2012; Wang *et al.*, 2014; Sevenster *et al.*, 2015; Anikushina, Taratukhin and von Stutterheim, 2018). In this paper, we identify the profession using the language approach. We do a search of sentence forms that give the possibility that it is a profession of someone.

2. Materials And Methods

In this research, we do several stages. Based on these stages, the system can identify the profession of a story in the narrative scenario document. The narrative scenario document that is included in this research is English text. The research stages are illustrated in Figure 1.

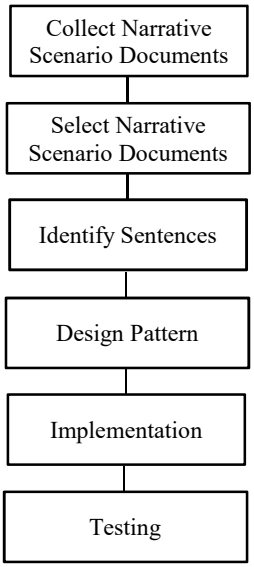

Fig. 1. Research Flow.

This research has 6 stages. The first stage is the collection of narrative scenario documents. Then the documents are selected according to the criteria of a good scenario. Then we search for and identify sentences which stating the profession of the stakeholders. When

the phrases are found, we create the patterns that the system can understand using natural language processing. Then we made a program to test the pattern. At last, we test the pattern using kappa statistic.

2.1 Collect Narrative Scenario Documents

The narrative scenario documents collected in the text document. The narrative scenario documents are collected from data already held from the past. The document collected is in English. There are 40 various documents collected. The narrative scenario document collected should have several things (Siahaan, 2012), namely as follows; scope; stakeholder's perspective; visualization; short, A4 size; recursive, decomposition, and refinement.

The aim is used to object that can be learned by systems analysts. The scope of a scenario must be well defined so that the discussion is not too extensive so that the given explanation can be specific. Scenarios should represent the problems of all stakeholder perspectives so as to produce better systems. Scenarios are expected to be short and simple but should include issues that are the primary concern of all stakeholders. Scenarios are based on essential occasions that tell the usefulness and advantages of the framework. At that point refined in view of client input and also extra data gathered from elicitation exercises.

2.2 Select Good Narrative Scenario Documents

Narrative scenario documents that have been collected are not necessarily good scenarios that can improve the quality of the software requirements. Good narrative scenario document parameters (Siahaan, 2012) are easy to access and easy to understand; important valuable, attractive; critical and challenging; often used and avoid things that are beyond habits; and specific. The final number of documents selected is 35 documents.

2.2.1 Easy to Access and Easy to Understand

Narrative scenario documents should be easily captured and understood by customers so they understand and understand what the contents of the scenario are.

2.2.2 Important, Valuable, Attractive

The contents of a given narrative scenario document should be important and have valuable information for interested customers to read. In addition, so that customers do not lose enthusiasm when reading narrative scenario documents and should be attractive and attractive.

2.2.3 Critical and Challenging

Narrative scenario documents must be critical and challenging so that developers can get feedback from customers.

2.2.4 Often Used and Avoid Things That Are Beyond Habits

Narrative scenario documents should use words or things that are often used in general and should avoid things that are considered outside of the habit in general.

2.2.5 Specific

Specific intentions are the individuals who play a role in the scenario should have a clear identity such as having a name, age, and other relevant attributes. In addition, there is a specific time and place if it is required and relevant. And the last is to have specific content.

2.3 Identify Sentences

The requirement of a good scenario is to be specific, especially in naming actors associated with the system. For example, the target must have a name, a job, an address, and so on. Whereas in the use case diagram, the necessary actors are professionally shaped actors, for example, customer, administrative officer, and others.

From the narrative scenario document we get, we try to find sentence patterns that indicate that there is a profession in the sentence. Some of the sentences we found are as follows.

- John is a customer.
- Budi, a young executive at Surabaya.
- Mr Heru Firman as a production manager requests the purchase of goods.
- A "broker" named Mr Bro who is willing to reserve a tugboat.
- This request is sent to the procurement in the document format and received by Mr Sueb who worked as an administrative officer.

Then sentences that have been found, performed the process POS (Part of Speech) tagging to help define the pattern of sentences. POS tagging is used to be the word identifier in the sentence (Wicaksono and Purwarianti, 2010; Fauzan and Siahaan, 2014). POS tagging gives certain labels on every word. The description of the POS (Wicaksono and Purwarianti, 2010) used is shown in Table 1.

Table 1. POS Tagging Description.

NO	POS	POS NAME
1	OP	Open Parenthesis
2	CP	Close Parenthesis
3	GM	Slash
4	;	Semicolon
5	:	Colon
6	“	Quotation
7	.	Sentence Terminator
8	,	Comma
9	-	Dash
10	...	Ellipsis
11	JJ	Adjective
12	RB	Adverb
13	NN	Common Noun
14	NNP	Proper Noun
15	NNG	Genitive Noun
16	VBI	Intransitive Verb
17	VBT	Transitive Verb
18	IN	Preposition
19	MD	Modal

NO	POS	POS NAME
20	CC	Coor-Conjunction
21	SC	Subor-Conjunction
22	DT	Determiner
23	UH	Interjection
24	CDO	Ordinal Numerals
25	CDC	Collective Numerals
26	CDP	Primary Numerals
27	CDI	Irregular Numerals
28	PRP	Personal Pronouns
29	WP	WH-Pronouns
30	PRN	Number Pronouns
31	PRL	Locative Pronouns
32	NEG	Negation
33	SYM	Symbols
34	RP	Particles
35	FW	Foreign Words

Here are the tagging results of several sentences that have been found.

- John/NNP is/VBZ a/DT customer/NN./.
- Budi/NNP /, a/DT young/JJ executive/NN at/IN surabaya/NN ./.
- Mr./NNP Heru/NNP Firman/NNP as/IN a/DT production/NN manager/NN requests/VBZ the/DT purchase/NN of/IN goods/NNS ./.
- A/DT "/" broker/NN "/" named/VBN Mr./NNP Bro/NNP who/WP is/VBZ willing/JJ to/TO reserve/VB a/DT tugboat/NN ./.
- This/DT request/NN is/VBZ sent/VBN to/TO the/DT procurement/NN in/IN the/DT document/NN format/NN ,/, and/CC received/VBD by/IN Mr./NNP Sueb/NNP who/WP worked/VBD as/IN an/DT administrative/JJ officer/NN ./.

From the sentences that have been found, we can find the pattern as follows.

- [name]NNP + is + DT + ... + [profession]NN
- [name]NNP + is + ... + [profession]NN
- PRP + is + DT + ... + [profession]NN
- PRP + is + ... + [profession]NN
- [name]NNP + , + DT + ... + [profession]NN
- [name]NNP + as + DT + ... + [profession]NN
- DT + ... + [profession]NN + named + [name]NNP
- [name]NNP + who + worked + as + ... + [profession]NN

Information:
... = can only be filled with adjectives (JJ) or emptied.
If there is an adjective before the noun, then both will be a profession.

2.3 Implementation

The software is built using Java programming language. In general, the software loads the document narrative scenario. Then the document is processed. And the resulting output is the profession of the stakeholders.

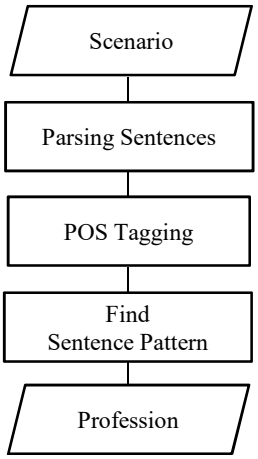

Fig. 2. Flow Chart Diagram.

Figure 2 explains how the system works. The system works from inputs in the form of narrative scenario documents. This system performs Natural Language Processing (NLP). This study uses StanfordNLP as an NLP tool. (Kumar and Dua, 2015). StanfordNLP in this study focused on parsing and tagging. Parsing is the process of extracting information about linguistic sentence formats for the advantages of machine translation applications, information extraction and more (Nivre, 2015; Dootio and Wagan, 2017; Kumari and Rao, 2017). Tagging is the word tagging in a sentence to define the label by using Part of Speech (POS) from the Pen Tree Bank model (Charniak *et al.*, 1996; Collobert *et al.*, 2011).
Here is an example of a scenario that is based on a narrative scenario.

> One day, Budi, a young executive in Surabaya, got a phone from his wife told him that phone, water, and electricity bills this month had not been paid, while the allowance for this month had already run out. This phone was included with 'will not get dinner at home before this problem is solved' threat.
>
> Budi has just realized that his wallet still contained approximately Rp. 150,000. Today he must attend meeting with several clients in different places, therefore he did not have time to queue and take money in the bank, queue in PLN, PDM, and Telkom counter. While he looked outside his office's window, he saw an ATM machine called "Inova". Before he walked to that ATM machine, he sent an SMS to that machine to ensure that the machine was able to be operated or under maintenance. Not too long after that, he

got a reply from that machine, showed that the machine was in a good condition and able to be operated.

In a hurry, he got out from his office, across the street towards the ATM. He directly inserted his ATM card, typed his PIN number, and did the transaction: checked the balance, paid electricity, water, and telephone bills, and took cash.

After Budi used it, ATM machine only left out with Rp. 1,000,000 cash, in which the minimum reserve requirement of cash. That machine sent an SMS to maintenance division, gave information and warning that the cash reserve inside that machine must be topped up.

Separation of the sentences and POS tagging using StanfordNLP is done in the above sentence. The result of the separation of the sentences and the provision of POS to the above scenario is as follows.

- One/CD day/NN, /, Budi/NNP, /, a/DT young/JJ executive/NN in/IN Surabaya/NNP, /, got/VBD a/DT phone/NN from/IN his/PRP$ wife/NN told/VBD him/PRP that/IN phone/NN, /, water/NN, /, and/CC electricity/NN bills/NNS this/DT month/NN had/VBD not/RB been/VBN paid/VBN, /, while/IN the/DT allowance/NN for/IN this/DT month/NN had/VBD already/RB run/VBN out/RP ./.
- This/DT phone/NN was/VBD included/VBN with/IN will/MD not/RB get/VB dinner/NN at/IN home/NN before/IN this/DT problem/NN is/VBZ solved/VBN threat/NN ./.
- Budi/NNP has/VBZ just/RB realized/VBN that/IN his/PRP$ wallet/NN still/RB contained/VBD approximately/RB Rp/NN 150,000/CD ./.
- Today/NN he/PRP must/MD attend/VB meeting/NN with/IN several/JJ clients/NNS in/IN different/JJ places/NNS ,/, therefore/RB he/PRP did/VBD not/RB have/VB time/NN to/TO queue/VB and/CC take/VB money/NN in/IN the/DT bank/NN ,/, queue/NN in/IN PLN/NNP ,/, PDM/NNP ,/, and/CC Telkom/NN counter/NN ./.
- While/IN he/PRP looked/VBD outside/IN his/PRP$ office/NN s/VBZ window/NN ,/, he/PRP saw/VBD an/DT ATM/NNP machine/NN called/VBD Inova/NNP ./.
- Before/IN he/PRP walked/VBD to/TO that/DT ATM/NNP machine/NN ,/, he/PRP sent/VBD an/DT SMS/NN to/TO that/DT machine/NN to/TO ensure/VB that/IN the/DT machine/NN was/VBD able/JJ to/TO be/VB operated/VBN or/CC under/IN maintenance/NN ./.
- Not/RB too/RB long/RB after/IN that/DT ,/, he/PRP got/VBD a/DT reply/NN from/IN that/DT machine/NN ,/, showed/VBD that/IN the/DT machine/NN was/VBD in/IN a/DT good/JJ condition/NN and/CC able/JJ to/TO be/VB operated/VBN ./.
- In/IN a/DT hurry/NN ,/, he/PRP got/VBD out/RP from/IN his/PRP$ office/NN ,/, acrossed/VBD the/DT street/NN towards/IN the/DT ATM/NNP ./.
- He/PRP directly/RB inserted/VBD his/PRP$ ATM/NNP card/NN ,/, typed/VBD his/PRP$ PIN/NN number/NN ,/, and/CC did/VBD the/DT transaction/NN :/: checked/VBD the/DT balance/NN ,/, paid/VBD electricity/NN ,/, water/NN ,/, and/CC telephone/NN bills/NNS ,/, and/CC took/VBD cash/NN ./.

- After/IN Budi/NNP used/VBD it/PRP ,/, ATM/NNP machine/NN only/RB left/VBD out/RP with/IN Rp/NN ./.
- 1,000,000/CD cash/NN ,/, in/IN which/WDT the/DT minimum/NN reserve/NN requirement/NN of/IN cash/NN ./.
- That/DT machine/NN sent/VBD an/DT SMS/NN to/TO maintenance/NN division/NN ,/, gave/VBD information/NN and/CC warning/NN that/IN the/DT cash/NN reserve/NN inside/IN that/DT machine/NN must/MD be/VB topped/VBN up/RP ./.
- After/IN that/DT ,/, the/DT ATM/NNP machine/NN remaining/VBG cash/NN amounting/VBG to/TO 1,000,000/CD rupiah/VBP which/WDT is/VBZ the/DT minimum/JJ cash/NN reserves/NNS of/IN cash/NN in/IN the/DT machine/NN ./.
- The/DT machine/NN sent/VBD a/DT message/NN to/TO the/DT maintenance/NN department/NN by/IN providing/VBG information/NN and/CC warning/VBG that/IN the/DT cash/NN must/MD be/VB immediately/RB added/VBN ./.

In the example of the previous scenario, after the identification of the profession, it was found that the profession was "young executive".

2.4 Testing

To test the accuracy of the proposed approach, the resulting output is compared with an expert's analysis. The expert used a system comparison is an analyst who has for two years working in software development efforts. The comparison result is calculated using kappa and yields kappa value. Kappa calculations are used because kappa gives value from the level of agreement between two experts (Carletta, 1996). The kappa value (Viera and Garrett, 2005; McHugh, 2012; Fauzan and Pramono, 2013) is indicated by the initial "k" of Formula 3 where P(A) denotes the value of both agreement and P(E) is the diversity of the two experts. The calculation of P(A) is described in Formula 2. The calculation of P(E) is described in Formula 3.

$$k = \frac{P(A) - P(E)}{1 - P(E)} \quad (1)$$

$$P(A) = \sum_{i=1}^{I} p_{ii} \quad (2)$$

$$P(E) = \sum_{i=1}^{I} p_i + p +_i \quad (3)$$

The kappa value can decide the level of assertion between specialists with the framework. Table 3 gives an interpretation of kappa values (Landis & Koch, 1977).

Table 3. Kappa Interpretation.

Index Kappa	Proportion of Agreements
< 0	less than chance agreement
0.01 – 0.20	slight agreement
0.21 – 0.40	fair agreement
0.41 – 0.60	moderate agreement
0.61 – 0.80	substantial agreement
0.81 – 1	almost perfect agreement

3. Result

To know the success of the professional identification of a narrative scenario document required a test. The scenarios used in the test are from scenarios that can provide the best results by the system until the most difficult scenarios processed by the system.

Table 4. Expert's Result.

		Expert 2 (System)		
		Y	N	Total
Expert 1 (Analyst)	Y	8	0	8
	N	1	1	2
	Total	9	1	10

Table 4 shows there are 8 trials said that systems and analysts have in common opinions with the answers generated. There is 1 system's answer that is not approved by an analyst. And there is 1 experiment where systems and experts do not find the profession written in the document. P(A), P(E), and k measures were calculated as follows.

$$P(A) = \frac{8+1}{10}$$
$$P(A) = 0{,}9$$
$$P(E) = \left(\frac{9}{10} \times \frac{8}{10}\right) + \left(\frac{1}{10} \times \frac{2}{10}\right)$$
$$P(E) = 0{,}72 + 0{,}02$$
$$P(E) = 0{,}74$$
$$k = \frac{0{,}9 - 0{,}74}{1 - 0{,}74}$$
$$k = \frac{0{,}16}{0{,}26}$$
$$k = 0{,}6154$$

We can find the value of P(A) of 0.9. And the value of P(E) is 0.74. From these two values, we can find the kappa kohen value of 0.6154.

From the test results, we found 2 document narrative scenarios that have different results with the analyst. After we look at the contents of the document, there is one document that has a sentence pattern that identifies the profession not included in the pattern made before. And there is another document that does not write down the true profession of the stakeholders.

4. Conclusions

This study was conducted to identify the profession in a narrative scenario document. The process is done from the collection of narrative scenario documents, document selection, sentence selection until the creation of sentence patterns. Then we built the software and test the built software.

The test result using kappa statistic gives Cohen's kappa value of 0.6154. The value proves that the established system can be used to identify the profession in the narrative scenario document with the result of "substantial agreement".

However, this system has deficiencies found during testing. Some scenarios that have not followed the scenario-making rule or miss-typed in the creation of a narrative scenario document. In addition, the system should be equipped with how the system can recognize new patterns that automatically identify the profession.

Acknowledgements

We thank our colleagues from Institut Teknologi Sepuluh Nopember. And also we thank Mr Daniel Oranova Siahaan for help with system and remarks that incredibly enhanced the composition.

References

[1] Alshazly, A. A., Elfatatry, A. M. and Abougabal, M. S. (2014) 'Detecting defects in software requirements specification', *Alexandria Engineering Journal*. Faculty of Engineering, Alexandria University, 53(3), pp. 513–527. doi: 10.1016/j.aej.2014.06.001.

[2] Alspaugh, T. A. and Anto, A. I. (2008) 'Scenario support for effective requirements', 50, pp. 198–220. doi: 10.1016/j.infsof.2006.12.003.

[3] Anikushina, V., Taratukhin, V. and von Stutterheim, C. (2018) 'Natural Language Oral Communication in Humans Under Stress. Linguistic Cognitive Coping Strategies for Enrichment of Artificial Intelligence', *Procedia Computer Science*. Elsevier B.V., 123, pp. 24–28. doi: 10.1016/j.procs.2018.01.005.

[4] Carletta, J. (1996) 'Assessing Agreement on Classification Tasks: The Kappa Statistic', *Computational Linguistics*, pp. 249–254.

[5] Charniak, E. *et al.* (1996) 'Taggers for parsers', *Artificial Intelligence*, 85(1–2), pp. 45–57. doi: 10.1016/0004-3702(95)00108-5.

[6] Choi, S., Park, S. and Sugumaran, V. (2012) 'A rule-based approach for estimating software development cost using function point and goal and scenario based requirements', *Expert Systems with Applications*. Elsevier Ltd, 39(1), pp. 406–418. doi: 10.1016/j.eswa.2011.07.029.

[7] Collobert, R. *et al.* (2011) 'Natural Language Processing (Almost) from Scratch', *Journal of Machine Learning Research*, 12, pp. 2493–2537. doi: 10.1.1.231.4614.

[8] Condori-Fernandez, N. and Lago, P. (2018) 'Characterizing the contribution of quality requirements to software sustainability', *Journal of Systems and Software*. Elsevier Inc., 137, pp. 289–305. doi: 10.1016/j.jss.2017.12.005.

[9] Curcio, K. *et al.* (2018) 'Requirements Engineering: a systematic mapping study in agile software development', *Journal of Systems and Software*. Elsevier Inc. doi: 10.1016/j.jss.2018.01.036.

[10] Davis, A. (1993) *Software requirements Objects, functions and state*. Prentice-Hall.

[11] Demner-Fushman, D., Chapman, W. W. and McDonald, C. J. (2009) 'What can natural language processing do for clinical decision support?', *Journal of Biomedical Informatics*. Elsevier Inc., 42(5), pp. 760–772. doi: 10.1016/j.jbi.2009.08.007.

[12] Dootio, M. A. and Wagan, A. I. (2017) 'Syntactic parsing and supervised analysis of Sindhi text', *Journal of King Saud University - Computer and Information Sciences*. The Authors. doi: 10.1016/j.jksuci.2017.10.004.

[13] Fauzan, R. and Pramono, D. (2013) 'Pemeriksaan Kemiripan Diagram Kasus Penggunaan Terhadap Skenario', *JUTI:Jurnal Teknologi Informasi*, 11, pp. 49–55. doi: http://dx.doi.org/10.12962/j24068535.v11i2.a11.

[14] Fauzan, R. and Siahaan, D. O. (2014) 'POS Tagging Bahasa Indonesia dengan Maximum Entropy Menggunakan Stanford Postagger', *POLHASAINS*, 2, pp. 26–30. Available at: http://linkinghub.elsevier.com/retrieve/pii/B9780124077683000173.

[15] Georg, G. *et al.* (2015) 'Synergy between Activity Theory and goal/scenario modeling for requirements elicitation, analysis, and evolution', *Information and Software Technology*. Elsevier B.V., 59, pp. 109–135. doi: 10.1016/j.infsof.2014.11.003.

[16] Irshad, M., Petersen, K. and Poulding, S. (2018) 'A systematic literature review of software requirements reuse approaches', *Information and Software Technology*. Elsevier B.V., 93, pp. 223–245. doi: 10.1016/j.infsof.2017.09.009.

[17] J. C. S. P. Leite, G. Hadad, J. Doorn, and G. K. (2000) 'A scenario construction process', *Requirements Engineering Journal*, pp. 38–61.

[18] Kumar, M. and Dua, M. (2015) 'Adapting Stanford Parser's Dependencies to Paninian Grammar's Karaka Relations Using VerbNet', *Procedia Computer Science*. Elsevier Masson SAS, 58, pp. 363–370. doi: 10.1016/j.procs.2015.08.033.

[19] Kumari, B. V. S. and Rao, R. R. (2017) 'Telugu dependency parsing using different statistical parsers', *Journal of King Saud University - Computer and Information Sciences*. King Saud University, 29(1), pp. 134–140. doi: 10.1016/j.jksuci.2014.12.006.

[20] McHugh, M. L. (2012) 'Interrater reliability: the kappa statistic', *Biochemia Medica*, pp. 276–282. doi: 10.11613/BM.2012.031.

[21] Nivre, J. (2015) 'Towards a Universal Grammar for Natural Language Processing', *Computational Linguistics and Intelligent Text Processing*, pp. 3–16. doi: https://doi.org/10.1007/978-3-319-18111-01.

[22] Sabatucci, L. *et al.* (2015) 'Ahab's legs in scenario-based requirements validation: An experiment to study communication mistakes', *Journal of Systems and Software*. Elsevier Ltd., 109, pp. 124–136. doi: 10.1016/j.jss.2015.07.039.

[23] Sarmiento, E. *et al.* (2016) 'Test Scenario Generation from Natural Language Requirements Descriptions based on Petri-Nets', *Electronic Notes in Theoretical Computer Science*. Elsevier B.V., 329, pp. 123–148. doi: 10.1016/j.entcs.2016.12.008.

[24] Sevenster, M. *et al.* (2015) 'A natural language processing pipeline for pairing measurements uniquely across free-text CT reports', *Journal of Biomedical Informatics*. Elsevier Inc., 53, pp. 36–48. doi: 10.1016/j.jbi.2014.08.015.

[25] Siahaan, D. O. (2012) *Rekayasa Perangkat Lunak*. Edited by D. Hardjono. Surabaya: Penerbit Andi.

[26] Viera, A. J. and Garrett, J. M. (2005) 'Interrater_agreement.Kappa_statistic', (May), pp. 360–363.

[27] Wang, H. *et al.* (2014) 'Extracting important information from Chinese Operation Notes with natural language processing methods', *Journal of Biomedical Informatics*. Elsevier Inc., 48, pp. 130–136. doi: 10.1016/j.jbi.2013.12.017.

[28] Wang, Y. *et al.* (2013) 'PLANT: A pattern language for transforming scenarios into requirements models', *International Journal of Human Computer Studies*. Elsevier, 71(11), pp. 1026–1043. doi: 10.1016/j.ijhcs.2013.08.001.

[29] Wicaksono, A. F. and Purwarianti, A. (2010) 'HMM Based Part-of-Speech Tagger for Bahasa Indonesia', *4th International MALINDO (Malay and Indonesian Language) Workshop*, (August).

[30] Yue, X. *et al.* (2012) 'Analysis of the combination of natural language processing and

search engine technology', *Procedia Engineering*, 29, pp. 1636–1639. doi: 10.1016/j.proeng.2012.01.186.