Design Answer Agent And Knowledge Repository For IQA (Interactive Question Answering) On Swamedikasi Mild Illness

Agus Mulyanto[1], Yeni Agus Nurhuda[2], Tony Reza Apriyanto[3]
Universitas Teknokrat Indonesia[1,2,3]
Jl. Z.A Pagar Alam, No 9-11, Labuhanratu, Bandar Lampung
agus.mulyanto@teknokrat.ac.id[1], agus.nurhuda@teknokrat.ac.id[2], tonyreza22@gmail.com[3]

ABSTRACT

Self-medication is a self-administered treatment, Self-medication is the most public effort to overcome the symptoms of disease without medical supervision. Lack of public knowledge about over-the-counter and drug-limited medicines in drug services. Therefore, there is a need for a system to find related information about Self-Medication knowledge, such as a system that can answer questions posed by user interactively about knowledge of Self-medication or commonly known as interactive question and answer system. Self-Medication Interactive Question Answering System (SMIQA) is a question and answer system about minor illnesses that can be tried the medicine. To be able to build SMIQA it is necessary to establish a knowledge repository of an agent's reply containing the knowledge of Self-Medication of a common health problems compiled using a tree structure stored in an AIML document. While the answer agent is a system that serves to find answers. To find answers in the knowledge repository will be used state concept information. This research produces Answer Agent design for IQA (Interactive Question Answering) on swadication of minor ailments based on free and restricted drug use according to the health department 2007. Testing conducted is a recall test where the data in the test is based on the sentence question, statement, dialogue grouding and clarification, and ellipsis dialogue with the results of the total accuracy of the test recall of 97.58%.

Keywords: Answer Agent, AIML, Interactive Answering Question, Knowledge Repository, Self-Medication.

1. Introduction

Knowledge Representation is a process for capturing the essential properties of a problem and making the information accessible to the problem solving procedure. Knowledge Representation is a field of Artificial Intelligence or artificial intelligence concerned with how a knowledge is symbolically represented and manipulated automatically by reasoning programs [3]. Representation of knowledge to be built that is about the answer agent Self-Medication mild disease.

Answer agent itself is an answering system where a system will be able to answer questions or statements submitted by users and allow users to control over the content to be displayed. This answer agent is built to answer related questions about Self-Medication, to obtain Self-Medication information about medicines in treating illness without a prescription.

Self-Medication is a self-administered treatment, in the "Training of Knowledge Enhancement and Skills of Choosing Drugs for Health Personnel" that Self-Medication is the most public effort to overcome symptoms of illness before seeking help from health personnel [4]. Self-Medication is usually done to overcome the complaints - minor health complaints such as dizziness, fever, mag, diarrhea, cough, influenza and others - which is often experienced by the community. However, treatment itself becomes not easy if you do not have knowledge about it. Ease of course is not the main thing, the more important it is how to do the Self-Medication correctly.

In the swamedication implementation, the community needs guidelines or integrated information to prevent medication errors. There is some knowledge about health that people should know about from the informer especially for drugs used in swamedication so that people can avoid drug misuse and drug misuse.

Lack of public knowledge about over-the-counter and limited medicines resulted in

the failure to make decisions in the use of medicines in swamedication. So the problem facing the public is the lack of public services that people use in asking questions about the use of drugs in Self-Medication correctly which can be accessed anywhere.

Information technology is needed at this time, because information is an important part in everyday life, where the information is expected to be obtained more quickly. Tools to obtain information is one of the obstacles that occur to date. Users of course want a more interactive information provider system or toolkit with its users, Interactive Question Answering or IQA is a branch of Question Answering or QA with increased interactivity that allows dialogue with users by stating information needs through questions using natural language and can give an answer to its users [5,6]. Interactive Qustion Answering System which will be built using Chatbot System concept as virtual assistant for question and answer. Chatbot itself is a computer program that can carry on conversations through written media using natural language [7, 8]. To build a chatbot will be done two stages of the knowledge repository and answer agent, the knowledge repository is a knowledge base that is structured using a tree structure and stored into the language AIML, (Artificial Intelligence Markup Language) AIML is a scripting language interpreter derived from Extensible Markup Languange XML) with a more specific function [9]. Answer Agent serves to find answers in the knowledge repository upon request from the dialog manager. To find answers in the knowledge repository will be used information state concept. this can be used to create applications that can be used for all sorts of fields, one of them in the field of health, especially Self-Medication.

Based on this, the focus of this research lies in the formation of a knowoledge repository answer agent which contains knowledge of swedish disease that is compiled using tree structure stored in AIML document with answer agent using information state model concept in search of answer. The results of this study are expected to obtain an Answer Design Solution for IQA (Interactive Question Answering) on swedicine mild disease based on guidelines for free and limited free drug use[4].

Literature Review

A. Interactive Question Answering

Interactive question answering (IQA) is one of the areas of research that is part of the field of questioning answering (QA) System and dialogue systems. The IQA system is defined as a QA system that supports the exchange of information between users and systems [2] like conversations between two people. The IQA system provides the user with the opportunity to ask questions and clarify questions until a desired information is obtained based on the feedback provided by the system, resulting in an interactive feedback between the user and the system in the sense that the system not only provides answers but also provides a back question to clarify.For example, the following conversations between IQA and user systems[2].

User: I want to find an inexpensive Japanese restaurant that takes reservations
Sys: I found 9 inexpensive Japanese restaurants that take reservations
Sys: Here are the first few:
Sys: GINZA JAPANESE RESTAURANT
Sys: OKI SUSHI CAFE
Sys: YONA SUSHI
Sys: Should I continue?

Here is an example of IQA OntoNLQA architecture [1].

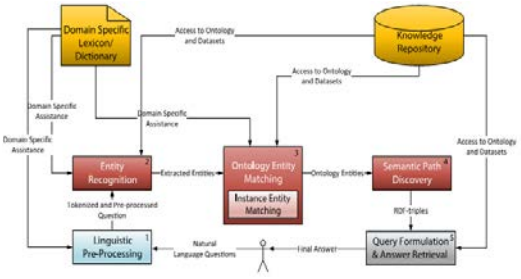

Figure 1. IQA architecture on OntoNLQA

OntoNLQA uses an ontology approach to frequently asked questions in natural language. OntoNLQA identifies key entities in the sentence, then searches the ontology and the simultaneous relationships between entities to process the questions and search

for answers. To support the process, OntoNLQA uses lexicon dictionaries for special domains and knowledge repository for ontology processes.

B. Issue of Modeling Dialogue

There are several problems of modeling dialogue[6]. including ellipsis dialogue, anphoric reference, grounding and clarification, turn talking. But in this study only discuss the issue of ellipsis dialogue and grounding and clarification, as well as explanation

1. Ellipsis

Ellipsis is incomplete sentence and has no verbal phrase. Interpretation and ellipsis resolution requires a dialogue context to resolve missing information from previous questions or statements. For example, an ellipsis dialogue scenario is shown below:

User: "penyakit flu apa?"
System : "Tahukah anda bahwa Flu adalah suatu infeksi saluran pernapasan atas"
User: " apa **obatnya** ?"

In the dialogue the missing information is the drug, what drug intended by the user, the drug intended by the user is a cold medicine, in this case the system is required to refer to the previous dialogue to lead to flu medicine.

2. Grouding and Clarification

Grouding and Clarification. In the dialogue scenario below, the system should clarify the meaning of "he", whether referring to cough or flu. Example:

User : "Apakah batuk termasuk dalam gejala flu ?"
System : "Batuk termasuk dalam gejala flu, tetapi batuk juga merupakan penyakit jika tidak disertai gejala flu lainnya.
User : " Apa **gejalanya** ?"

In the dialogue the information the user needs refers to what symptoms, whether the symptoms of cough or flu symptoms, the dialogue refers to flu symptoms, which the system must provide answers about flu symptoms.

C. Artificial Intelligence Markup Language (AIML)

Artificial Intelligence Markup Language (AIML) is a scripting language interpreter which is a derivative of Extensible Markup Language (XML) with more specific functions. One of its functions is to create a knowledge-based stimulus-response system [9]. Important parts of AIML are as follows:

1. Category

In AIML, category is the basic unit of knowledge. Category consists of at least

```
<category>
<pattern>siapa nama kamu </pattern>
<tamplate>Nama saya BOT</tamplate>
</category>
```

two AIML elements ie pattern and tamplate. Here's an example of a simple category:

At the time when the category above is loaded in memry then AIML bot will answer question "What is your name" with "My name is BOT".

2. Pattern

Patterns are a series of letters that are expected to match or match one or more of the user's input (Input). A pattern can be using wildcards that will match one or more user inputs. A pattern like the following.

```
Siapa nama*
```

Match the input "what is your name", "what is your lecturer name", and so on.

3. Tamplate

A tamplate determines the response of theappropriate pattern. A tamplate can be a simple literal text as follows:

```
Nama saya BOT.
```

Variables are the same value as the bot name and inserted into the sentence. Templates are also possible to forward to another pettern by using an AIML elm named srai.

The sieve element can be used to implement the meaning equations as in the following.

```
<category>
<pattern>siapa nama kamu </pattern>
<tamplate>Nama saya <bot
name=”name”/></tamplate>
</category>
<category>
<pattern>kamu dipanggil apa</pattern>
<tamplate>
```

example:

The first category will answer an input "what is your name" with a statement about the bot name. The second category will answer the "what you called" input by passing the question to the first category matching the "what is your name" input - in other words that the two phrases are comparable or the same.

4. That

That is often used in category creation, that is an AIML element that refers to the previous response or output.

2. Research Method.

A. Data Sources

The research that the authors do is qualitative research, which in the implementation will produce descriptive data in the form of fact data from drug free and free guidelines [2]. And collecting data from written words or dialogue conversations from social media persons associated with swamedication.

B. System Architecture Design

System architecture design to build both components can be seen in figure 2.

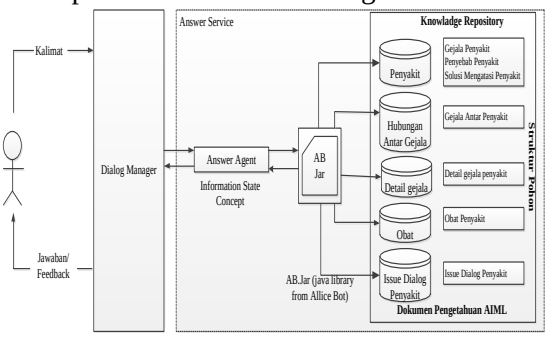

Figure 2. Design Architecture Answer Service built.

A brief description of the architecture in the picture. It starts when the user gives text to the system. Furthermore, the system will respond to sentences entered by the user to look for the main idea of the sentence through the dialog manager component. After that is given to the component of the answer agent to look for answers using information state concept that functions in recording the information received. In the process of searching for answers, the answer agent requires a java library developed by A.L.I.C.E (Artificial Linguistic Internet Computer Entity) that is Ab.jar. This library works to execute AIML brain file in which there are tags or commands to match the string of keywords obtained. Then, after the string matches the existing keywords in the AIML file, then the response is obtained in the form of answers in accordance with the keywords obtained. There are 4 AIML files stored in the knowledge repository: Symptoms. AIML, Diseases.AIML, causes.AIML, and Drugs.AIML. After the searching process is found, the answer is sent back by the Ab.jar library to the answer agent component for processing by the dialog manager. Then, after the answer given by the answer agent in the process by the dialog manager, the results of the response process are sent back by the user to respond to an answer or question again.

C. Answer Agent

The Answer agent in this research is used to record information derived from the dialog manager. Information state concept is used for answer agents. Answer Agent works to find answers in the knowledge repository upon request from dialog manager[4]. knowledge repository is built on tree structure, Tree is one of the most specific specific modeling and is usually used to describe hierarchy [8]. Here is an Answer agent based on the Information state model shown in Figure 3 below.

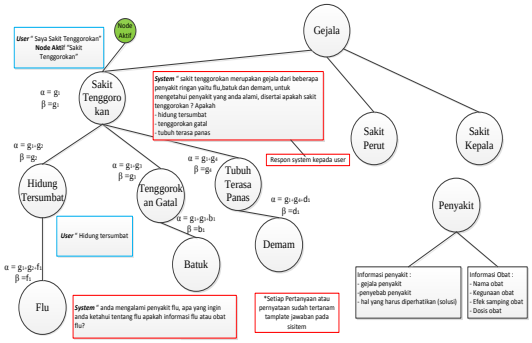

Figure 3. Answer agent based on the Information state model

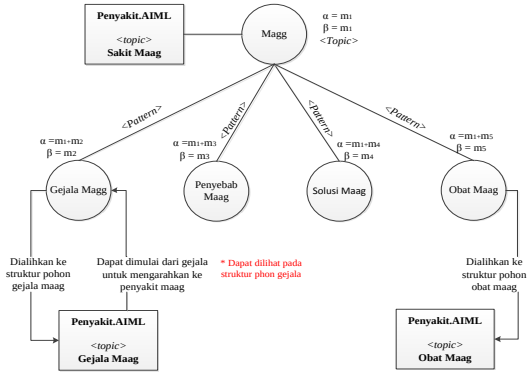

Figure 4. Representation of ulcer disease

The following information is recorded based on the information state model of the chosen knowledge representation on the "symptom" node and its target information on the "flu" illness.

$$\begin{vmatrix} A\alpha : & i \\ A\beta : List(i_1, i_2, i_3, \dots, i_n) & \end{vmatrix}$$

| Aα = g1|
| Aβ =g1 |
| Aα = g1+g2 |
| Aβ =g2 |
| Aα = g1+g2+f1 |
| Aβ =f1 |

g = gejala
f = penyakit flu

Based on the recording of information, obtained the results of a dialog or user conversation with the system that the user issued a statement for the symptoms of "my sore throat", then the user has symptoms of sore throat, and to know the disease experienced, the user input other symptoms of "nasal congestion" so that the information obtained more specific, that is flu disease.

Knowledge Repository

Knowledge repository built using tree structure. In the Tree Structure that has been designed there are three main variables in representing swamedication knowledge, namely: disease representation, symptom representation, and drug representation. the three variables are used to represent knowledge on 15 minor illnesses based on the free and free drug use guidebook. For example, the tree structure in ulcer disease is shown in Figure 4.

Symptoms of ulcers are pain in the pit of the stomach, nausea and abdominal bloating. Here is a symptom that occurs in an ulcer starting from a pain in the solar plexus shown in Figure 5.

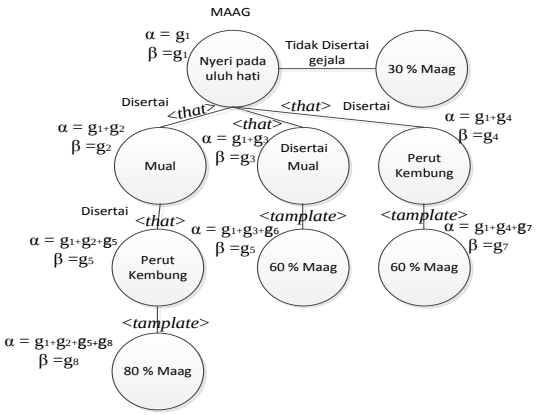

Figure 5. Representation of symptoms of ulcer

The following is a drug representation that can be used when the ulcer is shown in Figure 6.

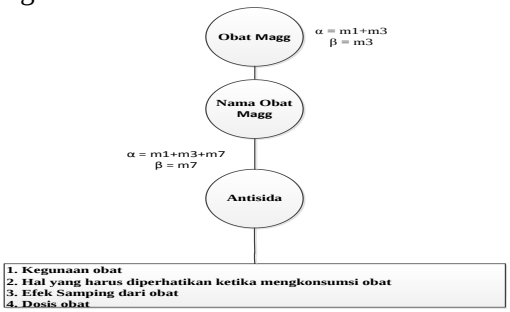

Figure 6. Drug Representation

3. Discussion.

The test is done on the component of the answer agent to find out the answer agent's ability to handle the dialogue problems: handle the question sentence, statement, handle the dialog containing the ellipsis element, handle the grouding and

clarification dialog. Here we will see the accuracy of the application response to the sentence forms in the dialog. The following test samples performed can be seen in table 1.

Table. 1. Test scenario

Dialog	Tepat	Tidak
alimat pernyataan ***User*** : saya sedang **sakit** ***System*** : apa keluhan anda ***User*** : saya sedang **sakit flu** ***System*** : Flu adalah suatu infeksi saluran pernapasan atas, Orang dengan daya tahan tubuh yang tinggi biasanya sembuh sendiri tanpa obat, Flu ditularkan melalui percikan udara pada saat batuk, bersin, dan tangan yang tidak dicuci setelah kontak dengan cairan hidung atau mulut	✔ ✔ ✔	
Kalimat pertanyaan ***User*** : **apa gejalan sakit flu** ***System*** : Demam, sakit kepala, nyeri otot, Mata berair, Batuk, bersin, hidung berair, Sakit tenggorokan, Kelelahan, Menggigil, Hidung Tersumbat, Kehilangan Nafsu Makan, Pegal Pegal	✔	
Dialog ellipsis ***User*** : **apa penyebabnya** ***System*** : Infeksi saluran pernapasan bagian atas oleh virus influenza	✔	
Dialog *Grouding and Clarification* ***User***: apakah batuk termasuk dalam gejala flu" ***System*** : ya batuk termasuk dalam gejala flu" ***User*** : apa gejalanya"		

B. Recall Testing

Tests using recall are used to obtain accuracy. Testing with recall can be seen in table 2. as follows:

Table. 2. Result of the recall testing

No	Sentence type	Number sentence	Total True	False Total	Accuracy *Recall*
1	Question	160	153	7	95.62%
2	statement	42	42	0	100 %
3	Dialog *grouding and Clarification*	7	7	0	100%
4	Dialogue *Ellipsis*	95	90	5	94.73%
Accuracytotal *recall*					97.58 %

$$recall = \frac{\text{amount of relevant data}}{\text{amount of relevant data} + \text{irrelevant da}}$$

Question:$recall = \frac{153}{160}\ x\ 100\% = 95{,}62\%$

Statement : $recall = \frac{42}{42}\ x\ 100\% = 100\%$

Dialog*Grouding and Clarification*

$$recall = \frac{7}{7}\ x\ 100\% = 100\%$$

Dialog *ellipsis*

$$recall = \frac{90}{95}\ x\ 100\% = 94{,}73\%$$

In table 2. shows the results of the recall test, from statement sentences and sentence questions and ellipsis dialogs and grouding and clarification dialogs. Sentence testing is randomly tested based on the sentence in the training data. Based on the test, the average of recall test is 97,58% (Ninety seven point fifty eight percent).

4. Conclusion.

The conclusion of this research is as follows In making the knowledge repository thing to note is the preparation of knowledge representation about swamedication. Where in compiling this swamedication knowledge can be constructed using a tree structure to facilitate the creation of AIML documents, and the ability of the answer agent in tracing the answers is determined by the many variations of the template patterns in the created AIML document. Thus, the more knowledge stored on the AIML document, the more responses given by SMIQA chatbot. This research is still very less seen from the data 4 sentence test

average with result of test using recall obtained equal to 97,58%.

Suggestions that can be given for the further development of this research, that is to add knowledge of the sentence of questions or statements that can't be answered by the system on AIML document, because the more knowledge made into the document AIML the knowledge of the system becomes more.

Bibliographies

[1] Asiaee, A. H., Minning, T., Doshi, P., & Tarleton, R. L., 2015, *A framework for ontology-based question answering with application to parasite immunology*, *Journal of Biomedical Semantics*, pp. 6:31, doi 10.1186/s13326-015-0029-x.

[2] Kelly D., Kantor, P. B., Morse, E. L., Scholetz, J. dan Sun, Y., 2009, *Questionnaires for eliciting evaluation data from users of interactive question answering systems*. Natural Language Engineering, 15 , pp 119-141 doi:10.1017/S1351324908004932.

[3] Ardianto, W., Anggraeni, W. and Mukhlason, A. (2012) 'Pembuatan Sistem Pakar Untuk Pendeteksian dan Penanganan Dini Pada Penyakit Sapi Berbasis Mobile Android Dengan Kajian Kinerja Teknik *Knowledge Representation*', *Jurnal Teknik ITS*, 1(1), pp. A310–A315. Available at: http://ejurnal.its.ac.id/index.php/teknik/article/view/1067.

[4] Bahartyan, E. and Bahtiar, N. (no date) '*web e-commerce*', 5, pp. 34–43.

[5] Departemen Kesehatan (2007) 'Pedoman penggunaan obat bebas dan bebas terbatas', *Direktorat Bina Farmasi Komunitas dan Klinik Ditjen Bina Kefarmasian dan Alat Kesehatan Departemen Kesehatan Republik Indonesia*, pp. 9–14.

[6] Pamungkas, W. C. (2011) '*Internet Relay Chatbot Dengan Menggunakan Artificial Intelligent Markup Language* (Aiml)', *Jurnal Informatika*, 4(April 2008), p. 23. Available at:https://labti.ukdw.ac.id/ojs/index.php/informatika/article/viewFile/42/23.

[7] Purwarianti, A. and Hakim, Z. (2012) '*Interactive Question Answering For Customer Service Representative Using*', pp. 1–6.

[8] Saputra, I M Wahyu, A.A.K. Oka Sudana, I. M. S. (2014) 'Implementasi Struktur Data Tree pada Sistem Informasi Upacara Yadnya berbasis Android', *Merpati*, 2(1004505060), pp. 1–10.

[9] Setiaji, B., Utami, E. and Fatta, H. Al (2013) 'Membangun Chatbot Berbasis AIML dengan Arsitektur Pengetahuan Modular', *Semnasteknomedia Online*, pp. 18–15.

[10] Syauqi, A. (2015) 'Sistem Tanya Jawab dengan Web Semantik', pp. 43–51