

PAPER • OPEN ACCESS

Intelligence chatbot for Indonesian law on electronic information and transaction

To cite this article: V A H Firdaus *et al* 2020 *IOP Conf. Ser.: Mater. Sci. Eng.* **830** 022089

View the [article online](#) for updates and enhancements.

You may also like

- [Developing Facebook Chatbot Based on Deep Learning Using RASA Framework for University Enquiries](#)

Yurio Windiatmoko, Ridho Rahmadi and Ahmad Fathan Hidayatullah

- [An AI-assisted chatbot for radiation safety education in radiotherapy](#)

David Kovacek and James C L Chow

- [Intelligent Chatbot Adapted from Question and Answer System Using RNN-LSTM Model](#)

P Anki, A Bustamam, H S Al-Ash et al.



245th ECS Meeting • May 26-30, 2024 • San Francisco, CA

Present your work at the leading electrochemistry & solid-state science conference.

Network with academic, government, and industry influencers!

Submit abstracts by December 1, 2023

[Learn more & submit!](#)



Intelligence chatbot for Indonesian law on electronic information and transaction

V A H Firdaus*, P Y Saputra, and D Suprianto

Informatics Engineering Department, State Polytechnic of Malang, Indonesia

*vipkas@polinema.ac.id

Abstract. In the Law Fiction Theory, it is assumed that once the government enforces legal norms at that time, everyone is considered to know the law. So that someone's ignorance of the law cannot free him from lawsuits. This condition led to the emergence of legal cases caused by a lack of public understanding of the law. This paper introduces a potential solution by providing chatbots platforms. We propose chatbots designed to provide information, for those who need information about applicable laws. Users can ask about anything about applicable legal documents. Furthermore, the bot performs a search according to requests related to legal documents. Various request commands are given so that bots can behave like humans and provide the information needed by users. The experimental results have shown that chatbot could recognize all of the questions from user and could answer correctly.

1. Introduction

In the field of law, there is a theory of legal fiction. Indonesia, as a country of law also applies the theory of legal fiction. According to Agus Surono, in his book defines that "Legal Fiction Theory assumes that once a legal norm is enforced, then at that time everyone is considered to know the law." So that everyone is considered to know (presumption iures de iure) and the provisions are binding so that ignorance someone will law can not free / forgive him from lawsuits (ignorantia jurist non excusat).

The existence of the legal fiction principle has been normalized in the explanation of Article 81 of the provisions of Law Number 12 of 2011 concerning Legislation namely "With the enactment of the Statutory Regulations in the official sheets as referred to in this provision, everyone is considered to have known." The Law Fiction Theory is implemented as part of the substance that regulates the enactment, namely in Article 45 of Law Number 10 Year 2004 concerning Formation of Laws and Regulations, with the explanation that reads "With the enactment of Laws and Regulations in the official sheet as referred to in the provisions of then everyone is assumed to know it."

As a result of the application of this legal fiction theory, a legal understanding is formed, where the public is considered to know the law. From this condition, there are problems, namely the emergence of many legal cases caused by the public's lack of understanding of the applicable law. Or even the public does not know that there are laws and regulations governing what is alleged or alleged to the people who are ensnared by the intended law.

Associated with the accessibility of the people to laws and regulations, the Indonesian government emphasizes the importance of legal socialization. One of the efforts made by the government is to disseminate every legislative product to the public. The obligation to spread the regulation was then set



forth by the President in Presidential Regulation No. 1 of 2007 concerning Ratification, Enactment and Dissemination of Laws and Regulations.

Based on the problems associated with the existence of legal fiction theory, the writer tries to provide a solution, namely by developing a chatbot to search for legal documents in Indonesia that specifically focuses on the Law on Information and Electronic Transactions. This research is in line with the form of legal socialization support at point 3, namely making applications that make it easier for the public to access the legislation. as well as in today's digital era communication with chat and speech is an important and inseparable part of everyday life [1], so it will be very helpful in socializing to the public.

This research utilizes Chatbot as its application, Chatbot is a computer program used to simulate an interactive conversation or communication for users [2, 3]. The response generated is a keyword in the user's input and produces a response that is considered the most suitable response from the database [4], so that the conversation that occurs as if carried out by two human persons who communicate with each other.

In developing this chatbot, the authors chose the Telegram instant messaging application platform. Telegram was chosen because, the first Telegram has many users, namely more than 200 million users. The second reason why Telegram was chosen is because the support for chatbot development is quite high where there are many libraries that can be utilized in developing chatbots.

In general, the chatbot application that will be developed by the author is aimed at facilitating users to search for laws by typing keywords in the conversation textbox. Another important benefit that can be used in implementing NLP on chatbots is by voice or text interaction [5]. Furthermore in our system, the system will send a request to the server based on these keywords. In the server section, the database is searched, to find documents that match the keywords. If you have found the document in question, the server will send the results to the user. The search results will be displayed as replies to conversations from keywords that have been typed by previous users. In this research, what will be used as a case study is the indonesia legal law namely ITE (Informasi dan Transaksi Elektronik) Law, which is Law No. 19 of 2016.

2. Methods

Our research aims to make it easy for users to interact with the chatbot system in seeking information about applicable ITE laws and regulations. Some functions in the context of conversations between users and the system will be used in the chatbot application [6]. In its implementation, input from the user to the chatbot application is not enough to provide the correct answer to the user, so it is necessary to have NLP as a method to be used together with chatbot so that the system can recognize the context of the user accurately [7]. We designed the system to be able to recognize every user's input by applying Natural Language Processing. Chatbot application using NLP is a system that we designed to help users get accurate information about ITE laws, where the system automatically responds by answering questions from users. In this system, we use the NLP method, which is a parser, lexicon, understander, knowledge base, and generator to get answers that correspond to the question in question. The system design is illustrated through the system flow design, the concept of the system flow can be seen in Figure 1. In our system consists of several parts of data processing involving the use of the NLP process. Here are the system designs that we proposed.

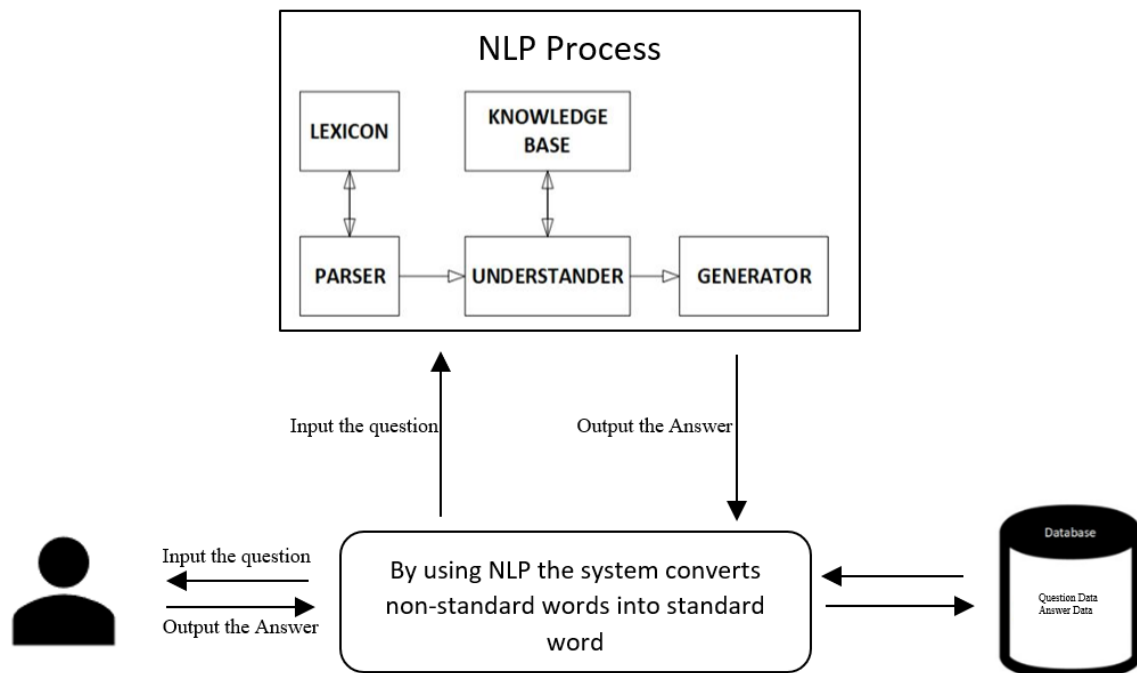


Figure 1. Network Development Life Cycle (NDLC).

The process of processing the chatbot system using the approach of NLP and implement some method other in NLP process such . Questions from the user become an input form for the system which will then be processed with the NLP approach followed by applying the laveinshtein distance and tf-idf consine similarity methods. With the laveinshtein distance method, it will be found the closest word for the conversion of non-standard words into standard words so that the questions will still be recognized by the system to then be processed with the intent in accordance with the user's input. The e tf-idf consine similarity method is used to compare consine similarity values between user questions and documents (questions in the database), consine similarity values of the largest documents that will be displayed in response to user questions.

In the system that we have developed, database used as store files or data are mutually related and can be accessed directly from the system, which consists of: The data file question is a data file that contains fields questions , and file response data yan is bi answer file sa directly accessed by the system.

The Natural Language Processing (NLP) method is the steps in processing a sentence to be recognized by the system [8] and then the system can provide answers to the user. The stages of the process are carried out by carrying out several stages, including: Parser, Lexicon, Understander, Knowledge Base, and Generator, as shown in Figure 2. In the final stages of data processing output will be obtained in the form of answers to questions asked by users.

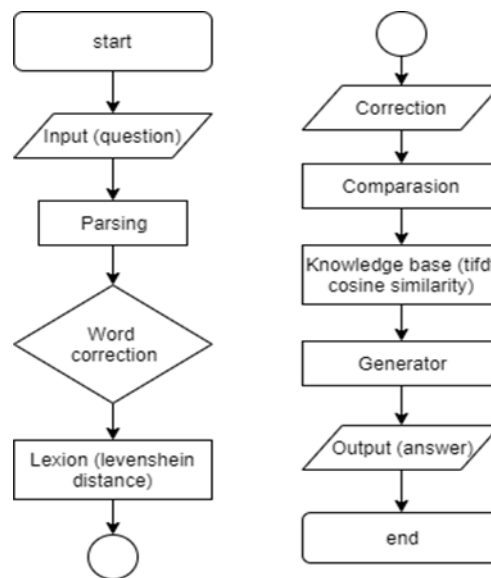


Figure 2. Network Development Life Cycle (NDLC).

First step in our system, user enter a sentence input in the application chatbot. Next step, system will process the sentence input that has been entered by the user, after sentence in a process that will continue in the parser [9]. Parsers function in analyzing the input, namely by checking the truth and sequence of the tokens that are formed by lexical analysis [10, 11]. In this process a noise disposal parser is used which uses a database for the analysis process.

In the process parser things first do is to change each letter into a letter of small or lowercase and break down each word of the sentence. After the sentence in broken into a collection of words, and then do the correction word to analyze the syntax of the sentence by comparing one by one word in a sentence input with the entire word in the database.

The last step is the lexicon, the words that have been separated to be in check one by one to do the justification of the word. Justification said in the parser is intended to carry out the conversion of the word that is the correction word is not standard on the mistake of writing the word. So that each word becomes input data (source word) for the word correction process by searching for a word in the dictionary (lexicon) that has a minimum edit distance to the target word. In this process using the calculation of the value of the edit distance Lavenshtein Distance. Rated distance will be calculated at each meeting of row and column begins from position of the index row and column of the first to the last of row and column as the result of the end of the value range edit. Any error in the form of loss of character letters, the excess characters letters, or error sequence of letters of two-character letters are different.

3. Results and discussion

We use the Black Box Testing for testing the system, Black box testing is a technique of testing that focuses on the output result of the response inputs, or in simple black box is the process of running the application to determine whether there is an error or there is a function that does not run in accordance expectations . Black box testing ignores the mechanism of internal systems, such as how the system works to process the input. The results of the testing are obtained the result that the system can be run with either. After testing the functional completion done, we do test methods. Method testing is done by calculating TFIDF and creating test scenarios. The scenario that is used by the method of Natural Language Processing. Results of the calculations TFIDF obtained as in Figure 3 and Figure 4.

```

-----NILAI TFIDF-----
Nilai TFIDF 1 =0
Nilai TFIDF 2 =0
Nilai TFIDF 3 =0
Nilai TFIDF 4 =0.24760365662646577
Nilai TFIDF 5 =0
Nilai TFIDF 6 =0
Nilai TFIDF 7 =0
Nilai TFIDF 1 =0
Nilai TFIDF 2 =0
Nilai TFIDF 3 =0
Nilai TFIDF 4 =0.2200921392235251
Nilai TFIDF 5 =0
Nilai TFIDF 6 =0

```

Figure 3. TFIDF.

```

-----NILAI TFIDF-COSINE -----
Nilai TFIDF-COSINE 1 =0.014796809304397305
Nilai TFIDF-COSINE 2 =0.011691306117054658
Nilai TFIDF-COSINE 3 =0.095499419690923
Nilai TFIDF-COSINE 4 =0.026305438763372976
Nilai TFIDF-COSINE 5 =0.0
Nilai TFIDF-COSINE 6 =0.0
Nilai TFIDF-COSINE 7 =0.0
Nilai TFIDF-COSINE 8 =0.056936002343520815
Nilai TFIDF-COSINE 9 =0.019326444805743418
Nilai TFIDF-COSINE 10 =0.0
Nilai TFIDF-COSINE 11 =0.0
Nilai TFIDF-COSINE 12 =0.0
Nilai TFIDF-COSINE 13 =0.0
Nilai TFIDF-COSINE 14 =0.0
Nilai TFIDF-COSINE 15 =0.0
Nilai TFIDF-COSINE 16 =0.0

```

Figure 4. TFIDF Cosine.

Testing with several scenarios is also done to see the accuracy of the developed chatbot system. There are three scenarios were used. First Scenario is a “using questions that come from the answers in the knowledge base”, and obtained the results is an average of 100%. Second scenario is the “using questions Random” and can concluded that testing result showed percentage matches the system with the data actually is an average of 77%. And the last scenario is the "using the Knowledge Base". Any question is still related to Knowledge Base so that obtained results of an accuracy of 83%.

Table 1. Precisson and recall.

No	Precisson	Recall
1	8/10 = 0.8	8/15 = 0.5
2	8/10 = 0.8	8/15 = 0.5
3	8/10 = 0.8	8/15 = 0.5
4	8/10 = 0.8	8/15 = 0.5
5	8/10 = 0.8	8/15 = 0.5
6	8/10 = 0.8	8/15 = 0.5
7	8/10 = 0.8	8/15 = 0.5
8	8/10 = 0.8	8/15 = 0.5
9	8/10 = 0.8	8/15 = 0.5
10	8/10 = 0.8	8/15 = 0.5

The last testing as shown in Table 1, we do test by using the Recall and Precision, the test use 10 question data. Recall is the proportion of the number of documents that can be recovered by a search process in the information retrieval system. We found the percentage of recall which is equal to 0.5. And Precision is the proportion of the number of documents found and considered relevant to the ratio of the number of relevant documents found back to the total number of documents in the document that is considered relevant. The result percentage of precision, which is equal to 0, 8

4. Conclusion

From the test results, it can be concluded that all system in chatbot application can run properly, based on the scenario that has been created. The system also has successfully applied the natural language processing method for word conversion on the chatbot system automatically. The test results using recall and Precision use 10 question data, the results obtained recall of 0.5 or 50%. Then the Precision results obtained by 0.8 or 80%.

In the future work, need to improve the vocabulary database for the staging process by using SPOK to do word conversions. And system optimization needs to be further improved when processing word improvements

Acknowledgment

The authors would like to thank to the Indonesian Directorate General of the Higher Education (DIKTI) and State Polytechnic of Malang who have supported this research project.

References

- [1] A. S and John Dr 2015 Survey on Chatbot Design Techniques in Speech Conversation Systems *International Journal of Advanced Computer Science and Applications*
- [2] Akhsan A A and Faizah F 2017 Analisis dan Perancangan Interaksi Chatbot Reminder dengan User-Centered Design *Jurnal Sistem Informasi*
- [3] Ciechanowski L, Przegalinska A, Magnuski M and Gloor P 2019 In the shades of the uncanny valley: An experimental study of human–chatbot interaction *Future Generation Computer Systems*
- [4] Handoyo E, Arfan M, Soetrisno Y A A, Somantri M, Sofwan A and Sinuraya E W 2018 Ticketing Chatbot Service using Serverless NLP Technology *Proceedings - 2018 5th International Conference on Information Technology, Computer and Electrical Engineering, ICITACEE 2018*
- [5] Uma M, Sneha V, Sneha G, Bhuvana J and Bharathi B 2019 Formation of SQL from natural language query using NLP *ICCIDS 2019 - 2nd International Conference on Computational Intelligence in Data Science, Proceedings* 1–5
- [6] Lieberman H and Mason C 2002 Intelligent agent software for medicine *Studies in Health Technology and Informatics*
- [7] Graesser A C, Chipman P, Haynes B C and Olney A 2005 Auto tutor: An intelligent tutoring system with mixed-initiative dialogue *IEEE Transactions on Education*
- [8] Jia J 2009 CSIEC: A computer assisted English learning chatbot based on textual knowledge and reasoning *Knowledge-Based Systems*
- [9] Rahman A M, al Mamun A and Islam A 2018 Programming challenges of chatbot: Current and future prospective *5th IEEE Region 10 Humanitarian Technology Conference 2017, R10-HTC 2017*
- [10] Huang B B, Zhang G and Sheu P C Y 2008 A natural language database interface based on a probabilistic context free grammar *Proceedings - 1st IEEE International Workshop on Semantic Computing and Systems, WSCS 2008*
- [11] Singh G and Solanki A 2016 An algorithm to transform natural language into SQL queries for relational databases *Selforganizology*