Journal of Theoretical and Applied Information Technology
31st January 2024. Vol.102. No 2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

AUTOMATIC GENERATION OF MULTIPLE-CHOICE QUESTIONS USING TEMPLATE-BASED SEMANTIC WEB IN INDONESIAN LANGUAGE

SHARFINA FAZA[1], ROMI FADILLAH RAHMAT[2*], DANI GUNAWAN[2], RINA ANUGRAHWATY[3], SILMI[4], GABRIELA DWI LADY[2], FARHAD NADI[5]

[1]Department of Computer Engineering and Informatics, Politeknik Negeri Medan, Indonesia

[2] Department of Information Technology, Universitas Sumatera Utara, Indonesia

[3] Department of Electrical Engineering, Politeknik Negeri Medan, Indonesia

[4] Department of Mechanical Engineering, Politeknik Negeri Medan, Indonesia

[5] Faculty of Business and Technology, UNITAR International University, Malaysia

*Corresponding E-mail: romi.fadillah@usu.ac.id

ABSTRACT

Questions serve as tools to assess a student's knowledge and understanding. However, creating questions on a large scale can be time-consuming. To address this issue, an automated question generation system offers a solution. An automatic question generation system, also known as a question generation system, is an application that automatically generates questions from a given text or document using a specific method. This study applies a template-based method using semantic web to generate questions categorized as non-factoid and factoid. In order to implement semantic web, the system utilizes ontology modeling, which includes main components such as classes, properties, and instances. The ontology modeling process is carried out using a software called Protégé. The dataset used in this study consists of biology and history subject matter for high school students, obtained from subject matter providers on the web. The system checks the dataset for keywords, categorizes the questions, and places them into the appropriate question template. The automated question generation system successfully generated a total of 2468 questions. During the testing phase, 2132 questions were deemed acceptable according to their respective categories, while 336 questions were not accepted. By calculating the percentage of accepted questions over the total questions, the overall accuracy rate was found to be 86.38%.

Keywords: *Questions, Automated Question Generation System, Template-Based Method, Semantic Web.*

1. INTRODUCTION

Questions are the most crucial element in learning activities. They play a vital role in advancing the quality of learning outcomes. By asking questions, teachers can effectively assess how well students understand the subject matter. It serves as a means to gauge students' comprehension and engagement with the material. Tan et al. (2011) suggested that answering questions is a crucial method for enhancing student understanding and performance, whether through electronic (online) or traditional offline learning processes [1]. However, providing learning activities that require resources, such as references and sufficient time, may result in the need to create a large quantity of questions with specific deadlines, which can be challenging to achieve [2]. To address this issue, an automatic question generation system becomes necessary.

An automatic question generation system is a tool that automatically generates questions from a given text or document using a specific method. This system is valuable for creating practice questions efficiently. The automatic question generation can be performed using various documents found on the internet (open domain) or from documents contained within the system itself (closed domain). Globally, questions are categorized into five types: factoid questions, non-factoid questions, yes/no questions, list questions, and opinion questions. This indicates that the automatic question generator system can generate questions from one or multiple categories or encompass all question categories.

Various studies on automatic question generation systems have been conducted previously. For instance, Liu & Calvo (2012) conducted research entitled " G-Asks: An Intelligent Automatic Question Generation System for Academic Writing Support". [3]. Subsequently, Rocha & Zucker (2017) conducted research entitled "Automatic Generation of Educational Quizzes from Domain Ontologies" [4]. Attar et al. (2018), also contributed to this area with their study entitled “Automatic Quiz Generator” [5]. Additionally, Ginanjar & Purnamasari (2017) conducted a study titled "Automatic Question Generation for Indonesian Texts Containing Compound Sentences" [6].

Various studies on automatic question generation systems have been conducted previously. For instance, Liu & Calvo (2012) conducted research entitled " G-Asks: An Intelligent Automatic Question Generation System forAcademic Writing Support". In their study, they employed a Machine Learning approach based on useful features like cue phrases. The results revealed that their new approach outperforms the rule-based approach in five citation categories. To gain further insights into how human experts generate specific trigger questions, they analyzed the questions generated by human supervisors to support writing a literature review. Their method utilized NLP and achieved a potential question accuracy of 57% without complicated inference processing. [3].

Subsequently, Rocha & Zucker (2017) conducted a study titled "Automatic Generation of Educational Quizzes from Domain Ontologies". This study created a multiple-choice question generation system in English using domain ontologies, encompassing a range of difficulty levels from low to high, including a mixture of low and high difficulty questions [4].

In another study, Attar et al. (2018), conducted research titled "Automatic Quiz Generator" to generate automatic quizzes using Heilman's question generation (QG) system. This system generates multiple-choice questions with each question assigned a rank or score based on linguistic features such as grammar, length, pronoun replacement, etc., which together constitute the linguistic score. [5].

Furthermore, Ginanjar & Purnamasari (2017) conducted a study titled "Automatic Question Generation for Indonesian Texts Containing Compound Sentences". Their research succeeded in generating factoid questions in the form of essay questions, including who, what, where, when, and how many, as well as non-factoid questions in the form of definition questions, methods, and reasons in the form of essay questions. The system achieved a question generation accuracy of 71.06%, with errors attributed to preprocessing (44.55%), syntactic analysis (21.78%), and question generation (33.66%) [6].

In the last decade, generating automated questions for educational purposes has attracted the attention of researchers from various disciplines. However, previous research has primarily focused on generating questions in English, languages of continental Europe, and languages from East Asia. Only a few studies have addressed automatic question generation in the context of the Indonesian language, this is one of problem need to be solved based on this study. Moreover, the accuracy of the generated questions in Indonesian has been relatively low in some cases, leading to the system producing repeated questions, thus it will become our second problem statement in this research.

Based on the aforementioned previous research, the researchers aim to develop a system that can automatically generate multiple-choice questions from subject matter using semantic web technology in Indonesian language thus this is our research gap to be fulfilled. The automatic question generator system employed in this study will utilize a template-based method to generate questions, categorizing them as non-factoid and factoid. The primary goal of this research is to achieve high accuracy and mitigate the occurrence of repeated questions.

2. METHODOLOGY

To automatically generate multiple-choice questions using a semantic web-based template-based method, the research is carried out through several stages. The initial stage involves data collection in the form of biology and high school history subject matter from subject matter providers and online tutoring platforms. Subsequently, RDF and ontology are constructed to establish relationships between the collected data, including domain determination and class definition with its hierarchy. Once the data is implemented in the form of an ontology, SPARQL is used to extract relationship results from the created classes.

The research then progresses to the text preprocessing phase, which includes tokenization, POS tagging, and Name Entity Recognition (NER). Following the preprocessing stage, the study identifies phrases and combinations of phrase patterns, as well as determines grammatical

functions. The subsequent process involves generating multiple-choice questions using the template-based method. Figure 1 illustrates the general architecture that outlines the stages of the research.

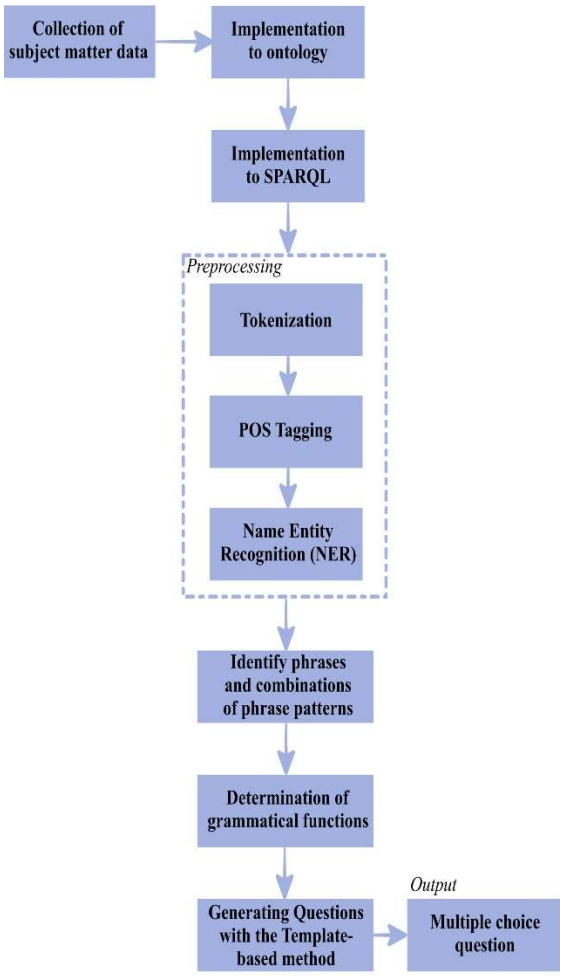

Figure 1. General Architecture

2.1 Dataset

The data used in this study comprises biology and history subject matter for high school students, sourced from an online tutoring website, as well as subject matter providers, namely "ruangguru" and "quipper". The selected subject matter specifically excludes materials that rely on pictures and formulas in their explanations.

2.2 Implementation to Ontology

In Greek, Ontology consists of two words, namely "ontos" and "logos" [7]. Ontology engineering is an emerging field that concerns the principles methods, methodologies and tools for developing and managing ontologies [8].In the field of artificial intelligence, ontology has two interconnected meanings. The first meaning is ontology as a glossary that identifies a specific subject or domain. The second meaning is ontology as a framework of knowledge that serves to explain certain limits or boundaries. Based on these two meanings of ontology, it can be concluded that ontology is a standard presentation that identifies a design of both classes and concepts within a domain. The properties derived from each class and concept are explained through attributes.

To implement the ontology, Protégé 5.5.0 software is used which is useful for designing ontologies for a domain [9]. The stages in developing the ontology are as follows [10]:

1. Selecting the Ontology Domain
   The ontology domain in this research is the high school subjects, which include modules from Biology and History subjects.
2. Defining Classes and Subclasses
   The definition of classes in this research includes "education level" and "subject module". Next, the class "StudyModules" is defined as a subclass of Biology and History, which are more specific, while "EducationalLevel" is a class without subclasses. The concepts of classes and subclasses are illustrated in Figure 2.

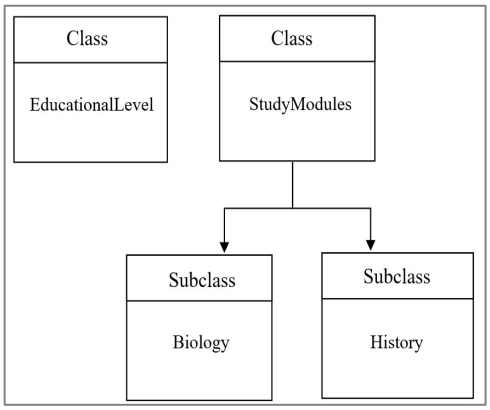

Figure 2. Design of Classes and Subclasses

3. Defining Data Properties and Object Properties
   Properties are relationships between objects and are typically defined using adjectives or verbs. In this stage, the definition is done with several properties, as shown in Table 1.

Table 1. Design of Properties in Classes

Property	Function	Type
HasAStudyModules	Connecting EducationalLevel with StudyModules	*Object Property*

ModuleLevel	Connecting StudyModules with EducationalLevel	*Object Property*
ModuleName	Connecting value string ModuleName from StudyModules	*Datatype Property*
ModuleContent	Connecting value string ModuleContent from StudyModules	*Datatype Property*

4. Defining Instances or Individuals
   Instances or individuals are the earliest instances of a class that are associated with properties, as shown in Figure 3.

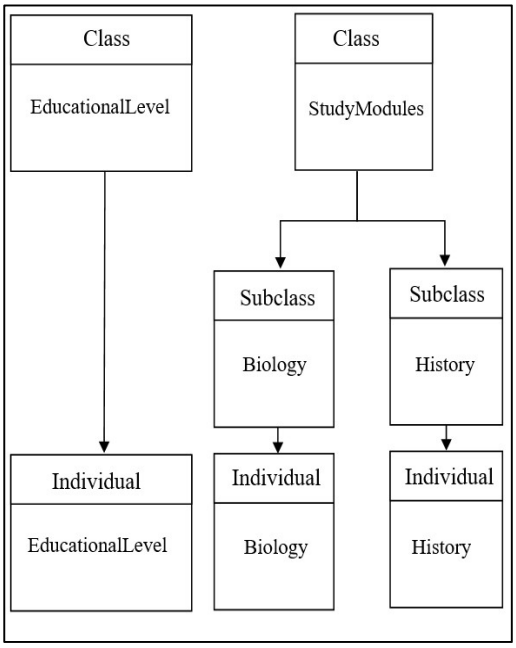

Figure 3. Design of Instances or Individuals

2.3 Implementation to SPARQL

After transforming the data into an ontology form, it is processed using SPARQL to obtain the results of each class relationship. SPARQL is processed using the Apache application Jena Fuseki to retrieve data from SPARQL [11]. Subsequently, the SPARQL results are exported to the database to facilitate the next process, namely the preprocessing stage.

SPARQL is a query language used to obtain and manipulate semantics from documents contained in RDF rules. Mapping a SPARQL query can impact the application's response time when displaying search results [12]. SPARQL provides a standard query language for RDF graphs [13]. Similar to SQL and XQuery, the SPARQL language allows a declarative user interface for interacting with RDF databases. It was created to combine data from various sources and foster the development of web 2.0 applications, leading to a standardized web service for expressing queries [14].

2.4 Preprocessing

The preprocessing stage is conducted to prepare the text in a more structured manner. This stage is divided into several step.

2.4.1 Tokenization

Tokenization involves breaking down text in a paragraph, sentence, or page into individual units called tokens for further analysis. The purpose of Tokenization is to convert words in paragraphs, sentences, or pages into discrete units. During the tokenization process, unnecessary punctuation marks are removed. The results of the tokenization stage are presented in Table 2.

Table 2. Table of Tokenization Results

Before the Tokenization Process	After the Tokenization Process
Vertebrata adalah golongan hewan yang memiliki tulang belakang. Tulang belakang berasal dari perkembangan sumbu penyokong tubuh primer atau notokorda (korda dorsalis).	*["Vertebrata", "adalah", "golongan", "hewan", "yang", "memiliki", "tulang", "belakang", "Tulang", "belakang", "berasal", "dari", "perkembangan", "sumbu", "penyokong", "tubuh", "primer", "atau", "notokorda", "korda", "dorsalis"]*

2.4.2 POS Tagging

Case folding is the process of converting all letters in a document to lowercase, ensuring uniformity in the letter types being processed. During this process, only the letters 'a' to 'z' undergo changes. It is performed because not all text documents maintain consistent capitalization.

The next stage is POS Tagging (Part-of-Speech tagging). In the POS Tagging process, a set of text that has been input into the system will be labeled with their respective word classes. POS Tagging is a stage for categorizing words and managing them during the sentence elaboration phase. The word labels include verbs, nouns, adjectives, prepositions, and so on. POS tagging is a phenomenon of allotting the words in a textual matter as matching to a picky component of speech. In general, POS tagging is as well denoted to as grammatical tagging of textual matter as representing to a specific component of

speech because of both its definition and context [15]. The results of POS Tagging are shown in Table 3.

Table 3. Tabel Hasil Proses POS Tagging

Before POS Tagging Process	After POS Tagging Process
["Vertebrata", "adalah", "golongan", "hewan", "yang", "memiliki", "tulang", "belakang", "Tulang", "belakang", "berasal", "dari", "perkembangan", "sumbu", "penyokong", "tubuh", "primer", "atau", "notokorda", "korda", "dorsalis"]	*["Vertebrata", "NNP"], ["adalah", "VB"], ["golongan", "NN"], ["hewan", "NN"], ["yang", "SC"], ["memiliki", "VB"], ["tulang", "NN"], ["belakang", "NN"], ["Tulang", "NN"], ["belakang", "NN"], ["berasal", "VB"], ["dari", "IN"], ["perkembangan", "NN"], ["sumbu", "NN"], ["penyokong", "VB"], ["tubuh", "NN"], ["primer", "JJ"], ["atau", "CC"], ["notokorda", "VB"], ["korda", "VB"], ["dorsalis", "NN"]*

2.4.3 Named Entity Recognition (NER)

Named Entity Recognition (NER) is a crucial stage in natural language processing, aimed at identifying and classifying entities present in a given text. Entities refer to specific objects, persons, locations, dates, and other named elements within the text [16]. The effectiveness of the NER process depends heavily on the accuracy of the input data and the context in which the entities appear, which poses challenges for the NER system in correctly identifying and labeling entities.

The NER system analyzes the contextual information of each word, taking into account the surrounding words and their grammatical classes to determine the entity type. For instance, if a word is preceded by "Mr." or "Ms.," it is likely to be recognized as a person (person entity). Similarly, mentions of well-known organizations or locations would be tagged as organization and location entities, respectively. NER's ability to discern these entities is essential for downstream natural language processing tasks, such as question generation, information retrieval, and sentiment analysis.

In the domain of question generation, recognizing entities becomes particularly valuable. Incorporating identified entities in generated questions enhances the system's ability to produce contextually relevant and accurate questions. For example, by identifying a location entity, the question generation system can frame location-specific queries, improving the overall quality of generated questions. The results of the NER process are shown in Table 4.

Table 4. Table of POS Tagging Process Results

Before POS Tagging Process	After POS Tagging Process
["Vertebrata", "NNP"], ["adalah", "VB"], ["golongan", "NN"], ["hewan", "NN"], ["yang", "SC"], ["memiliki", "VB"], ["tulang", "NN"], ["belakang", "NN"], ["Tulang", "NN"], ["belakang", "NN"], ["berasal", "VB"], ["dari", "IN"], ["perkembangan", "NN"], ["sumbu", "NN"], ["penyokong", "VB"], ["tubuh", "NN"], ["primer", "JJ"], ["atau", "CC"], ["notokorda", "VB"], ["korda", "VB"], ["dorsalis", "NN"]	*["Vertebrata", "ORG"], ["adalah", "O"], ["golongan", "O"], ["hewan", "ORG"], ["yang", "O"], ["memiliki", "O"], ["tulang", "ORG"], ["belakang", "ORG"], ["Tulang", "O"], ["berasal", "O"], ["dari", "O"], ["perkembangan", "O"], ["sumbu", "O"], ["penyokong", "O"], ["tubuh", "O"], ["primer", "O"], ["atau", "O"], ["notokorda", "O"], ["korda", "O"], ["dorsalis", "O"]*

2.5 Identify Phrases and Combination of Phrase Patterns

After the preprocessing is complete, the next step is to identify phrases and combinations of phrase patterns. Phrases are grammatical units formed by two or more words that support each other in a grammatical function. As a construction, phrases are formed by several constituent elements that are related to one another. Phrases are classified into several types, namely; nominal phrases (NP), verbal phrases (VP), adjective phrases (AP), numeral phrases (NUMP), pronominal phrases, adverbial phrases (ADVP), and prepositional phrases (PP).

During the introduction of phrases into this clause, the process involves comparing the similarity of the combination of words in the clause with the defined phrase patterns. The detection of phrase patterns is carried out sequentially from the first word to the last word in the clause. Table 5 is a summary of phrase patterns, which serve as a reference for combining phrase patterns.

Table 5. Combination of Phrase Patterns.

No	Combination	Phrase Level
1	NN + NN	NP
2	NN + NN + NN	NP
3	NN + CD	NP
4	NN + DT	NP

5	NN + PR	NP
6	IN + JJ	NP
7	NN + VB + DT	NP
8	NN + JJ	NP
9	NN + DT	NP
10	NN + JJ + JJ + PR + DT	NP
11	NN + PR + CC + JJ + DT	NP
12	VB + RB	VP
13	VB + DT	VP
14	VB + JJ	VP
15	JJ + JJ	AP
16	JJ + RB	AP
17	JJ + NN	AP
18	JJ + VB	AP
19	CD + CD	NUMP
20	CD + RB	NUMP
21	CD + NN	NUMP
22	CD + NN + DT	NUMP
23	RB + NN	NP
24	RB + VB	VP
25	RB + JJ	AP
26	RB + CD	NUMP
27	RB + DT	ADVP
28	RB + RB + VB	VP
29	RB + RB + DT	ADVP
30	MD + RB + VB	VP
31	MD + JJ	AP
32	MD + JJ +VB	VP
33	MD + VB	VP
34	IN + NN	PP
35	IN + NN + NN	PP
36	IN + NN + JJ	PP
37	IN + NN + VB	PP
38	IN + NN + VB + DT	PP
39	IN + PR	PP
40	NEG + VB	VP
41	NEG + VB + VB	VP
42	VB + VB	VP
43	VB + VB + VB	VP
44	NN + CC	NP
45	NN + CC + NN	NP
46	RB + NN + NN	NP
47	CD + NN + NN	NUMP
48	SC + JJ	AP
49	IN + NN + NN + NN	PP
50	SC + NN	NP

2.6 Determination of Grammatical Function

The stages of determining grammatical functions aim to identify grammatical functions (subject, predicate, object, complement, and description) for each word or phrase in a clause. This process is essential in the stage of generating factoid category questions, as it allows the question generation system to understand the structure and relationships within the sentence. By identifying the grammatical functions, the system can frame accurate and contextually relevant questions that extract specific pieces of information from the text. This level of linguistic analysis enhances the overall quality and precision of the generated factoid questions, making the question generation process more effective and informative. The flowchart for determining grammatical functions is shown in Figure 4.

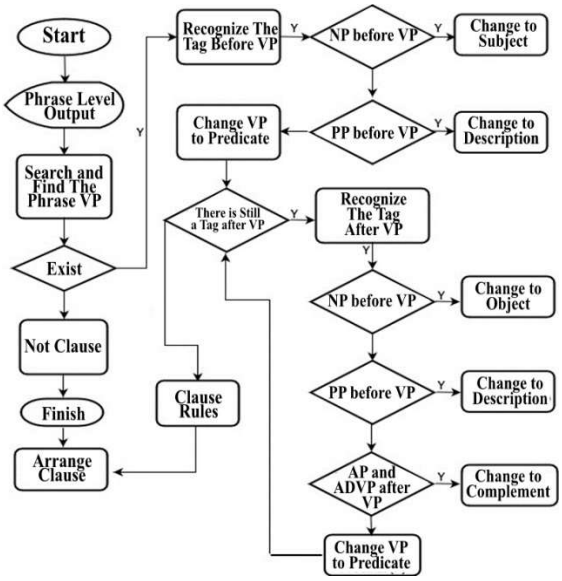

Figure 4. Flowchart of Grammatical Functions

2.7 Questions Generating

After completing the preprocessing, phrase identification, and determination of grammatical functions, the input sentences in the system contain various types of essential information used in the process of generating questions. The questions generated consist of both factoid and non-factoid questions.

The generation of question sentences is facilitated by templates based on the presence of keywords and grammatical functions obtained from the previous analysis [17]. The generation of non-factoid questions involves identifying keywords present in the datasets. Once the keywords are identified, the system selects a question template that matches the keyword category and then generates questions and answers accordingly using the selected template. Below is a list of keywords and question templates that can be used to generate non-factoid questions from a sentence.

Table 1. Keyword of Non-factoid

Non-factoid Category	Keywords	
	Before Target	After Target

Definition	*disebut, dikenal, dinamakan, mendefinisikan*	*adalah, yaitu, ialah, merupakan, diartikan*
Reason	*oleh sebab itu, jadi, memungkinkan adanya, dengan demikian, maka, dikatakan, penyebab terjadinya, sehingga, walau demikian, namun demikian*	*sebab, karena, bertujuan*
Method	*berfungsi untuk, berguna untuk*	*dengan cara*

A clause or sentence that is identified as non-factoid through a non-factoid detection process will prompt the system to generate questions using a non-factoid template. The non-factoid templates are shown in Table 2.

Table 2. Template of Non-factoid Questions

Question Category	Template
Definition	*Apa yang dimaksud + target + ?*
Reason	*Mengapa + target + ?*
Method	*Bagaimana cara. + target + ?*

The process of generating factoid questions will be conducted if the clause is not detected as a non-factoid question. Factoid question generation process relies on a customized template derived from the results of NER and syntactic analysis. The factoid question template is shown in Table 3.

Table 3. Template Question Factoid

Question Word	Template
Apa	*Apa yang + predicate + subject + complement + description + ?*
Siapa	*Siapa yang + predicate + subject + complement + description + ?*
Dimana	*Dimana + predicate + subject + complement + ?*

3. RESULT AND TESTING

3.1 Experimental Result

The system generated a total of 1371 multiple-choice questions. These questions were extracted from 30 datasets of high school biology and history subject topics, with each dataset serving as input. Below is an example of a question successfully generated by the system.

Table 4. Examples of Generated Questions

No	Subject Topics	Questions raised
1	Types of Circulatory System Technologies (Biology-XI class)	**Question :** *Apa yang di maksud dengan elektrokardiograf ecg?* **Answer:** *a.) menyuntikkan bahan ke tubuh pasien* *b.) detak jantung menstabilkan detak jantung memberi impuls listrik berkekuatan* *c.) teknologi sistem peredaran darah yang pertama akan dibahas (Real Answer)* *d.) membuat saluran aliran darah ke*
		Question: *Apa yang di maksud dengan angioplasti ?* **Answer:** *a.) alat pemacu detak jantung untuk menstabilkan detak jantung dengan memberi impuls listrik berkekuatan ringan* *b.) teknologi selanjutnya (Real Answer)* *c.) detak jantung menstabilkan detak jantung memberi impuls listrik berkekuatan* *d.) di kedokteran ya*
2	History as Stories and Events (History-X Class)	**Question**: *Apa yang di maksud sejarah dikatakan sebagai kisah karena sejarah ?* **Answer:** *a.) sejarah* *b.) peristiwa* *c.) kejadiankejadian pada masa lalu yang kemudian dibangun kembali (Real Answer)* *d.) kejadiankejadian*
		Question: *Apa yang di maksud sebagai peristiwa sejarah ?* **Answer:** *a.) sebuah fakta yang hadir dari masa lalu (Real Answer)* *b.) sejarah*

		c.) orang *d.) secara tidak langsung setiap cerita yang dibuat pastinya tidak dapat langsung*
3	Virus (Biology-X Class)	**Question:** *Apa yang di maksud juga sebagai parasit intraseluler obligat ?* **Answer:** *a.) virus yang akan memperlemah kekebalan tubuh manusia* *b.) virus* *c.) virus hanya dapat hidup dan berkembang biak pada sel hidup atau (Real Answer)* *d.) penyakitpenyakit*
		Question *: Mengapa virus hanya akan berkembangbiak di dalam sel hidup ?* **Answer** *:* *a.) virus berperan* *b.) virus bersifat aseluler atau tidak mempunyai sel (Real Answer)* *c.) dengan menjaga kebersihan lingkungan sekitar* *d.) virus*

3.2 System Testing and Evaluating

Testing of the system was carried out by 15 people consisting of teachers and students. In testing the questions raised by the system, categories are used to indicate the questions raised are declared true or not true, the following categories are used in the test:

1. Writing questions generated by the system is correct.
2. The questions generated by the system are understandable.
3. The questions generated by the system are appropriate to the context.
4. If the reader is given the same text, will the reader make a question similar to the one generated by the system.
5. The questions generated by the system are useful questions.
6. The answer in multiple choice is correct

The results of the tests carried out by 15 examiners based on the test categories can be seen in table 5.

Table 5. Table of Question Generation Test Results

No	Subject Topics	Number of Generated Questions	Number of Correct Questions
1	Animalia	108	85
2	Biological Sciences Branches	59	40
3	Fungi	78	69
4	Pre-literacy Period	32	23
5	Historical Research Methods	34	27
6	Environmental pollution	65	48
7	Plantae	37	25
8	The Process of Entering Hinduism and Buddhism in Indonesia	19	14
9	Protist	26	21
10	History as Stories and Events	25	19
11	Virus	125	116
12	Indonesian democracy	22	17
13	Enzyme	30	25
14	Reformation Period	47	35
15	Meiosis	31	22
16	Economic Politics in the New Order	37	26
17	Period Green Revolution	38	27
18	Theory of the Origin of Human Life	31	24
19	The Dynamics of RI's First Government	16	12
20	Circulatory System Technology	29	23
21	Colonialism and Imperialism	63	54
22	Muscular System in humans	35	26
23	Ethical Politics	58	43
24	Industrial Revolution	36	30
25	Endocrine System	24	20
26	Human Body Defense System	31	25
27	Menstruation and Ovulation	36	31

28	Types of Inherited Diseases in Humans	32	27
29	History of PPKI	92	73
30	History of the Proclamation	107	90

Based on Table 5, the total number of questions generated by the system is 2468. Among these, 2132 questions were accepted based on the test categories, while 336 questions were not accepted. The overall accuracy percentage, calculated by dividing the number of accepted questions by the total number of questions, is 86.38%.

The evaluation results indicate that this research is not yet perfect, as there are still many errors in the questions generated, both in the factoid and non-factoid categories. These errors include issues caused by preprocessing and inaccuracies in the question generation process. However, compared with previous research in [3] and [6] we can see our result is much better. This is due to our methodology used in this research.

Several examples of errors in the questions generated by the system are presented in Table 6

Table 6. Examples of Errors in Non-Factoid Questions of the Definition Category.

Questions
Question : *Apa yang di maksud mereka ?* Answer : *a.) organisme multiseluler mereka benarbenar tidak memiliki dinding sel (Real Answer)* *b.) bagian terluar tubuh vertebrata* *c.) coelentrata* *d.) invertebrate*
Question : *Apa yang di maksud beberapa diantaranya ?* Answer : *a.) •rhizophus stolonifera jamur* *b.) jamur* *c.) hifa berdiferensiasi* *d.) jamur pada makanan* (Real Answer)
Question : *Apa yang di maksud berikut ?* Answer : *a.) basidiomycota* *b.) jadi, askus* *c.) beberapa contoh jamur anggota divisi ascomycotina (Real Answer)* *d.) aspergillus fumigatus*
Question : *Apa yang di maksud contoh lumut hati ?* Answer : *a.) spagnum* *b.) memiliki alat perkembangbiakan berbentuk kerucut yang* *c.) marchantia polimorpha* (Real Answer) *d.) paku*
Question : *Apa yang di maksud berikut ini ?* Answer : *a.) rajaraja* *b.) nama-nama kerajaan hindu buddha* (Real Answer) *c.) tersebut,* *d.) krom*

In Table 6, examples of errors in non-factoid questions of the definition category can be observed. This is due to the presence of certain words like "is," "are," "and," and "consists of" in the selected subject matter, which leads the system to generate questions from these sentences under the definition category, even though they are not actual definitions. To reduce such errors, it is recommended to conduct a detection process for non-definition sentences before proceeding to the question generation stage.

In Table 7, examples of errors in non-factoid questions of the reason category are shown. The questions, in terms of the predetermined question templates, do not have errors since the generated question sentences contain relevant keywords for non-factoid questions of the reason category. However, these questions are categorized as incorrect because their meaning is not understood and they do not align with the context according to the testing categories that determine whether a question is acceptable or not. To reduce such errors, it is recommended to add grammatical rules for sentences containing conjunctions to ensure that the generated questions are clearer and easier to understand.

Table 7. Examples of Errors in Non-Factoid Questions of the Reason Category.

Questions
Question : *Mengapa oleh ?* Answer : *a.) virus influenza tipe ini dapat menyebabkan epidemik pada hewan epizootik dan berbagai spesies hewan dalam cakupan wilayah yang luas panzootik* *b.) ekor* *c.) virus corona paramyxovirus menyerang sistem pernafasan* *d.) itu ada baiknya untuk selalu berhati hati dengan hewan liar dan melakukan vaksinasi untuk hewan peliharaan di rumah* (Real Answer)
Question : *Mengapa murah ?* Answer : *a.) revolusi industri* *b.) harga barang-barang kebutuhan* (Real Answer) *c.) negara-negara* *d.) perubahan dalam bidang ekonomi yang sangat cepat*

Tabel 7. Examples of Errors in Factoid Questions.

Questions
Question *: Apa yang ada yang hanya hidup* Answer *:* *a.) kulit vertebrata* *b.) vertebrata* *c.) jenis caecilian* (Real Answer) *d.) latin, reptil melata*
Question *: Apa yang dihasilkan terdapat di daerah perut* Answer *:* *a.) organisme yang paling beragam hadir di bumi* *b.) notokord* *c.) telur aves bercangkang* *d.) susu* (Real Answer)
Question *: Apa yang terdapat di dalam lambung berfungsi* Answer *:* *a.) mudah dipengaruhi oleh kondisi lingkungan misalnya suhu dan ph* *b.) katalis* *c.) substrat* *d.) enzim* (Real Answer)
Question *: Apa yang memiliki tubuh* Answer *:* *a.) manusia purba* (Real Answer) *b.) manusia purba* *c.) zaman kuarter sendiri* *d.) kedua*
Question *: Apa yang keenam pemindaian dengan bahan* Answer *:* *a.) teknologi sistem peredaran darah* (Real Answer) *b.) untuk mengembalikan pasokan darah ke jantung dengan cara membuat saluran baru supaya aliran darah ke jantungnya lancar* *c.) teknologi sistem peredaran darah* *d.) teknologi kelima*

4. CONCLUSION

The template-based method can be utilized to automatically generate multiple-choice questions, encompassing both factoid questions in the form of "Apa" (what), "Siapa" (who), and "Dimana" (where) questions, and non-factoid questions in the form of definition questions, methods, and reasons.

The total number of questions generated by the system is 2468. Among these, 2132 questions were deemed acceptable based on the test categories, while 336 questions were not accepted. The overall accuracy percentage, calculated by dividing the number of accepted questions by the total number of questions, is 86.38%, this result shows promising achievement from previous research. The gap in accuracy increment is obtained through this methodology.

Based on the results of system testing, it shown that this research is good but still need more improvement, as there are still errors in the questions generated, both in factoid and non-factoid category questions. These errors include issues arising from preprocessing and inaccuracies in the question generation process. This will also open another methodology such as Indonesian BERT deep learning to be generated for multiple choice question in Indonesian language.

ACKNOWLEDGMENT

The authors would like to thank profusely for the financial support provided by DIPA funds Polmed (Politeknik Negeri Medan) to carry out research with contract number: B/437/PL5/PT.01.05/2023. The funds received from the DIPA POLMED fund played a very important role in the successful completion of this study. All authors share equal contributions in the completion of this study. Sharfina Faza and Romi Fadillah Rahmat responsible to design the methodology, implementation are in Dani Gunawan and Gabriela Dwi Lady, while Silmi, Rina and Farhat responsible for data preparation and publication improvement.

REFERENCES:

[1] S.-Y. Tan, C.-C. Kiu, and D. Lukose, "Evaluating Multiple Choice Question Generator," in *Knowledge Technology*, Berlin, Heidelberg, 2012, pp. 283-292: Springer Berlin Heidelberg.

[2] A. Aminudin, A. SN, and B. Ahmad, "Automatic Question Generation (AQG) dari Dokumen Teks Bahasa Indonesia Berdasarkan Non-Factoid Quesion," *2018,* Automatic Question Generation (AQG); Naive Bayes Classifier; Chunking Labelling vol. 5, no. 2, p. 8, 2018-05-24 2018.

[3] M. Liu, R. Calvo, and V. Rus, "G-Asks: An Intelligent Automatic Question Generation System for Academic Writing Support," *Dialogue and Discourse,* vol. 3, pp. 101-124, 03/16 2012.

[4] O. Rodríguez Rocha and C. Faron-Zucker, *Automatic Generation Of Educational Quizzes From Domain Ontologies*. 2017, pp. 4024-4030.

[5] A. Bongir, V. Attar, and R. Janardhanan, *Automated Quiz Generator*. 2018, pp. 174-188.

[6] K. K. P. Dendy Dwi Ginanjar, "Pembangkit Pertanyaan Otomatis Untuk Teks Berbahasa Indonesia Yang Mengandung Kalimat Majemuk," 2018.

[7] M. Ali and F. M. Falakh, "Design of Vocational Education Self-Evaluation System Based-on

Semantic Web Ontology," *Journal of Physics: Conference Series,* vol. 1737, no. 1, p. 012024, 2021/01/01 2021.

[8] A. Abdelghany, N. Ramadan, and H. Hefni, "An Agile Methodology for Ontology Development," *International Journal of Intelligent Engineering and Systems,* vol. 12, pp. 170-181, 04/30 2019.

[9] L. Bunnell, K.-M. Osei-Bryson, and V. Y. Yoon, "Development of a consumer financial goals ontology for use with FinTech applications for improving financial capability," *Expert Systems with Applications,* vol. 165, p. 113843, 2021/03/01/ 2021.

[10] N. Noy and D. McGuinness, "Ontology Development 101: A Guide to Creating Your First Ontology," *Knowledge Systems Laboratory,* vol. 32, 01/01 2001.

[11] P. Thapar and L. S. Sharma, "Implementing SPARQL-Based Prefiltering on Jena Fuseki TDB Store to Reduce the Semantic Web Services Search Space," in *Evolutionary Computing and Mobile Sustainable Networks*, Singapore, 2022, pp. 319-333: Springer Singapore.

[12] H. Wijayanto, *Monograf - Pencarian Katalog Buku berbasis Semantik Web*. 2020.

[13] T. Segaran, J. Taylor, and C. Evans, *Programming the Semantic Web*. 2009.

[14] T. Kanimozhi and D. A. Christy, "Applications of Ontology and Semantic Web in Image Retrieval and Research Issues," 2014.

[15] D. Kumawat and V. Jain, "POS Tagging Approaches: A Comparison," *International Journal of Computer Applications,* vol. 118, pp. 32-38, 05/20 2015.

[16] P. Sun, X. Yang, X. Zhao, and Z. Wang, "An Overview of Named Entity Recognition," in *2018 International Conference on Asian Language Processing (IALP)*, 2018, pp. 273-278.

[17] S. F. Kusuma, R. Z. Alhamri, D. O. Siahaan, C. Fatichah, and M. F. Naufal, "Indonesian Question Generation Based on Bloom's Taxonomy Using Text Analysis," in *2018 International Seminar on Intelligent Technology and Its Applications (ISITIA)*, 2018, pp. 269-274.