Hispabot-Covid19: the official Spanish conversational system about Covid-19

David Griol[1], Zoraida Callejas[1], David Pérez-Fernández[2]

[1]Department of Languages and Computer Systems, University of Granada,
Periodista Daniel Saucedo Aranda s.n., 18071, Granada, Spain
[2]Ministry of Inclusion, Social Security and Migration,
Paseo de la Castellana 63, 28071, Madrid, Spain

dgriol@ugr.es, zoraida@ugr.es, david.perez@inv.uam.es

Abstract

Hispabot-Covid19 is a conversational system developed for the Spanish Government to provide responses to frequently asked questions related to the pandemic originated by Covid-19 and its implications in Spain. The system received more than 350,000 queries between April and June 2020, being a clear example of how conversational systems can be applied to reduce the pressure on the health emergency phone lines and to provide 24/7 access to information and services using natural language. In this paper, we describe the main features of the Hispabot-Covid19 system.

Index Terms: Covid-19, Conversational Systems, FAQ, Spanish.

1. Introduction

The Secretary of State for Digitalization and Artificial Intelligence (SEDIA)[1], through the Spanish Plan for the Advancement of Language Technology[2], coordinated in March 2020 the development of a conversational assistant to answer frequent questions about COVID-19.

The Hispabot-Covid19 conversational system [1] accessed information from official sources from the Spanish Ministry of Health and the World Health Organization to report on various issues, such as: symptoms, vulnerable groups, transmission, prevention, coexistence with infected people, conditions for quarantine and isolation, attention telephone numbers, among many others. The assistant also incorporated the information published in the Official Spanish State Bulletin (Boletín Oficial del Estado, BOE) regarding the application of the *State of Alarm* and the *Plan for the Transition towards a New Normality*. The assistant did not retrieve or analyze any personal data.

Hispabot-Covid19 is an example of the application of conversational systems as a solution to provide information to citizens and facilitate contact with public institutions [2, 3, 4]. This objective can be achieved with:

1. systems for solving frequent doubts related to public services, so that it is possible to have automated FAQs where questions can be asked in natural language. In this case the questions are isolated, it is not necessary to take into account the questions previously asked by the citizen.

2. systems that can hold a dialogue with the citizen to automate services, e.g. to make arrangements with an administration. In this case it is necessary to have systems that process the previous history of the dialogue.

Hispabot-Covid19 falls within the first group. For the development of the system it has been necessary to develop the different modules and models, deploy it in several interaction channels (web-based assistant, Telegram and WhatsApp), and optimize the system on the basis of the observed interactions. For the development phase, a model was built for language processing. For this purpose, a list of the main entities or concepts handled (e.g. autonomous communities, synonyms of coronavirus, names of related diseases, etc.) has been elaborated and, on the other hand, a training of the natural language understanding model with an initial corpus of training phrases was carried out.

All dialogues have been recorded in the system and processed for continuous retraining and improvement of the assistant. To this end, new training phrases and new question categories were incorporated daily, misunderstandings have been detected and corrected, and the knowledge base has been restructured according to updates in the information provided.

On April 3rd 2020, the Ministry of Economic Affairs and Digital Transformation launched a pilot project to firstly integrate the assistant into the web portals of the Government of La Rioja. From April 3rd to 7th, the assistant received more than 5,000 queries on these portals. From April 8th until the end of June 2020 it was accessible through WhatsApp and Telegram. More than 350,000 interactions took place during this period.

In this paper we describe the main features of the Hispabot-Covid19 system, the main phases that were followed to develop the system, the set categories of FAQs that were defined and the main statistics related to the operation of the system.

2. State of the art

Establishing a natural, agile and fluent conversation with a machine using natural language has been one of the main research challenges in the fields of Natural Language Processing, Computational Linguistics and Speech Technologies. This challenge has captured since years ago the interest in the academic, commercial and industrial fields, especially considering the wide range of applications of this kind of systems [2].

In fact, several recent reports highlight how the use of natural language (and mainly voice) is changing the way we relate to technology, with growth prospects in technologies, sectors and devices that indicate that we are in the "era of conversational interfaces" and that have been accentuated during the pandemic caused by Covid-19 with the increased use of e-government.

The reference legislation on e-Government in Spain dates back to June 2007. The circumstances of the pandemic have prevented the staff of the Public Administrations from performing their functions in the offices, have generalized the implementation of teleworking tools and have also made possible a

---

[1]https://avancedigital.mineco.gob.es (Last access: February 2021)
[2]https://plantl.mineco.gob.es (Last access: February 2021)

Figure 1: *Screenshots of the integration of the Hispabot-Covid19 in the website of La Rioja Salud*

generalized use of telematic means by the citizens. Information and communication technologies have played a fundamental role in maintaining the activity of the Public Administration during this crisis, and have shown the future importance of continuing with the digital transformation.

The use of electronic media has also proved to be a magnificent opportunity to accelerate the number of developments and the application of conversational systems in different areas of the Administration, both to facilitate access to information and answers to frequently asked questions, as well as to carry out e-government procedures.

Chatbots have become a tool to combat misinformation and even to channel people's anxiety about possible contagions. The use of chatbots as a strategy to deal with the pandemic has been both nationally and globally. Regarding the resolution of frequently asked questions, it is worth noting the large number of conversational systems that have been developed in a very short time worldwide to answer questions related to the pandemic caused by Covid-19, some of them also including triage services related to the symptoms of the disease: Hispabot-Covid19 [1], Carina[3], COVID19AragónBot[4], IMPAI[5], World Health Organization chatbot[6], the UCSF Health and Northwell Health chatbot[7], the CDC chatbot[8], CovidBot[9], Boti[10], etc. The usage statistics of these assistants have shown that they have become indispensable global tools in the fight against the virus, in combination with the Web channel and messaging services as an ideal environment to carry out action protocols, provide information and raise awareness about the pandemic generated by COVID-19.

3. The Hispabot-Covid19 system

As described in the introduction section, since the beginning of the pandemic, official State agencies have made a great effort to provide data and information in real time and to reduce the collapse of the citizen health telephone lines. The Hispabot-Covid19 system has contributed to this effort by providing easy access to official information, allowing citizens to ask questions in their own words, and providing a practical example of the great potential that these systems will have for communication between citizens and the Public Administration.

For the development of the system it has been necessary to work on 4 main lines:

- Supervision of contents: The Health Service of La Rioja

[3]https://1millionbot.com/chatbot-coronavirus/ (Last access: February 2021)

[4]https://www.europapress.es/aragon/noticia-itainnova-pone-marcha-covid19aragonbot-telegram-responder-dudas-ciudadanos-20200403100820.html (Last access: February 2021)

[5]https://campussanofi.es/e-professionals/noticias/impai-chatbot-coronavirus/ (Last access: February 2021)

[6]https://www.whatsapp.com/coronavirus/who (Last access: February 2021)

[7]https://www.healthcareitnews.com/news/northwell-ucsf-unc-using-chatbot-and-related-tech-manage-covid-19-patients (Last access: February 2021)

[8]https://www.theverge.com/2020/3/21/21189227/cdc-microsoft-chatbot-coronavirus-symptom-checker (Last access: February 2021)

[9]https://www.infobae.com/tecno/2020/04/01/asi-es-el-chatbot-mexicano-que-lucha-contra-la-desinformacion-sobre-coronavirus/ (Last access: February 2021)

[10]https://planetachatbot.com/buenos-aires-lanza-boti-chatbot-en-app-mas-utilizada-mundo-9ca18f2d5678 (Last access: February 2021)

Table 1: *Distribution of users' queries for each one of the intents defined for the Hispabot-Covid19 System*

Intent	Perc.	Intent	Perc.	Intent	Perc.
Visits to dental centers	0.12	Meetings during the lockdown	0.31	How to care for patients	0.48
Attendance to health centers	1.00	Rites and ceremonies	0.77	School - educational institutions	1.27
Effects in different ages	0.17	Questions without answer	2.38	Welcome and greetings	1.10
Droplet transmission	0.32	Patients with diabetes	0.23	Differences other diseases	0.28
Home isolation to avoid spread	0.78	Apologies for misunderstandings	0.69	Interpersonal distance	0.20
Nutritional advice	0.50	Official and technical documents	0.19	Donations	0.08
Hospital discharge	0.08	Blood donations	0.24	Duration of isolation for patients	0.20
Pets and transmission in animals	1.13	How to entertain children	0.31	Role of state security forces	0.12
Use of antibiotics	0.13	Pregnancy-related questions	0.50	Use of nasal swabs	0.05
Anticoagulated patients	0.12	Patients with epilepsy	0.11	Chronic respiratory diseases	0.09
Flattening the COVID-19 peak	0.25	Questions about the alarm state	0.69	Stigmatizing behaviors	0.11
Applause for healthcare workers	0.20	Explain the disease to children	0.37	Questions about pharmacies	0.23
Asymptomatic patients	0.17	Different phases and provinces	1.53	End date of lockdown.	3.12
Asthmatic and allergic patients	0.42	Time slots and allowed activities	3.43	Fruits and vegetables	0.25
Patients home care	0.16	Official sources of the contents	0.82	Thank you and goodbye	0.85
Self-employed workers	0.24	Questions on vulnerable groups	1.00	Recommendations for shopping	1.35
Emergency phone numbers	2.50	Transmission through feces	0.09	Patients with hypertension	0.42
When to use emergency phones	0.39	Tourism and accommodation	0.09	Questions about ibuprofen	0.23
Workers sick leave	0.62	How to use the system	0.36	Insults and bad language	0.23
Widespread fakes about Covid	0.80	Congresses and scientific events	0.13	Encouraging queries	0.57
Specific fake about using oil	0.04	Breastfeeding	0.20	Advice on hand washing	0.58
Specific fake about eating garlic	0.04	Washing face masks	0.23	Use of contact lenses	0.13
Specific fake about using alcohol	0.08	Use of the car and urban transport	1.55	Clean the shopping basket	0.32
Killing the virus with a hair dryer	0.04	Hygiene recommendations	0.86	Disinfection commercial spaces	0.08
Virus transmission through shoes	0.20	Disinfection of clothes	0.25	Days of national mourning	0.05
Specific fake of using UV light	0.07	Gender violence	0.31	Information about face masks	2.14
How to stop spreading fakes	0.27	Information purchase of masks	0.54	Medicines and vaccines	1.09
Influence of temperature, zones	0.28	Information on social measures	0.77	Information on virus mortality	0.09
Oncology patients	0.20	Using/cleaning mobile phones	0.33	Fines and associated regulations	0.37
COVID-19 and viral load	0.05	Not touching nose-eyes-mouth	0.05	Relation with pneumonia	0.13
Questions about cybersecurity	0.31	Access to official press releases	0.25	News and updates	0.48
Reported cases and deaths	5.51	Updates related to the system	0.11	Works, reforms and construction	0.20
How report. cases are calculated	0.36	Origin of the virus	0.29	Information on specific jobs	0.17
How the disease is spread	0.53	Meaning of pandemic	0.12	Use of paracetamol	0.31
Infection through water	0.15	Information about home orders	0.48	Specific information hairdressers	1.45
Transmission in the workplace	0.76	Incubation period of the virus	0.74	Information staying at home	0.23
Transmission by direct contact	0.36	Inform. recoverable paid leave	0.37	System's privacy policy	0.07
Neighborhood coexistence	0.64	Immunity and recovery	0.08	Inform. how to handle worry	0.28
Quarantine related questions	0.54	Private health care information	0.12	Inform. probability of infection	0.23
Inform. healthcare professionals	0.21	Protection and prevention meas.	1.49	Information about covid-19 tests	1.54
General inform. about disease	0.88	Inform. system's functionalities	1.29	What not to do	0.13
Providing the name of the bot	0.61	How to safely remove gloves	0.20	Antibodies and immunity	0.82
People with kidney conditions	0.12	Residences - sheltered housing	0.11	Agricultural and livestock activit.	0.77
People with autism spectrum	0.23	Bars, terraces and restaurant	0.76	Care of orchards and gardens	0.17
Recommend. for leaving home	0.48	Regulations for leaving home	2.68	Permissible duration of outings	0.21
Recommend. on children walks	2.97	Libraries, cinemas, theaters	0.35	Visits to beaches - green spaces	0.53
Making and receiving visits	1.22	Walking the dog	0.49	Garbage disposal	0.11
Work permits and safeguards	1.39	Use of public transport	0.48	Inform. related to mental health	1.41
General information about SARS	0.16	Seroprevalence	0.16	Share Hispabot on WhatsApp	0.32
Symptoms in children	0.65	How to act when symptoms	5.32	Essential supplies lockdown	0.21
Transmission of virus on surfaces	0.44	Information about smoking	0.19	Information about teleworking	0.31
Information about open stores	2.31	Information about scams	0.12	Inform. of working conditions	0.96
Administrative procedures	0.86	Interprovincial travel regulations	2.16	Treatment of the disease	0.27
Information on the use of gloves	0.35	Vaccine information	0.07	Incoming international travels	0.25
Outgoing international travels	1.06	Transmission area	0.19		

and the Ministry of Health participated in the revision and accessibility of the main contents provided by the conversational assistant.

• Development of the conversational system: A natural language understanding model was built using Google technologies. For this purpose, a list of 90 entities with

more of 2100 different as possible values for these entities were defined (e.g. autonomous communities, synonyms of coronavirus, names of related diseases, etc.). More than 8,000 training sentences were gradually incorporated to process a total of 164 intents (categories of questions that can be answered by the system). Table 1 shows the list of intents defined for the system. The numbers indicated in the table show the percentage of users' sentences that were classified in each category during the operation of the system. In Hispabot-Covid19 there was only one defined answer for each intent. These responses have been designed either as direct answers in text mode (including text, emojis and bold and italic highlighting) or defined after processing them using a script (e.g. to indicate the telephone number corresponding to the autonomous community indicated by the user). The most frequently asked questions corresponded with the hot topics of each day or the most prominent news in the media. For example, during the first weekend of April, the number of questions on the use of masks increased considerably, and when there was news of the infection of pets, the number of queries on the transmission of COVID-19 between humans and animals increased. The daily analysis of the queries provides very valuable information to know in real time which were the most frequent doubts of the citizens.

- Deployment in interaction channels: The system was deployed for the interaction in messaging services (WhatsApp and Telegram, Figure 1) and web environments (Figure 2) More than 350,000 interactions were recorded with the system from April to June 2020.

- System maintenance: The system was daily maintained to detect understanding failures, restructure the knowledge base according to the updates and incorporate entities, categories, and training sentences. Rates higher than 94% were achieved with regard the correct classification of users' queries in the correct intent. A set of scripts were developed to dump the data (intents and entities) to CSV format in order to have a readable version of the data for monitoring.

Figure 2: *Screenshots of the integration of the Hispabot-Covid19 assistant in WhatsApp*

4. Conclusions and future work

In an emergency situation such as the one we are currently experiencing, providing information, updates and support to citizens are key objectives. Covid-19 has accelerated the digitization process and made it urgent to provide timely and permanent support to citizens rooted in trustworthy sources of information.

In this context, the implementation of chatbots has experienced a significant increase and revealed the full potential of these systems, especially in the healthcare sector. They are a valid complementary help to provide information and services and lighten the workload of official phone numbers in emergency situations.

The Hispabot-Covid19 system has contributed to this aim by providing an easy access to official information, allowing citizens to formulate questions in their own words and providing a practical example of the great potential that these systems will have for communication between citizens and the Public Administration. The number of users' queries received by the system show the potential for the application of conversational systems in information and citizen services as consolidated in the Administration as the 012 telephone number (it received an average of 900,000 queries per year in the Community of Madrid between 2004 and 2007), the average of nearly 900,000 calls received monthly by the 060 service at present, or the recent application of these systems in the Social Security portal to provide a guide of steps to follow for the resolution of procedures and answer frequently asked questions.

The corpus acquired by means of the interaction of citizens with the Hispabot-Covid19 conversational system will be published in a open access repository in the next months with detailed instructions about its labeling (sentences IDs, entities, date and hour). We want to extend this information with the labeling of predicate-argument structures, the definition of different kinds of word embedding for the task, and the additional definition of train and test partitions.

5. Acknowledgements

6. References

[1] D. Pérez, D. Griol, and Z. Callejas, "El sistema Hispabot-Covid19: sistemas conversacionales para la Administración y la Atención al Ciudadano," *Boletic*, vol. 86, 2020. [Online]. Available: Available in: https://www.astic.es/la-asociacion/boletic/boletic-no-86-julio-2020f

[2] M. McTear, Z. Callejas, and D. Griol, *The Conversational Interface: Talking to Smart Devices*. Springer, 2016.

[3] M. McTear, *Conversational AI. Dialogue systems, Conversational Agents, and Chatbots*. Morgan and Claypool Publishers, 2020.

[4] J. Quesada, Z. Callejas, and D. Griol, *Informe sobre sistemas conversacionales multimodales y multilingües*. Plan de Impulso de las Tecnologías del Lenguaje. Available in: https://plantl.mineco.gob.es/tecnologias-lenguaje/actividades/estudios/Paginas/sistemas-conversacionales.aspx, 2019.