FOM - Hochschule für Oekonomie & Management

Hamburg

Master-Studiengang Big Data & Business Analytics

3. Semester

Development of a system to control and monitor blood pressure measurements to prevent cardiovascular disease

Betreuer: Prof. Dr. Kai Brüssau

Autor: Jacqueline Franßen

Matrikel-Nr: 496804

3. Fachsemester

Hamburg, den 28.02.2020

Contents

List of Figures

List of Tables

1 Abstract

Business Case: 1) With the developed solution, doctors are getting an overview of the patient's blood pressure values. This makes them react more precisely to any special values. 2) Since the patient is lead through a tutorial and the chatbot is 'controlling'/'checking' his measured blood pressure values, all measurements are taken more accurately. This improves the process of documentation. 3) With the service of sending a report via email every two weeks to the doctor. The doctor has always the current value and can interprete them faster. (now they are only getting a long list of all measured values of their patients which they have to interprete on their own and

4) Recommendation system of the nearest doctor helps the patient to directly go to his doctor Jack: 'Hey, i found these nearest doctors/specialists in your neighbourhood. Just select one of them and make an appointment with them.

2 Introduction

2.1 Problem statement

2.2 Aim and scope of this work

Aim of this scientific work is to develop a solution to document blood pressure in order to react preventively against heart disease. To recommend an appropriate doctor in one's surrounding (approximately 5 kilometers of distance). The application shall send every week/every two weeks a report (including a diagramm of all measured blood pressure values of the patient) to the doctor so that the doctor will be informed in real-time. In the diagrams/frontend, it is possible to select different scales, e.g. like the values of last week/last month/last year. At the beginning of using the Chatbot, the user is being led through a tutorial which shows him how to measure correctly his blood pressure. One instruction is for example not to drink coffee before measuring your blood pressure or to sit for at least 5 minutes.

3 Fundamentals

3.1 Reporting and Big Data

3.2 Software Architecture: Best Practices

To describe the 'best practices' of software architecture, in this section the architecture of a blockchain Peer to Peer (P2P) network which is backed with a distributed ledger system (see figure 3.1) will be explained. As stated by Talukder et al.[1], this model is an appropriate solution for health applications because they support multiple stakeholders.

As a common problem in every big data project there are multiple data sources and systems which provide relevant information for the particular use case. For instance, these information can be handwritten human readable and human understandable medical notes. Some information are computer readable and human understandable and the third 'generation' of information describes computer readable and understandable algorithms.

In order to provide an effective treatment of any disease, all health related data of a person on a spatial and temporal basis from birth is needed. These data will be examined by a panel of experts to reach a consensus Proof of Disease (POC) and include all illness episodes, lab tests, pathological test results (which are outside the normal range), genomic data (to evaluate the genomic state of the individual), environmental and health events, lifestyle related data captured by Internet of Things (IOT), therapeutic data and outcome analysis results.

According to Talukder et al.[2], there are three different types of mining:

[1] cf.[Talukder et al. 2018]
[2] cf.[Talukder et al. 2018]

- medical episode mining (MEM)
- health state mining (HSM)
- payment (financial/coin) mining

As can be seen in figure 3.1, medical systems need many resources from which all relevant medical data iare loaded. As described by Talukder et al.[3], all medical data is processed by Natural Language Processing (NLP) techniques, evidence based medicine as well as big data analytics (see figure 3.1). In most health applications, patients' participation increases when they have access to their health and lab records. In the solution provided by Talukder et al. genomic tests and non-communicable disease (NCD) data are stored in the blockchain as a transaction. Moreover, the blockchain technology is deployed in the cloud (cf. figure 3.1 Ethereum).

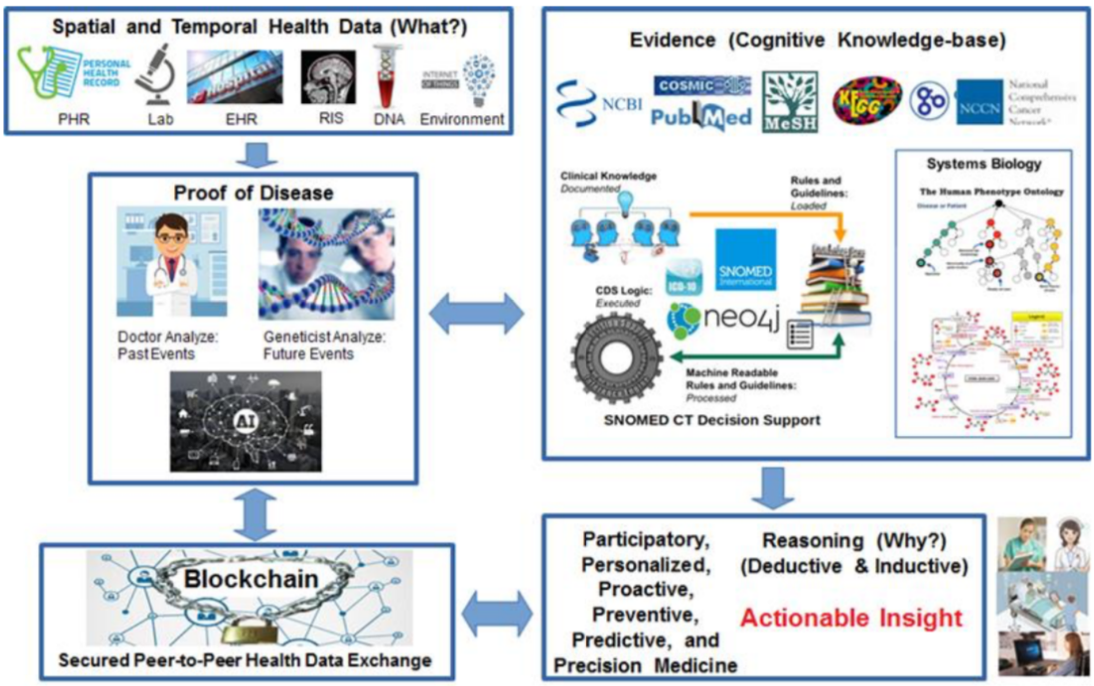

Figure 3.1: Example software architecture cf.[Talukder et al. 2018]

What is more, there is a medical miner which validates every transaction, then translates all clinical notes into structured International Classification of Disease (ICD) or Systemized NOmenclature of MEDicine Clinical Terms (SNOWMED CT) codes. After that, all codes are stored into a smart contract. During that process, a medical expert validates whether current the onset matches any clinical pathway. Finally, medical

[3]cf.[Talukder et al. 2018]

experts discuss in a proper medical consensus if the data is useful for an accurate diagnosis and public health.

3.2.1 SOA for big data applications in the cloud

The state-of-the art architecture for any project is SOA and has many advantages, such as flexibility, agility, process orientation, time-to-market and innovation[4]. What is more, SOA is convenient for cloud computing since it is ready for extended service models. Figure 3.2 shows the architecture 'ESARC', developed by Zimmermann et al.[5]. It helps to cluster, classify, examine, compare, evaluate quality and optimize enterprise architectures. As depicted by figure 3.2, there is a link between enterprise and business information and design for supporting strategic initiatives. What is more, ESARC enables integration capacities for **IT!** (**IT!**) management, software engineering, service and operations management as well as process improvement initiatives.

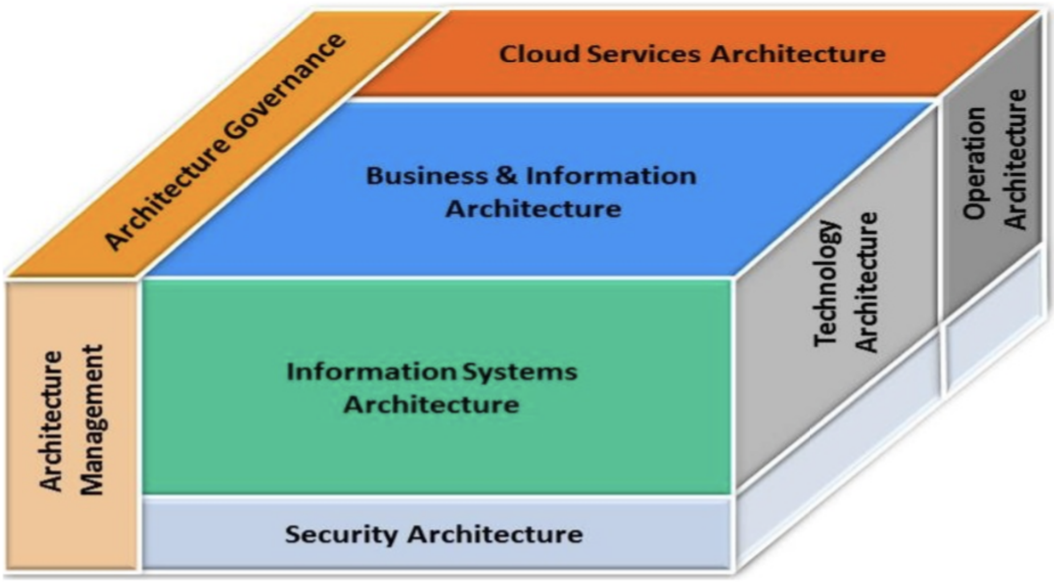

Figure 3.2: ESARC as an example for big data architecture cf.[Zimmermann et al. 2013]

As can be seen in figure 3.2, metamodels are used to define model elements in architectures. They relate architectural elements to ontologies which represent a common vocabulary for enterprise architectures. Zimmermann et al. recommend that operations of tasks and entity services should not have any knowledge about their process

[4]cf.[Zimmermann et al. 2013]
[5]cf.[Zimmermann et al. 2013]

or interactive usage context [6]. Instead, task service operations should be independent from users and sessions and should only implement business functionality.

Figure 3.3: ESARC business and information reference architecture cf.[Zimmermann et al. 2013]

Figure 3.4 shows a more detailed view of 'ESARC', the procedural framework for architecture assessment processes and questionnaire design. On top of the graphic, with orange background, there are business vision, drivers, goals and objectives. To be more precuse, architecture governance has the goal to manage activities such as plan, define, enable, measure, control and sets rules for architecture complicance to internal and external standards.

Actors in cloud computing According to Zimmermann et al., the main actors in cloud computing are cloud consumers, providers, auditors and broker [7]. In general, all SOA services are cloud services and follow a reference architecture: 'Jericho-Security-focused Service-oriented Reference architecture for cloud computing'. Thereby, management perspectives from Information Technology Infrastructure Library (ITIL) and The Open Group Architecture Framework (TOGAF) standards are integrated.

[6] cf.[Zimmermann et al. 2013]
[7] cf.[Zimmermann et al. 2013]

3.3 Medical Documentation Apps

As stated by Lupton et al.[8] the current technical 'era' we are living in is the web 2.0 or social web. Social web includes sharing health and medical information with each other, e.g. patients and caregivers write about experiences and the individual health status. Often, the aim of these social webs is to control the health status by using online information and imaging. Conforming to Lupton et al.[9], in healthcare projects, big data can be used to generate knowledge about healthcare, health behaviours and disease patterns. Such applications can assist in calculating diagnosis, identifying risks, facilitating health, fitness self-tracking as well as patient self-care regimes. As reported by a study which surveyed American doctors[10] medication interaction apps are the first most-used and diagnosis apps the second most used category of apps. Moreover, pregnancy apps offer greater opportunities, such as that women can engage obsessive self-surveillance because of producing detailed data, such as heart rates, in real-time [11]. Pursuant to Lupton et al., the future potential of medical application lies in systems which enable lay people to access medical information (such as the electronic medical record) that was previously only available to healthcare practitioners or students.

In another article, Lupton et al. reported that the potential lies in automation of news or notifications which can be personalized or targeted so that doctors could contact patients directly to remind them to adhere to their tratment programs [12]. A further example of medical applications are 'smart pillboxes' for patients suffering from diabetes[13]. 'Smart pillboxes' are wireless devices that remind patients to take their medication and alert a doctor if the patient had failed to conform to their medication regimen. Continuing, medical health (M-HEALTH) technologies have a feedback, also called cybernetic mechanism in that they react with their users as opposed to passively provide information. To give an example, modern prosthesis or technological extensions of the body are a kind of cybernetic mechanism[14]. A big part of

[8] cf.[Lupton 2014]
[9] cf.[Lupton 2014]
[10] cf.[Lupton 2014]
[11] cf.[Lupton 2014]
[12] cf.[Lupton 2012]
[13] cf.[Lupton 2012]
[14] cf.[Lupton 2012]

today's medical applications are surveillance systems in order to record and monitor cases of illnesses, such as obesity or infections[15]. These records might be useful to early detect epidemiological changes in the disease pattern. To give an example, 'individual medical encounters' which are conducted online enable doctors to flexibly practice personalized surveillance over each of their patients. At this point, another term occurs: 'surveillance knowledge' which refers to the digital data produced in the surveillance and can be useful for the individuate users.

Blockchain solution for accurate medical decisions As stated by Talukder et al. a significant amount of today's diagnosis in NCD is erroneos or unwanted[16]. The term NCD implicates disease caused by an unheallthy lifestyle, the proper environment or genomic causes over a long time and come up with confusing signs and symptoms. Talukder et al. mention 'P6'-Medicine which describes medicine using six adjectives starting with the letter 'p': medicine needs to be participatory, personalized, proactive, preventive, predictive, precision medicine. As a requirement list for health data, Talukder et al. describe important features as follows:

- secured (the anonymity, privacy, confidentiality of health data must be approved)
- systems which provide health data must have a zero down-time
- the integrity of the health data must be ensured
- the systems must be ubiquitous which implies an unlimited availability
- machine understandable (all health data should be conform to international standards and should be able to be distributed over multiple systems)
- health systems should be resistant against fraudulent hacking

[17]

[18] [19]

[20] https://www.heart.org/en

Blood Pressure Apps: Was können die schon ?

[15]cf.[Lupton 2012]
[16]cf.[Talukder et al. 2018]
[17]cf.[Kawohl and Haß 2019]
[18]cf.[White 2007]
[19]cf.[Beevers, Lip, and O'Brien 2001]
[20]cf.[Alessa et al. 2019]

3.4 Chatbots

3.4.1 The potential of chatbots

Modelling, profiling, analyzing and understanding users becomes increasingly important in many different indrustries and count as key to success in todays data driven world. The main advantage of chatbots is to provide a 24-hour customer service with personalized interaction and no waiting time[21]. Akhtar et al. analyzed chat conversations between customers and the chatbot of a telecommunication company in order to find out the user's topics of interest and how to satisfy users. As described by Akhtar et al., the tests of the chatbot were splitted into different activities, such as text mining techniques (feedback comments), event sequence analysis, frequent term extraction, analysis of bigrams/trigrams. During data preprocessing, Akthar et al. used the following methods:

1. corpus generation
2. eliminating extra white space
3. stopwords removal
4. tokenizing
5. stemming
6. creating term-document matrices

The main challenges during the data analysis process are data availability, the access to further user information (e.g. contract details or age in order to generate an user model) and the distinction between different user types and different personality structures.

Question Answering Paradigms There are several types of conversations which can be designed by building a chatbot[22]. Generally, there can be distinguished between two different paradigms: information-retrieval based Question and Answering and knowledge-based Question and Answering. The first type describes the mechanism to define short texts as answers to a user's intent. On the opposite, the second

[21] cf.[Akhtar, Neidhardt, and Werthner 2019]
[22] cf.[Akhtar, Neidhardt, and Werthner 2019]

type describes how to in natural language. The answers are stored into a full-related database and the conversation works simply with a rule-based method.

Types of dialog systems In general, dialog systems can be divided into two kinds of systems. On the one hand, there are task-oriented systems which are appropriate for short conversations and built for a certain purpose. On the other hand, there are non-task-oriented systems which are built for longer and more complex interactions with the purpose of imitating human conversations[23].

3.4.2 A deep learning question-answering specialized chatbot for medical students

During their studies, medical students have to take an exam which is called Objective Structured Clinical Examination (OSCE) where they interact with a 'standardized' patient played by an actor who simulates the symptoms and intents of the patient[24]. The aim of this exam is to test and assess the students' abilities and social interaction and diagnosis skills. Since in practice, there are not many actors who can play a patient's role, Zini et al. developed a virtual patient and chatbot system which works with NLP techniques.

Figure 3.4 shows the architecture of the developed system to create a virtual patient. Zini et al. used a Convolutional Neural Network (CNN) and Long Short Term Memory (LSTM) network in order to learn domain specific word embeddings, sentence embedding and answer selection models. The embeddings model which is outlined by a red rectangle in figure 3.4 is trained on a corpus of medical documents. In Figure 3.4, there is a NLP engine outlined by a red rectangle. By using a supervised learning scheme to learn a mapping between question and answering pairs and judgement of correct match, this NLP engine should correctly answer questions based on a script.

The aim of the developed system was to create a deep learning framework for answer selection in the medical domain and to create domain-specific word and sentence e,bedding models. Additionally, a question and answering corpus should be created for OSCEs.

[23] cf.[Akhtar, Neidhardt, and Werthner 2019]

[24] cf.[Zini et al. 2019]

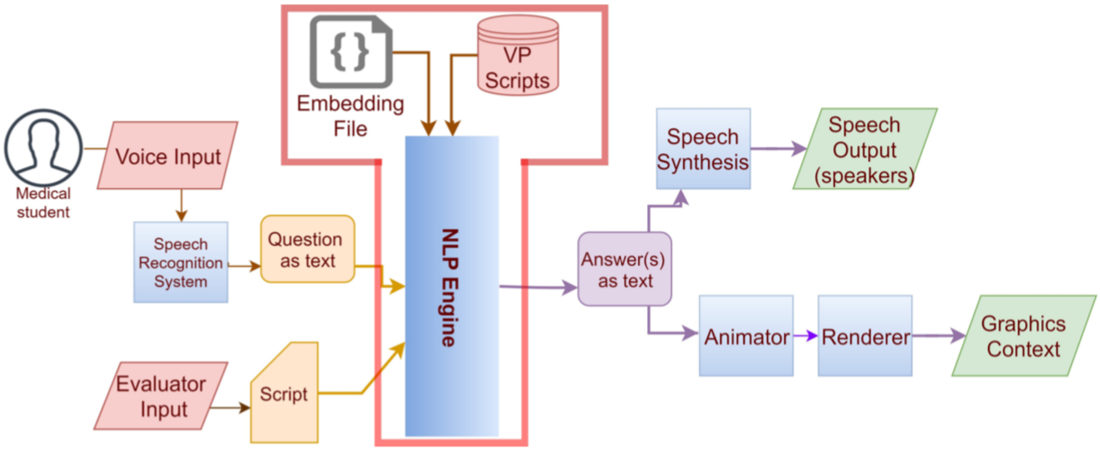

Figure 3.4: Virtual patient software architecture cf.[Zini et al. 2019]

Question and Answering systems According to Zini et al. there are two types of question and answering systems[25]. First of all, there is open domain question and answering which uses very specific terminology. Secondly there are restricted domain question and answering systems which are broader in their scope. These are for example insurance-related deep learning question and answering systems which make use of two baseline models: Bag Of Words (BOW) and Information Retrieval (IR) model.

[25]cf.[Zini et al. 2019]

4 Analysis and Development

First of all, the developed chatbot includes information about blood pressure and was built to remind patients of their measurements. Secondly, based on the measured data, analysis can be done in order to react earlier to outliers. Thirdly, a generated report is sent to the doctor so that he can get more insights about the blood pressure values of his patients and improve the treatment. One of the main challenges during development was the process of providing information about the disease to the patient. Is it possible to include information about different types of blood pressure into the automated conversation with a chatbot? Or does it overwhelm the conversation's use case? Is it useful to let the user ask questions like: 'What are the different types of hypertension? Am i a high-risk patient?' Or should these information be provided as a video or a simple web page with long articles to read? Might the patient or user be aborred after a while of talking to a chatbot who only knows answering his questions in the same way? Of course, a chatbot can be developed more intelligent to never provide the same answer and to answer more precisely to a users' intent. But this requires a lot of training and testing. For that reason, in the first version of this chatbot, five simple intents and dialogs have been designed and implemented with the focus of the instructions to measure correctly and regularly. In a second or third version, it is possible to focus more on the improvement of providing information about the disease (by not doing this in the style of question and answering).

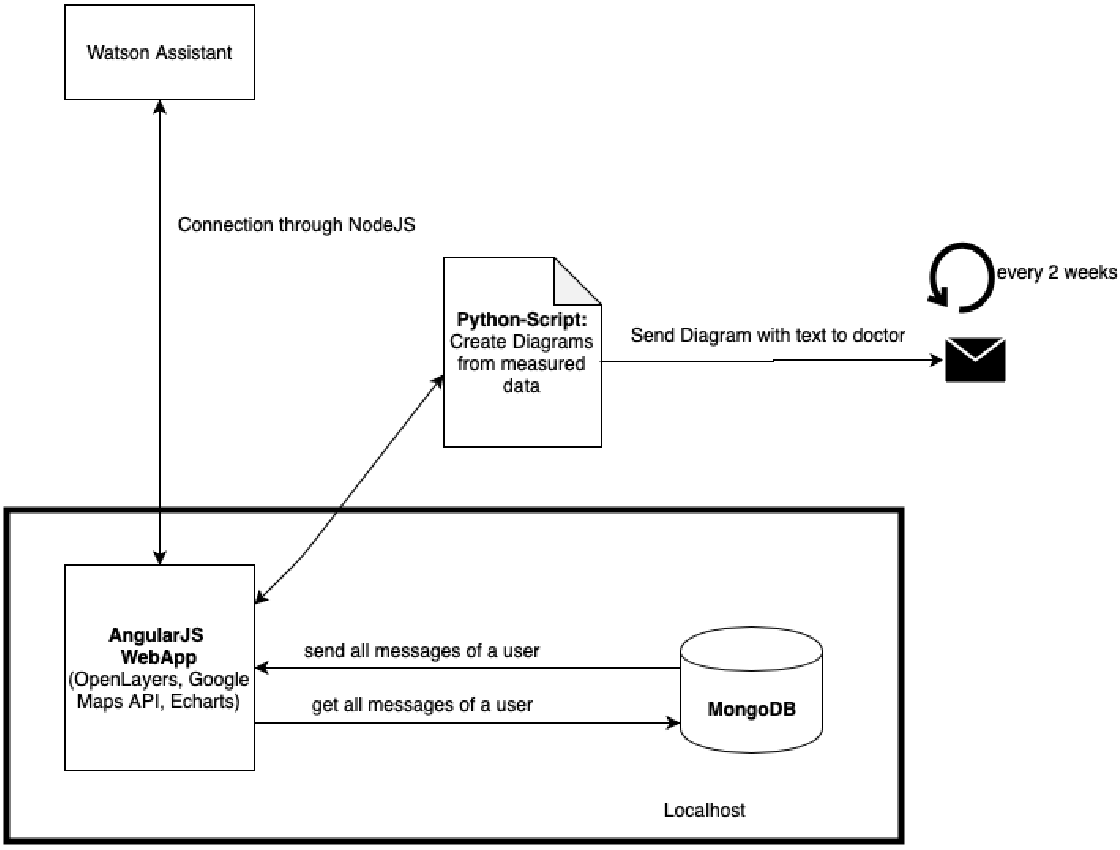

Figure 4.1: Component diagram of developed solution

4.1 Experimental set-up

4.1.1 Software architecture

4.1.2 Components

Kaggle Dataset

Development of Watson Assistant Dialog

Intent model The chatbot was built according to the description of "Deutsche Hochdruckliga", a german organization for patients with hypertonia [1]. To better understand the users and patients a basic intent model with four intents was developed. The four intents include

[1] cf.[*Bluthochdruck vermeiden, behandeln und senken - Aktiv gegen Bluthochdruck* 2019]

Intent	User input
Definition of hypertonia	What is hypertonia?
Curses of hypertonia	What are the implications or curses of hypertonia?
Blood pressure measurement	I am measuring my blood pressure.
Measurement tutorial	How should i measure my blood pressure?

These five intents were used to define and develop four typical dialogs, displayed in figure 4.2, 4.3, 4.4 and 4.5.

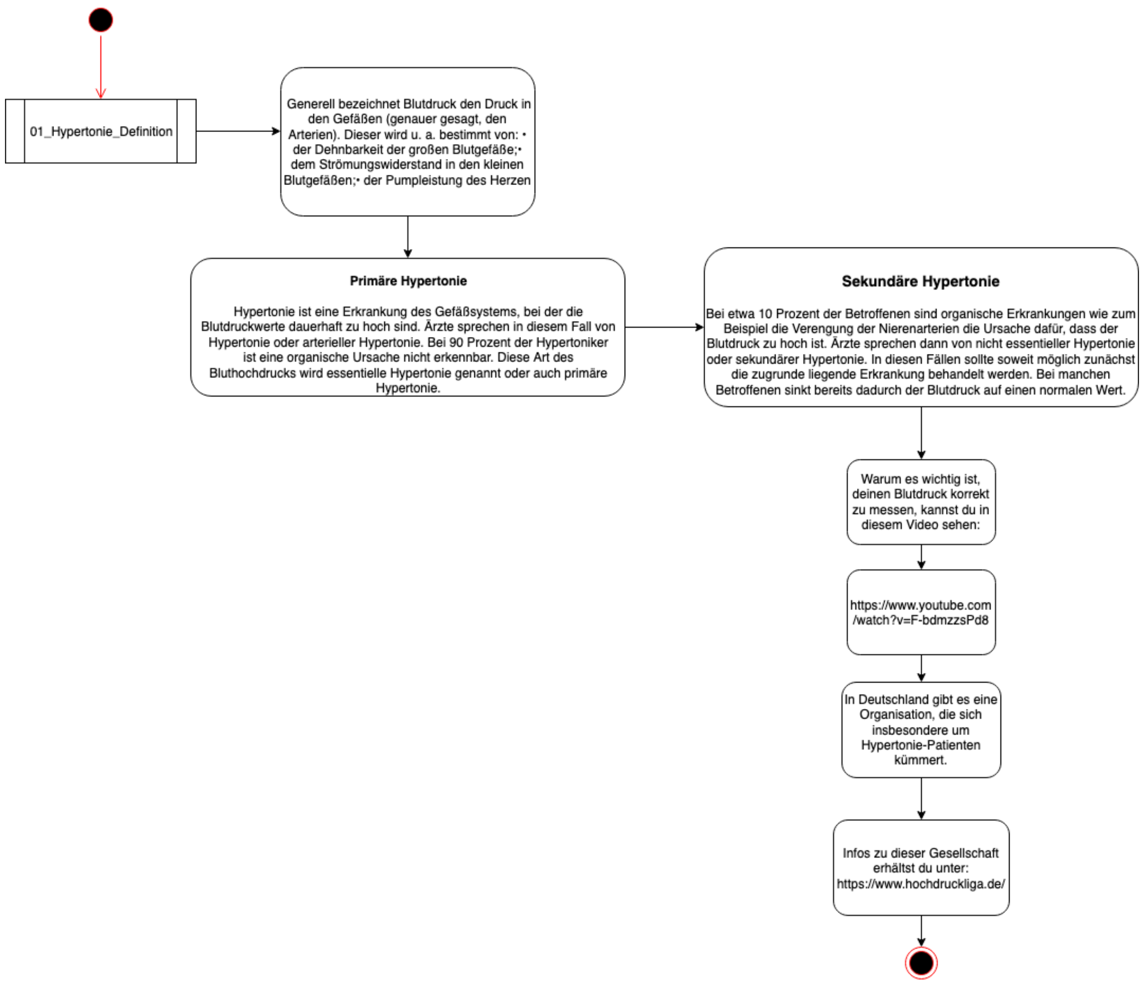

Figure 4.2: Dialog diagram: definition of hypertonia

Watson Assistant implementation In the following, the implemented dialog as well as all entities and intents are described. They have been developed according to the Watson Assistant documentation [2].

[2]cf.[*Übersicht über die Watson Assistant-API* 2020]

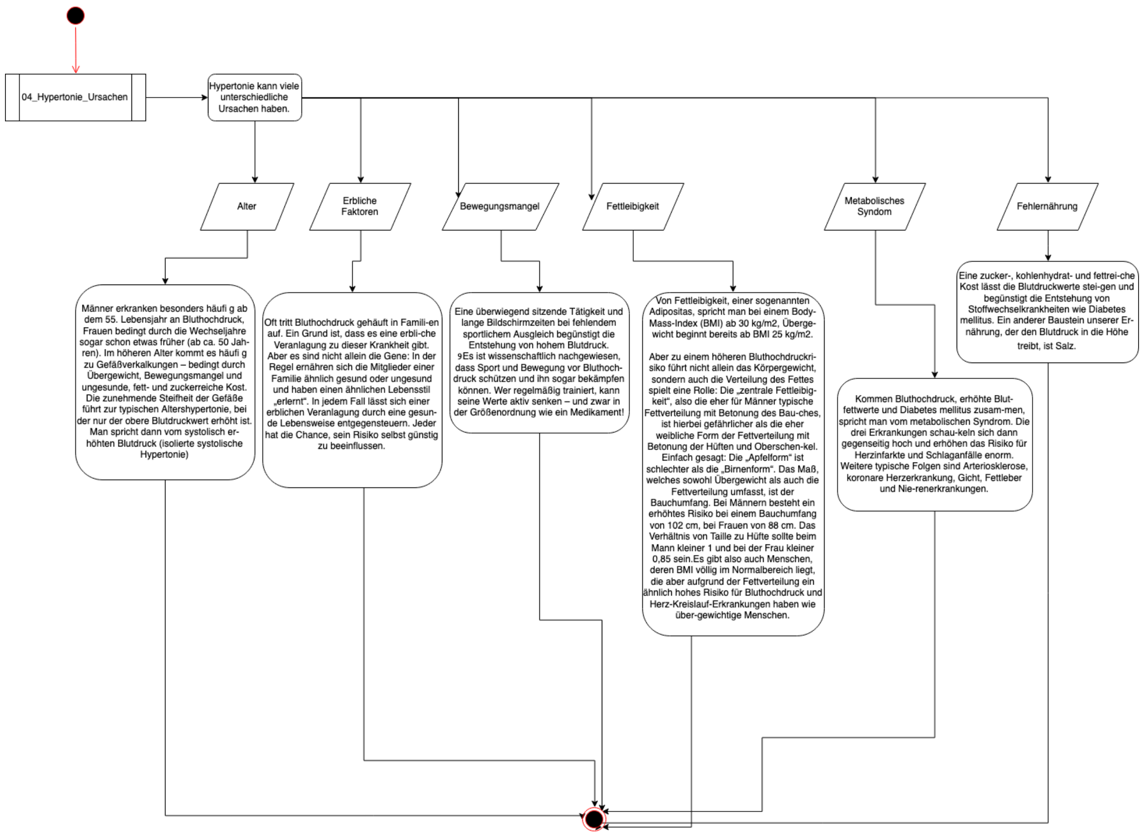

Figure 4.3: Dialog diagram: curses of hypertonia

Setup of MongoDB and basic AngularJS Frontend

Connect Watson Assistant to Frontend and MongoDB To be able to connect to the Watson Assistant instance on IBM cloud, the API Version 2.0 had to be called. First of all, the current sessionId has to be requested to be able to interact with Watson Assistant. After that, a simple get request is executed to let the chatbot start the conversation. Everytime, the user sends a message to Watson Assistant, a post request is sent to the API and the response is the best fitting answer (with the highest confidence value) from Watson, which is calculated [3].

Data visualization: Development of a Python Script to show all measured values

Development of recommendation of nearest doctors to patient outlook: maybe in future to connect to the doctors' calendar to directly make an appointment through the chatbot

[3]cf.[*Watson Assistant v2 - IBM Cloud API Docs* 2020]

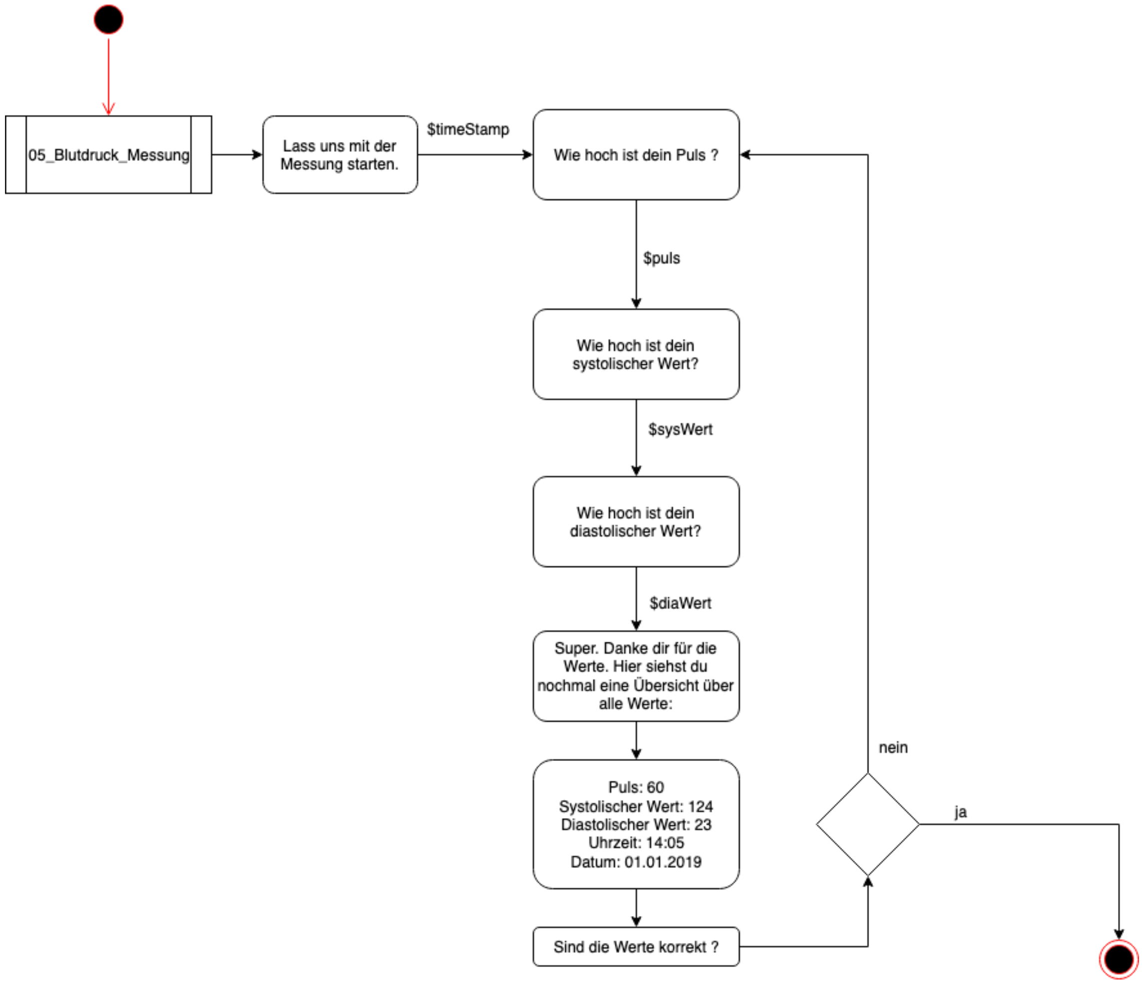

Figure 4.4: Dialog diagram: blood pressure measurement

Development of email service to send reports every two weeks to the doctor

Another challenge was the way to automatically ask the user for his measured data. A 'usual' chatbot only helps in certain situations including precise user intents, e.g. the question 'When should i measure my blood pressure?'. But they are not constructed to ping a user every five hours or once a day in order to retrieve his measured data, analyze these and send them to a doctor. In order to face this problem or use case, a routine including a timer had to be implemented.

4.2 Problem solving

4.2.1 Tests

4.2.2 Dataset

4.3 Results

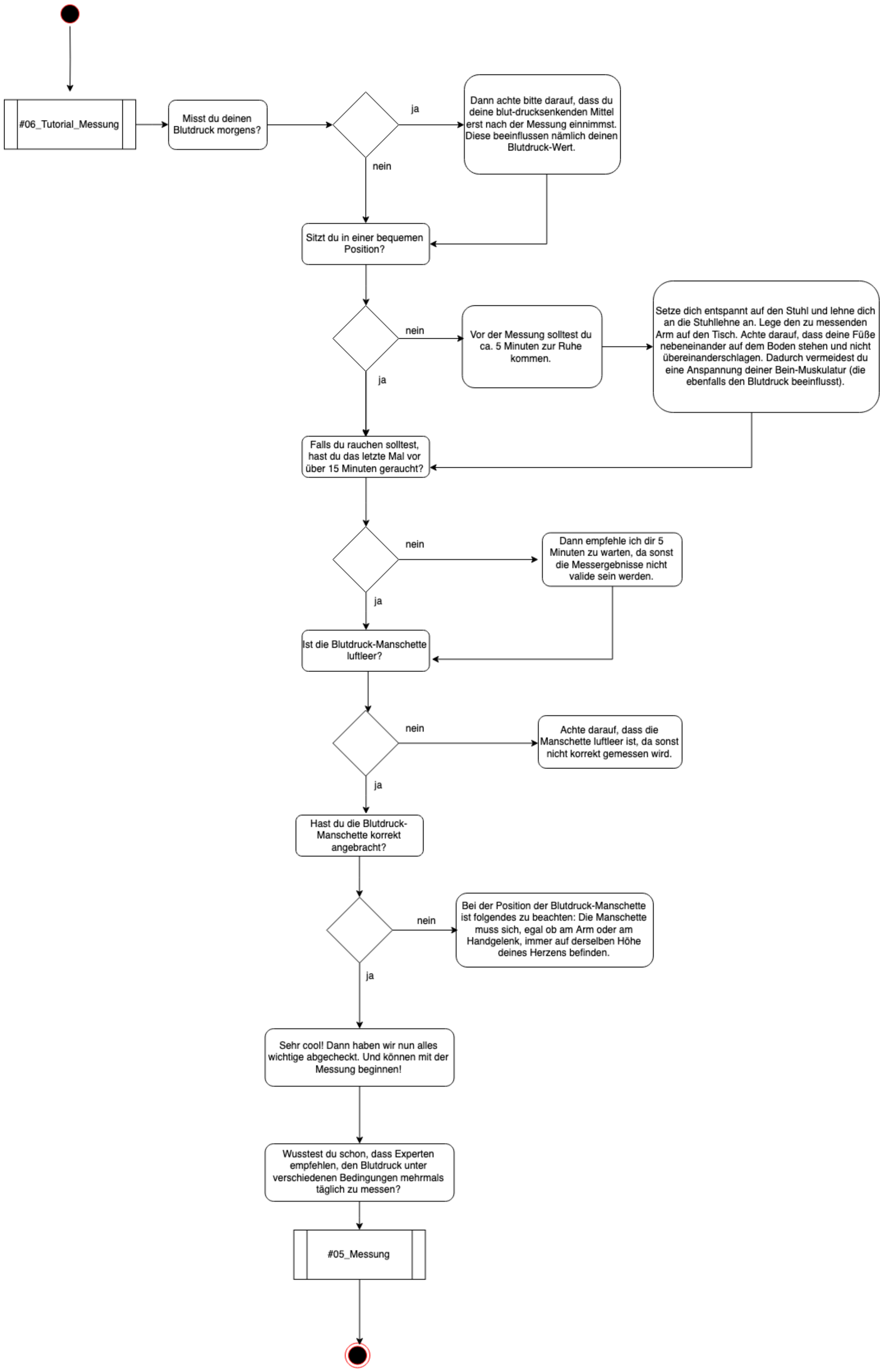

Figure 4.5: Dialog diagram: measurement tutorial

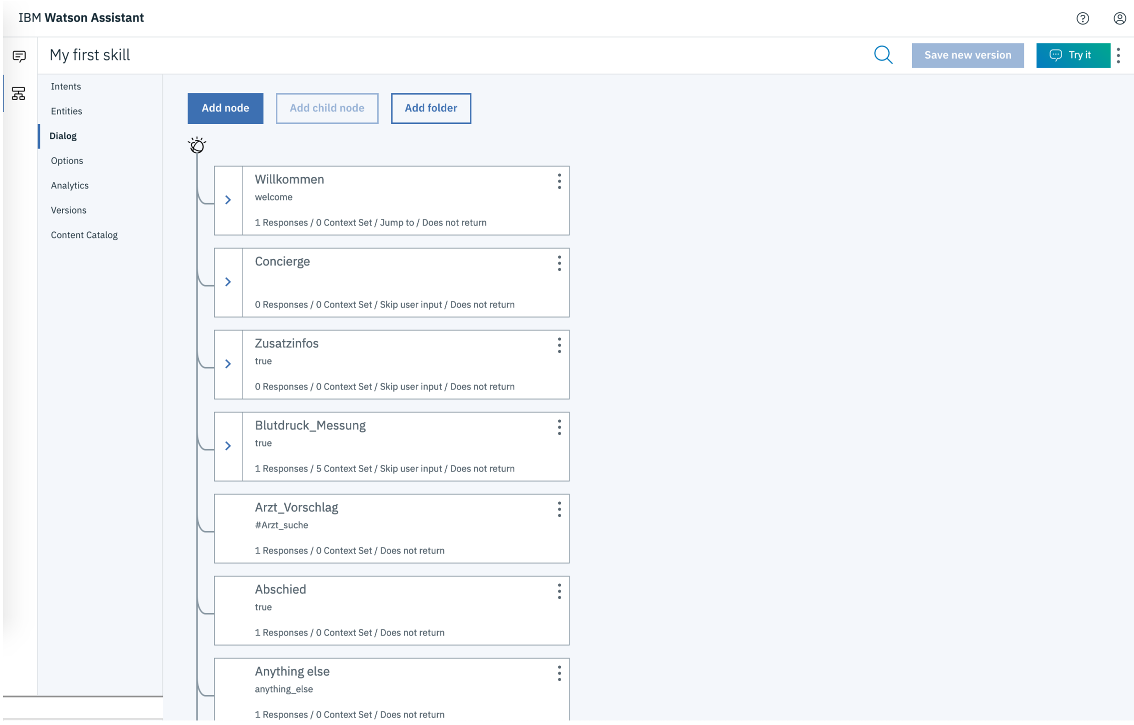

Figure 4.6: Watson Assistant dialog

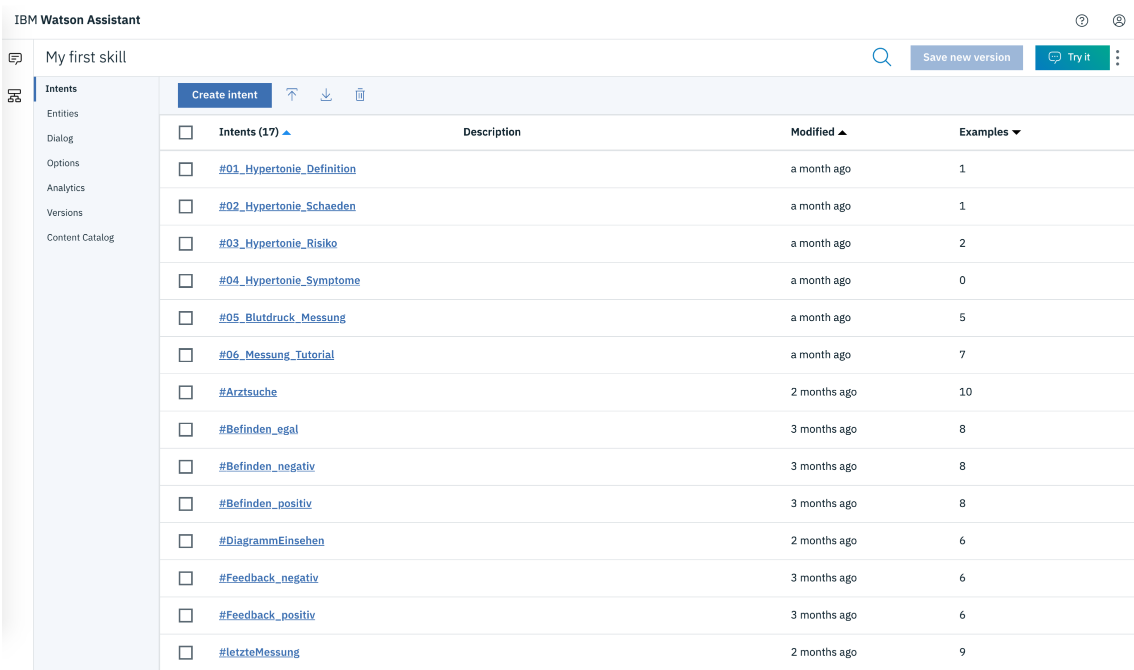

Figure 4.7: Watson Assistant intents

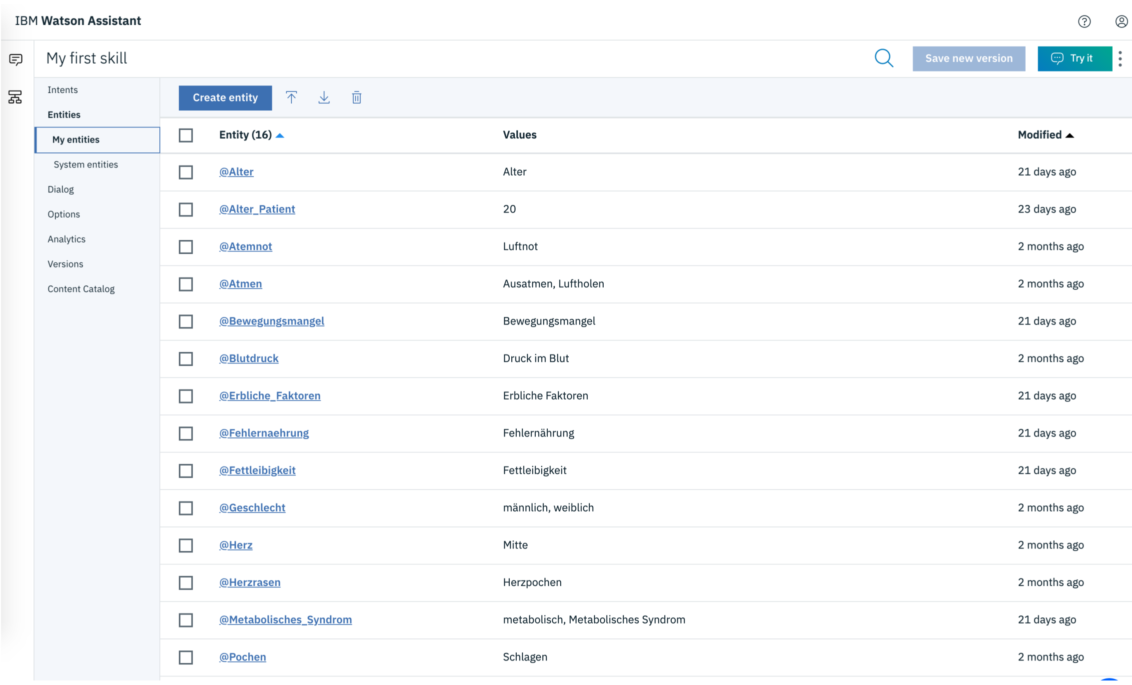

Figure 4.8: Watson Assistant entitities

5 Conclusion and Outlook

5.1 Conclusion

5.2 Outlook

Is a chatbot an appropriate solution for recording and reminding patients to measure their blood pressure ? In practice, most chatbots are created to solve and help multiple intents of their users and not to only 'retrieve' information. On the one hand, the retrieved information can be used to run several analysis and to find out trends in the data. But on the other hand, the developed solution supports users to never forget to measure and doctors to better understand trends in the measured data. During development and research, there came up another issue: To inform patients precisely about their illness. Most patients go to doctors and describe their symptoms, the doctor makes some diagnosis and provides them some medication. But in most cases, doctors do not have enough time to answer all the questions of their patients. For that reason it might be useful to provide a 24/7 service for patients with chronic disease to both record their symptoms and measurements and to answer all their questions.
connection frontend (angularjs webpage) to mongodb through socket.io connection.
Connection between Watson Assistant and Frontend through mongodb and via socket.io connection.

6 Abbreviations

NLP	Natural Language Processing
ESARC	Enterprise Software Architecture Reference Cube
M-HEALTH	medical health
P2P	Peer to Peer
NCD	non-communicable disease
POC	Proof of Disease
IOT	Internet of Things
MEM	medical episode mining
HSM	health state mining
ICD	International Classification of Disease
SNOWMED CT	Systemized NOmenclature of MEDicine Clinical Terms
OSCE	Objective Structured Clinical Examination
CNN	Convolutional Neural Network
LSTM	Long Short Term Memory
BOW	Bag Of Words
IR	Information Retrieval
SOA	Service-Oriented enterprise Architecture
ITIL	Information Technology Infrastructure Library
TOGAF	The Open Group Architecture Framework

Bibliography

Akhtar, Mubashra, Julia Neidhardt, and Hannes Werthner (July 2019). "The Potential of Chatbots: Analysis of Chatbot Conversations". In: *2019 IEEE 21st Conference on Business Informatics (CBI)*. Vol. 01. ISSN: 2378-1963, pp. 397–404. DOI: 10.1109/CBI.2019.00052.

Alessa, Tourkiah et al. (May 2019). "Smartphone Apps to Support Self-Management of Hypertension: Review and Content Analysis". In: *JMIR mHealth and uHealth* 7.5. ISSN: 2291-5222. DOI: 10.2196/13645. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6658295/ (visited on 01/05/2020).

Beevers, Gareth, Gregory Y H Lip, and Eoin O'Brien (Apr. 2001). "Blood pressure measurement". In: *BMJ : British Medical Journal* 322.7293, pp. 1043–1047. ISSN: 0959-8138. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1120188/ (visited on 01/03/2020).

Bluthochdruck vermeiden, behandeln und senken - Aktiv gegen Bluthochdruck (2019). de. URL: https://www.hochdruckliga.de/ (visited on 11/23/2019).

Kawohl, Dr Julian M and Stefanie Haß (2019). "Customer Service 4.0 - Wie gut sind Chatbots?" de. In: (), p. 30. URL: https://www.heise.de/downloads/18/2/5/4/1/3/4/2/Studie_chatbots.pdf (visited on 01/01/2019).

Lupton, Deborah (Aug. 2012). "M-health and health promotion: The digital cyborg and surveillance society". en. In: *Social Theory & Health* 10.3, pp. 229–244. ISSN: 1477-8211, 1477-822X. DOI: 10.1057/sth.2012.6. URL: http://link.springer.com/10.1057/sth.2012.6 (visited on 12/27/2019).

– (Dec. 2014). "Apps as Artefacts: Towards a Critical Perspective on Mobile Health and Medical Apps". en. In: *Societies* 4.4, pp. 606–622. DOI: 10.3390/soc4040606. URL: https://www.mdpi.com/2075-4698/4/4/606 (visited on 12/27/2019).

Talukder, Asoke K et al. (Oct. 2018). "Proof of Disease: A Blockchain Consensus Protocol for Accurate Medical Decisions and Reducing the Disease Burden". In: *2018*

IEEE SmartWorld, Ubiquitous Intelligence Computing, Advanced Trusted Computing, Scalable Computing Communications, Cloud Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI). ISSN: null, pp. 257–262. DOI: `10.1109/SmartWorld.2018.00079`.

Übersicht über die Watson Assistant-API (2020). URL: `https://cloud.ibm.com/docs/services/assistant?topic=assistant-api-overview` (visited on 01/10/2020).

Watson Assistant v2 - IBM Cloud API Docs (2020). URL: `https://cloud.ibm.com/apidocs/assistant/assistant-v2` (visited on 01/10/2020).

White, William B (2007). *Blood pressure monitoring in cardiovsacular medicine and therapeutics*. en. OCLC: 915957468. Totowa, N.J.: Humana Press. ISBN: 978-1-59259-978-3 978-1-58829-512-5. URL: `http://www.springerlink.com/openurl.asp?genre=book&isbn=978-1-58829-512-5` (visited on 01/03/2020).

Zimmermann, Alfred et al. (Sept. 2013). “Towards Service-Oriented Enterprise Architectures for Big Data Applications in the Cloud”. In: *2013 17th IEEE International Enterprise Distributed Object Computing Conference Workshops*. ISSN: 2325-6605, pp. 130–135. DOI: `10.1109/EDOCW.2013.21`.

Zini, Julia El et al. (July 2019). “Towards A Deep Learning Question-Answering Specialized Chatbot for Objective Structured Clinical Examinations”. In: *2019 International Joint Conference on Neural Networks (IJCNN)*. ISSN: 2161-4393, pp. 1–9. DOI: `10.1109/IJCNN.2019.8851729`.

7 Appendix A

Ehrenwörtliche Erklärung

Hiermit versichere ich, dass die vorliegende Arbeit von mir selbstständig und ohne unerlaubte Hilfe angefertigt worden ist, insbesondere dass ich alle Stellen, die wörtlich oder annähernd wörtlich aus Veröffentlichungen entnommen sind, durch Zitate als solche gekennzeichnet habe. Ich versichere auch, dass die von mir eingereichte schriftliche Version mit der digitalen Version übereinstimmt. Weiterhin erkläre ich, dass die Arbeit in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde / Prüfungsstelle vorgelegen hat. Ich erkläre mich damit nicht einverstanden, dass die Arbeit der Öffentlichkeit zugänglich gemacht wird. Ich erkläre mich damit einverstanden, dass die Digitalversion dieser Arbeit zwecks Plagiatsprüfung auf die Server externer Anbieter hochgeladen werden darf. Die Plagiatsprüfung stellt keine Zurverfügungstellung für die Öffentlichkeit dar.

Ort, Datum (Vorname Nachname)

8 Appendix B