

# Development of the text analysis software agent (chat bot) for the library based on the question and answer system TWIN

Anastasia Ivanovskaya<sup>1</sup>, Konstantin Aksyonov<sup>1,\*</sup>, Igor Kalinin<sup>1</sup>, Yuriy Chiryshev<sup>1</sup>, and Olga Aksyonova<sup>1</sup>

<sup>1</sup>Ural Federal University, 620002 Mira 32, Ekaterinburg, Russia

**Abstract.** The development of a text analysis software agent is presented for a library based on the TWIN question-response system. A review of modern platforms for creating chat bots. The results of experiments with a trained text analysis software agent are described. The trained agent fully provided correct information during the experiments.

## 1 Introduction

Information technologies are applied in all spheres of human activity, this allows to reduce the time for processing a large amount of information. For example, in the library system, when the reader asks for help in selecting a book, the librarian often has difficulty quickly providing an answer about the availability of a book or advice on choosing literature. These difficulties give rise to problems of quick help and relevant search. To solve these problems, a text analysis software agent based on the TWIN question-answer system was developed.

## 2 Overview of question-answer systems for the processing of Russian speech, focused on the voice interface and chat-bots interface

A question-answer system is an information system that is a hybrid of search, reference, and intelligent systems that uses a natural-language interface. The input to such a system is a request formulated in natural language, after which it is processed using NLP methods, and a natural language response is generated. The QA system uses either local storage or a global network, or both, as a source of information [1]. At the moment, there are many cloud solutions for the needs of NLP and the development of software agents for analyzing texts based on the question-answer system. Table 1 provides an overview of several chat bots that can recognize text and voice [2]. TWIN question-answer system was chosen for the development of a text analysis software agent for the library.

---

\* Corresponding author: [wiper99@mail.ru](mailto:wiper99@mail.ru)

**Table 1.** Overview of platforms for building voice and chat bots.

Chat development environment	bot	Cost	Machine learning simulation	Messaging Platform Integration	Notes
Wit.ai (Facebook, 2013) <a href="https://wit.ai">https://wit.ai</a>		free	Training with examples	Not. Web Services API	There is a limit in the ability to transmit only one request per second.
Api.ai (Google, 2010) <a href="https://api.ai">https://api.ai</a>		free	Learning is present	Has paid funds for deployment on Heroku	
IBM WATSON, (IBM, 2010) <a href="https://www.ibm.com/watson">https://www.ibm.com/watson</a>		Trial free version 30 days	Learning is present	Integrates with Azure	There is a large functionality that will require skills and a long time to master.
TWIN, (TWIN, 2018) <a href="https://twin24.ai/">https://twin24.ai/</a>		free	Training with examples	It has integration with various communication systems.	

### 3 Description of the TWIN system

The TWIN system is an omni channel communication platform for building head-on and chat bots, which is able, in particular, to receive data in one language and transmit it in another. TWIN can keep voice and text recording, display detailed statistics and analytics for each call or dialogue. TWIN system supports and integrates 6 communication channels - SMS, calls, instant messengers, online chat, mail, social networks.

The platform provides the ability to connect to the customer's IP-PBX, use of its telecom operator and integration with virtually any CRM or ERP system. The robot can work with any modern IP-PBX. The basis of the formation of TWIN knowledge is the work with Big Data and neural networks [3].

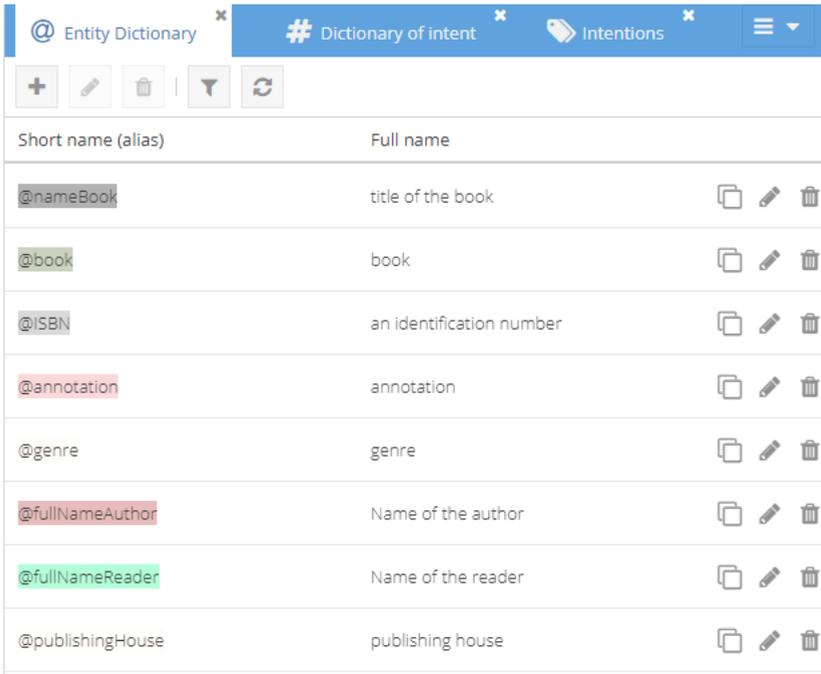
### 4 Application of TWIN system

It is necessary to distinguish named entities and intentions before developing a software agent. Named entity is a word or phrase intended for a specific, well-defined object or phenomenon, distinguishing this object or phenomenon from a number of similar objects or phenomena. Intention can be defined as the meaning of what has been said, i.e. what the user meant when he said a certain phrase [4].

The main entities of this subject area are the following (Figure 1): book; author name; genre; description.

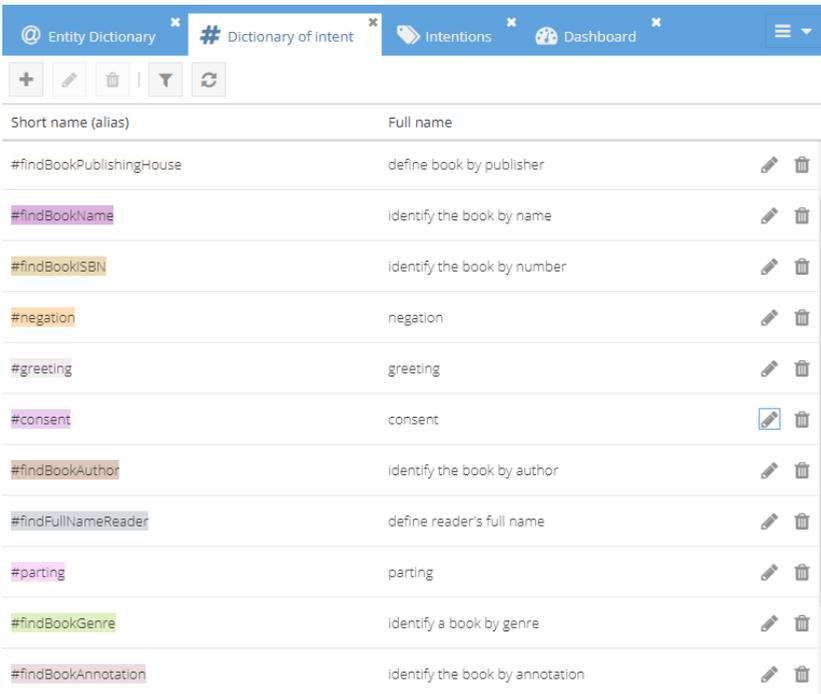
The following intentions are the ones of this subject area: determine the year of publication; define a book description; identify a book by name; identify a book by ISBN; identify a book by author; identify a book by genre (see. fig. 2).

Composite intentions solve the main problem of understanding human speech, namely, people very often express several intentions in one phrase. Therefore, the markup of compound phrases is carried out after adding entities and intentions for the task of training the agent and thereby the agent's knowledge base is filled. After the preparatory actions, the agent was trained.



Short name (alias)	Full name	
@nameBook	title of the book	  
@book	book	  
@ISBN	an identification number	  
@annotation	annotation	  
@genre	genre	  
@fullNameAuthor	Name of the author	  
@fullNameReader	Name of the reader	  
@publishingHouse	publishing house	  

**Fig. 1.** The entity of the text analysis software agent for the library.



Short name (alias)	Full name	
#findBookPublishingHouse	define book by publisher	 
#findBookName	Identify the book by name	 
#findBookISBN	Identify the book by number	 
#negation	negation	 
#greeting	greeting	 
#consent	consent	 
#findBookAuthor	Identify the book by author	 
#findFullNameReader	define reader's full name	 
#parting	parting	 
#findBookGenre	Identify a book by genre	 
#findBookAnnotation	Identify the book by annotation	 

**Fig. 2.** The intent of the text analysis software agent for the library.

After filling the agent's knowledge base, the agent is trained using a neural network. The results of processing a phrase by a trained agent are presented in Figure 3.

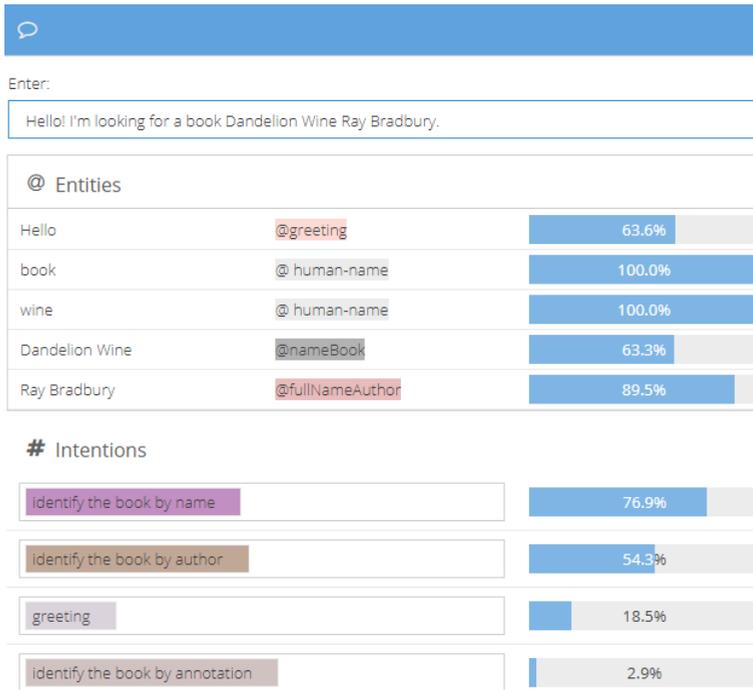


Fig. 3. An example of processing a phrase by a text analysis program agent.

Figure 4 shows a summary chart that displays the percentage of recognition of intentions.

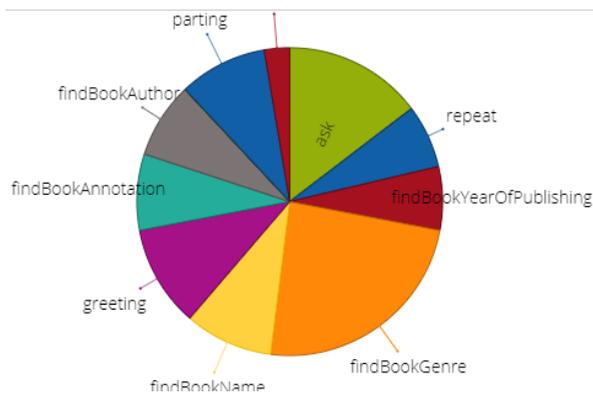
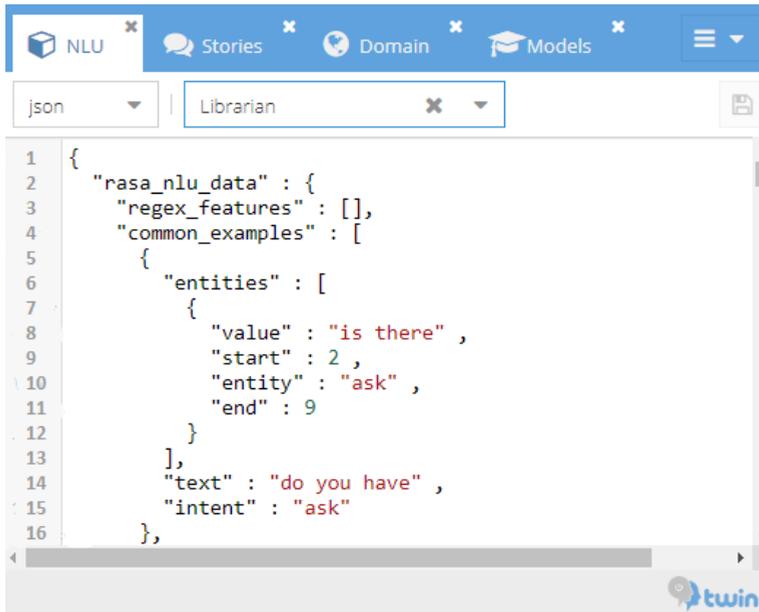


Fig. 4. Recognized Intent Relation Chart.

The TWIN system is developed using the RASA platform [5]. This is an open and very popular chat bot framework. It consists of two independent components, Rasa NLU and Rasa Core. Rasa NLU (Natural Language Understanding). The main goal of this component is to convert user input in natural language into objects that the program can work with. Rasa Core - this component is responsible for creating a chat bot script based on intentions and entities.

The previously created agent is automatically transferred to the synonyms directory in the NLU section (see figure 5), while determining the text of the users message; the intent which should be associated with the text; the entities are specific parts of the text which need to be identified.



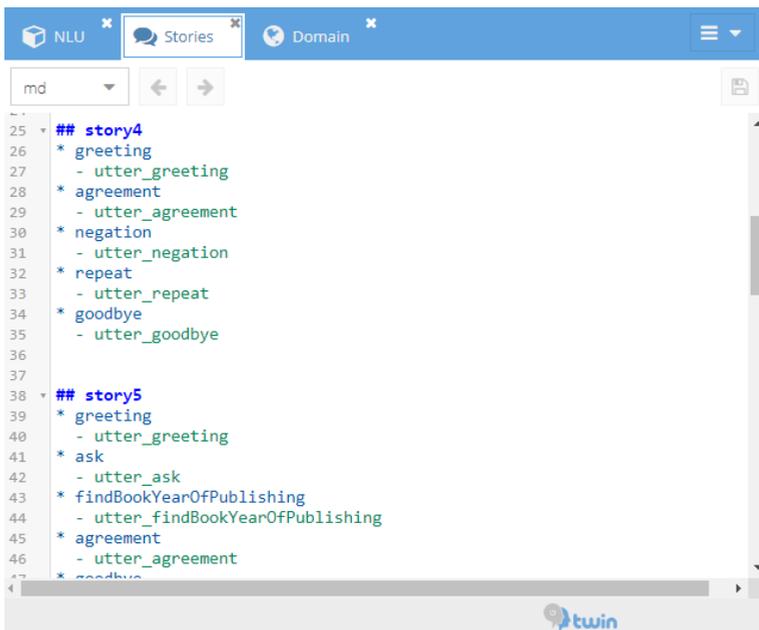
```

1  {
2  "rasa_nlu_data" : {
3    "regex_features" : [],
4    "common_examples" : [
5      {
6        "entities" : [
7          {
8            "value" : "is there" ,
9            "start" : 2 ,
10           "entity" : "ask" ,
11           "end" : 9
12         }
13       ],
14       "text" : "do you have" ,
15       "intent" : "ask"
16     }

```

**Fig. 5.** Synonyms directory.

The Stories page defines training examples for the interactive system (see figure 6). A story starts with a name preceded by two hashes `## story1`. Messages sent by the user are shown as lines starting with `*` in the format `intent{"entity1": "value", "entity2": "value"}`. Actions executed by the bot are shown as lines starting with `-` and contain the name of the action.



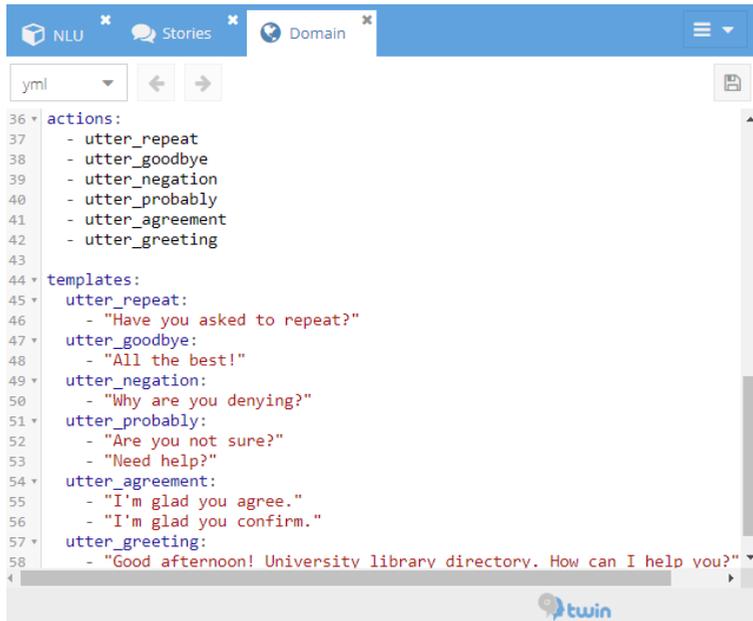
```

25  ## story4
26  * greeting
27  - utter_greeting
28  * agreement
29  - utter_agreement
30  * negation
31  - utter_negation
32  * repeat
33  - utter_repeat
34  * goodbye
35  - utter_goodbye
36
37
38  ## story5
39  * greeting
40  - utter_greeting
41  * ask
42  - utter_ask
43  * findBookYearOfPublishing
44  - utter_findBookYearOfPublishing
45  * agreement
46  - utter_agreement
47  * goodbye

```

**Fig. 6.** Examples for the interactive system.

The Domain defines the universe in which assistant operates (see figure 7). It specifies the intents, entities, slots, and actions bot should know about. Also, phrases patterns that the bot can speak are defined in this section.



```

36 actions:
37   - utter_repeat
38   - utter_goodbye
39   - utter_negation
40   - utter_probably
41   - utter_agreement
42   - utter_greeting
43
44 templates:
45   utter_repeat:
46     - "Have you asked to repeat?"
47   utter_goodbye:
48     - "All the best!"
49   utter_negation:
50     - "Why are you denying?"
51   utter_probably:
52     - "Are you not sure?"
53     - "Need help?"
54   utter_agreement:
55     - "I'm glad you agree."
56     - "I'm glad you confirm."
57   utter_greeting:
58     - "Good afternoon! University library directory. How can I help you?"

```

**Fig. 6.** Intentions, entities, slots and actions of the bot.

A chat bot model was created after all the steps taken. The bot is trained during the creation of the model. Further, experiments with the bot were carried out with communication during which the developed bot demonstrated accurate and necessary answers to user replicas.

## 4 Conclusion

The developed text analysis software agent (chat bot) for the library on the basis of the question-answer system solves the problems posed: it allows you to answer quickly interesting questions from the reader about the availability and selection of books. The trained agent fully provided correct information during the experiments.

This work is supported by Act 211 Government of the Russian Federation, contract № 02.A03.21.0006.

## References

1. A. Nikitin, P. Raikov. *Question-answer systems* (2018) Access mode: <https://gigabaza.ru/doc/67598.html>
2. *Comparative analysis of chatbots API* (2018) Access mode: <https://activewizards.com/blog/a-comparative-analysis-of-chatbots-apis/>
3. *Future Bunking. Twin - omni-channel communication platform for building voice and chat bots* (2018) Access mode: <http://futurebanking.ru/post/3647>
4. K. Aksonov, D. Antipin, T. Afanaseva, I. Kalinin, I. Evdokimov, A. Shevchuk, A. Karavaev, U. Chiryshv, E. Talancev. *Testing of the speech recognition systems using Russian language models*. CEUR, **2298**, 7 (2018), <http://ceur-ws.org/Vol-2298/paper13.pdf>
5. *Opencorpora. Instructions for determining named entities* [Electronic resource] - Access Mode: <http://opencorpora.org/wiki/Nermanual/1>