POLITECNICO
MILANO 1863

Emotional Storytelling

Design & technology

Advanced User Interface

2020 - 2021

5th February 2021

Abstract

Children with cognitive disorders experiment the world differently that neurotypical people, so they may not be able to understand and express emotions in a similar way, which can have an impact on their social interactions. Using the tools of storytelling, we try to help them improve their emotional intelligence and better connect with people. Throug research and creative problem solving, we have designed and developed a conversational storyteller, which helps children to identfy and manage emotions in story settings, that they can later apply in daily situations.

Team 13

Aida Denisa Opîrlesc
ID : 10735889
aidadenisa.opirlesc@mail.polimi.it
Computer Science and Engineering

Akash Aloysius James
ID : 10687690
akashaloysius.james@mail.polimi.it
Computer Science and Engineering

Gerardo Frutos Dordelly
ID : 10692616
gerardo.frutos@mail.polimi.it
Design & Engineering

Mélodie Marie Anne Jacob
ID : 10621817
melodie.jacob@mail.polimi.it
Design & Engineering

Table of contents

p. 5 Introduction

p. 6-8 1. Target Group & User Needs

p. 6 1.1. Target, User & Context

p. 6 1.2. Goal & Objectives

p. 6 1.3. Constraints

p. 7 1.4. Requirements

p. 9 - 11 2. State of the Art

p. 9 2.1. Research Papers

p. 10 2.2. Research Project

p. 11 2.3. Project on the Market

p. 12 - 22 3. Solution - UX Design

p. 12 3.1. Exploration through Digital Ethnography

p. 13 3.2. Persona & Scenario

p. 13 3.3. Structure of Hyperstory

p. 19 3.4. User Journey

p. 21 3.5. Design Choice / Interface

p. 23 - 28 4. Implementation

p. 23 4.1. High Level Software Architecture

p. 24 4.2. Client Application (Front-End)

p. 25 4.3. Server Application (Back-End)

p. 27 4.4. Database Schema

p. 28 4.5. External Services

p. 28 4.6. Technologies

p. 29 - 30 5. Value Proposition

p. 31 - 32 6. Future Work

p. 33 7. Bibliography

p. 34 - 37 8. Annexes

Introduction

Children with cognitive disabilities may have difficulties to express, understand and manage their emotions. As they may not be able to identify their emotions or empathize with other persons, it is difficult for them to adopt an appropriate reaction to a specific situation. Their perception of the world may not help them to interpret properly the sensory information they receive, which can prevent them from understanding the emotions of their interlocutor.

Storytelling can positively impact emotional intelligence of children by staging emotion through fictive stories. By seeing characters experimenting the same feelings, the child can make comparison between the fictive situation and his own life. Storytelling can help him to externalise his emotion and be more sensitive to others' feelings.

Thus, we are developing a storyteller application, using conversational agents, to help the child (aged between 6-12 years old) with cognitive disabilities to learn more about emotions and how to manage them, in order to better interact within society. This is achieved through short interactive stories and conversations with the chatbot, that last between 2-4 minutes long, which are helping the user train and learn through storytelling, association and repetition.

We have decided to build it as a mobile-friendly web application in order to make it accessible and portable, so that it can be used on both desktop and mobile devices. With the help of sounds and a narrator's voice, as well as interactive and playful story with images, we aim to keep the child focused and present the story.

Through the use of repetition and alternate-ending stories we aim to engage and help the child remember relatable scenarios and know how to react in certain situations. In order to offer a personalized experience to children, the story is adapted real-time according to the emotions detected in their answers. Although this application is not intended to replace therapy, it can enhance practicing at home, intended for the child to have fun and learn at the same time.

1. TARGET GROUP & USER NEEDS

1.1 Target group, user needs & context

The target group of this project is composed by children with cognitive disabilities, aged from 6 to 12 years old. The potential users may have difficulties in understanding their emotion and other people's emotion, as well as adopt an appropriate response to a situation. In general, children with special needs also have problems to interact with other members of the society and to maintain a visual contact with their interlocutor. On the practical level, they may be unable to read or write.

Children with cognitive disabilities may suffer from different pathologies, such as autism, Down Syndrome, Traumatic Brain Injury (TBI), and dementia. Less severe cognitive conditions include the Attention Deficit Disorder (ADD), dyslexia (difficulty reading), dyscalculia (difficulty with math), and learning disabilities in general. The user presents specific needs that we must take in consideration throughout the creative and technical process. Children with cognitive disabilities need help to recognize, understand and manage their emotion, which require a lot of effort for them. They must practice listening and speaking in order to better interact with society. Moreover, they need to learn to connect with individuals by empathizing with their circumstances.

In order to improve their emotional education, we have created an application which can be used at home. No specific framework is required for the use this conversational agent, but the parents can supervise the daily training.

1.2 Goals & objectives

The main goal of the application is to help children with cognitive disabilities understand and manage their emotions, in order to better interact within society, through the use of interactive stories with diverse activities and multiple endings.

In addition to the main goal, we also have several objectives. The chatbot needs to provide a **personalized learning**, by taking in consideration the emotions identified in the child's messages. Then, the main tool for **providing emotional education** support we want to use is Storytelling, which has been proven to improve their learning. Moreover, we want to **improve the child's engagement and create an immersive experience** by using fictive scenarios and interactive stories. Last but not least, we want to provide a tool which can help the child **practice their emotional understanding in between therapy sessions**, in the comfort of their home.

1.3 Constraints

As listed in the diagram below, among the constraints of this project we can find the technology used, which is a conversational agent which is capable of emotion recognition from text phrases. This project has a team of two engineers and two designers, which have been given two months for implementing it. Because it deals with children with cognitive disabilities, this application has to make sure the user's data is protected.

1.4 Requirements

Based on the goal and objectives of the project, we created the list of requirement :

FR1. The system should be able to engage in fluent conversations, directed towards the user's experience

FR2. The system should be able to narrate (audio) the messages that it communicates to the user

FR3. The system should perceive what the user is saying and act according to it.

FR4. The system should display the dialog between itself and the user.

FR5. The system should be able to detect the emotion expressed in the user's messages.

FR6. The system should be able to choose an appropriate story based on the detected emotions.

FR7. The system should be able to narrate the story, pause it, ask questions and wait for the user's response to them, before continuing the story.

FR8. The stories should be interactive, include questions about the character's emotions or actions.

FR9. The system should provide an explanation for why the user or the character in the story feels a specific emotion.

FR10. The system should offer suggestions about how a detected emotion can be managed.

FR11. The system must provide the child with the continuation of the story based on his previous answers.

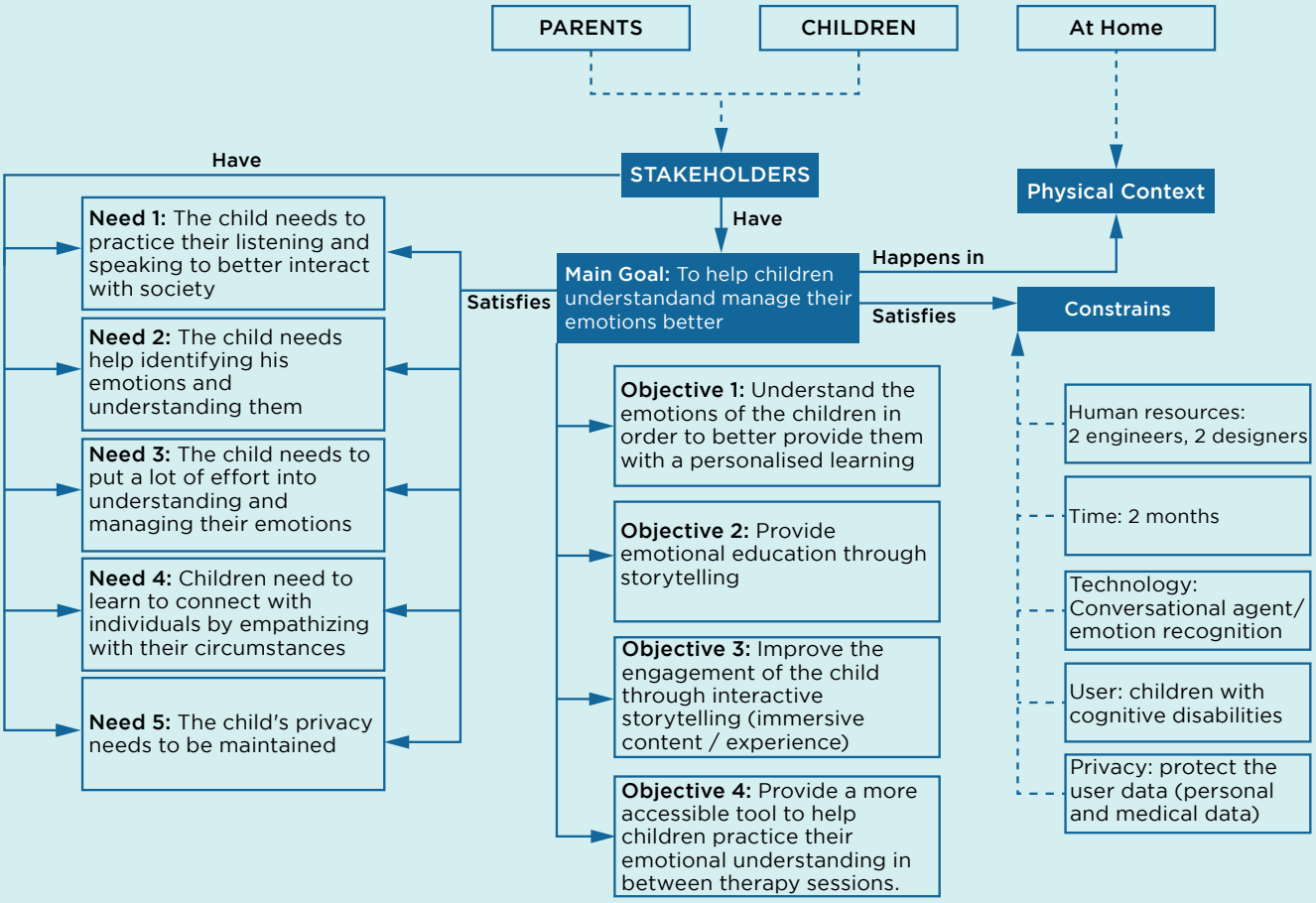
FR12. The system must have a predefined set of stories that are composed by sections that have description, questions, possible answers (suggestions), and the next steps (sections) that can be reached, depending on the emotion and the answer of the child.

The non-functional requirements we have identified are the following:

NFR1. The system must store the user's activity and information securely.

NFR2. The system must make use of a chatbot or conversational agent

NFR3. The right permissions must be given for the microphone by the user, in order to send vocal messages.



REQUIREMENTS

FR1.The system should be able to engage in fluent conversations, directed towards the user's experience

FR2. The system should be able to narrate (audio) the messages that it communicates to the user

FR3. The system should perceive what the user is saying and act according to it.

FR4. The system should display the dialog between itself and the user.

FR5. The system should be able to detect the emotion expressed in the user's messages.

FR6. The system should be able to choose an appropriate story based on the detected emotions.

FR7. The system should be able to narrate the story, pause it, ask questions and wait for the user's response to them, before continuing the story.

FR8. The stories should be interactive, include questions about the character's emotions or actions.

FR9. The system should provide an explanation for why the user or the character in the story feels a specific emotion.

FR10. The system should offer suggestions about how a detected emotion can be managed.

FR11.The system must provide the child with the continuation of the story based on his previous answers.

FR12.The system must have a predefined set of stories that are composed by sections that have description, questions, possible answers (suggestions), and the next steps (sections) that can be reached, depending on the emotion and the answer of the child.

NFR1: The system must store the user's activity and information securely.

NFR2: The system must make use of a chatbot or conversational agent

NFR3: The right permissions must be given for the microphone by the user, in order to send vocal messages.

NGR diagram : Needs, goals, requirements

2. State of the Art

The state of art of solutions to similar problems is divided in three parts. The first one is about papers dealing with emotional conversational agents and emotion detection. We have also done research about studies on the notion of emotion for children with cognitive disabilities. Finally, we have analysed an existing project on the market.

2.1. Research papers

The state of art of this project is divided in three parts. The first one is about papers dealing with emotional conversational agents and emotion detection. We also made research about studies on the notion of emotion for children with cognitive disabilities. Finally, we inspired from existing project on the market.

The first step of our design process was to decide the nature of the agent, will it be a virtual or a robotic agent? We found several answer in the paper written by Sandra Costa, Alberto Brunete, Byung-Chull, Nikolaos Mavridis, **Emotional Storytelling using virtual and Robotic Agents**. It deals with the effect of having a physical embodiment of the agent on the listener's attention. A storyteller can adapt to the user's speech, it is about real-time and face-to-face interaction.

Recreating the characteristics of a storyteller into a conversational agent requires to take multiple details in consideration such as the nature of the voice. A synthetic voice is less engaging than a human voice, which will be perceive as more natural. Moreover, in order to create a personalized conversation, the storyteller is aware of both the internal and external emotions of the audience. When the listener enters the story world and empathizes with the character it corresponds to internal emotions. External emotions occur when the listener adopts the good narrative discourse and expresses emotion described in the story. In order to help children to understand their emotion we need to resort to both internal and external emotions.

The research, **Personalised Avatar on Social Stories and Digital Storytelling: Fostering Positive Behavioural Skills for Children with Autism Spectrum Disorder**, made by Kuan Tian Ying focuses on solving the problem for communication, interaction and action of ASD children between the age of 5 to 10 based on a digital storyteller. This storyteller was aiming to overcome social behaviour and communication skills, through the use of an avatar character that was well designed in order to trigger the child's attention so they could be more aware on road safety, in conclusion the overall research found out that the attention and awareness of all the children improved although the sample was small, so further research is needed. This paper allows us to see what works currently and how to improve on the investigation done by other persons as well as techniques and suggestions made by the researchers.

"A Survey on Conversational Agents/Chatbots Classification and Design Techniques" (published in Springer Nature Switzerland AG in 2019), by Hussain S., Ameri Sianaki O., Ababneh N., is a comprehensive study detailing the various techniques that are currently being implemented regarding conversation agents in general. It dissects the agent into its principal components which include the media through which it carries out its conversation, the methodology with which decides a response, the functionality the agent needs to carry out and the very nature of the conversation to be carried out.

In the paper **"Speech Emotion Recognition Using Deep Learning Techniques: A Review (2019)"** written by R. A. Khalil, E. Jones, M. I. Babar, T. Jan, M. H. Zafar and T. Alhussain, we can see an overview of the performances of various Machine Learning algorithms used for emotion detection from speech. The latest improvements have been done usign Deep Learning techniques, which do not require manual feature extraction and classifiacation. Unfortunately, we have not had time to build and train an accurate Deep Neural Network, therefore we have used an external service for emotion recognition from text, Natural Language Understanding from IBM Watson.

2.2. Research project

Emotion is the demonstration of the affective life of an individual, therefore it is subjective. Several centres of research base their studies on this unpredictable characteristic of human being. The child Emotion Lab at Weismer Center University of Winsconsin works mainly on the identification of emotion. In 2020, a study the impact of the mask on the emotion detection by children has been conducted by Ruba AL and Pollak SD and described in this paper**Children's emotion inferences from masked faces: Implications for social interactions during COVID19.**

Reasearch studying the impact of mask on the emotions detection by children

The aim of this research is to understand what is the impact of occulting part of the face, on the identification of emotions by children. The target group of this study is children without disabilities aged from 7 – 13 years old. The results demonstrate that covering the eyes or the mouth impacts the identification of several emotions. For instance, identifying 'fear' requires both the information given by the eyes and the mouth, otherwise this emotion can be mixed up with 'surprise'. However, it is unlikely that children draw emotional inferences based only on facial configuration. The understanding of the emotion depends on the context, the body posture and gesture and vocal information. This study demonstrates that children base on both the eyes and the mouth to identify an emotion but also on elements which are not related to facial expression.

This research can be crossed with the paper of Marine Cesmat, **L'anomalie de discrimination des emotions du visage sous-tendrait le déficit social caractéristique de l'autiste**. The author demonstrates that children with cognitive disabilities overlook several information of a situation during the identification process. Indeed, they don't take in consideration, for example, the context, the body gestures or the shape of the eye. Only the shape of the mouth will be used to understand the emotion. These two studies demonstrate that children are not able to identify an emotion by relying only on one element of the face. In our project, we will need to highlight all the elements of a situation, which can help children to understand character's feelings. In fact, our service needs to train children ability to cross several information of a situation (gesture, mouth, eyes, vocal) which is a necessary step to understand emotions.

2.3. Project on the market

In order to understand how to convey emotion trough storytelling we analysed several projects based on conversational agents, virtual interfaces but also tangible artefacts. In 2002, Telecom developed a conversational agent called Greta. It is a three-dimensional realistic avatar, which was animated. The virtual woman is able to communicate using a rich range of non-verbal behaviours. Greta can hold a conversation and adapt her gesture, facial expression, gaze and head movements. In the same way, Papous, a virtual storyteller can express six basic emotions through its facial expression and body language. Thus, it can manage joy, sadness, anger, disgust, surprise and fear through its speech and virtual representation. All the emotion can also be expressed with different intensities.

A child with cognitive disabilities needs more time to understand and succeed in an activity than other children. He or she will follow his or her own rhythm and requires an adapted training. Conversational agents and artificial intelligence are able to adapt to their users. Today, there are several services on the market able to offer a personalized experience. The social express is an application using storytelling and interactive activities to illustrates emotions. Through images and spoken dialogs, the characters express their feelings. The child is asked to understand the situation and to continue the story. The app also proposes strategy to help the child managing his or her emotions. Understanding an emotion can also occur in the physical world trough teddy bears. The project Kimochis helps children to associate facial expression to an emotion. It offers a physical support to the child who can visually materialise an emotion and externalise his or her own feelings through the toy.

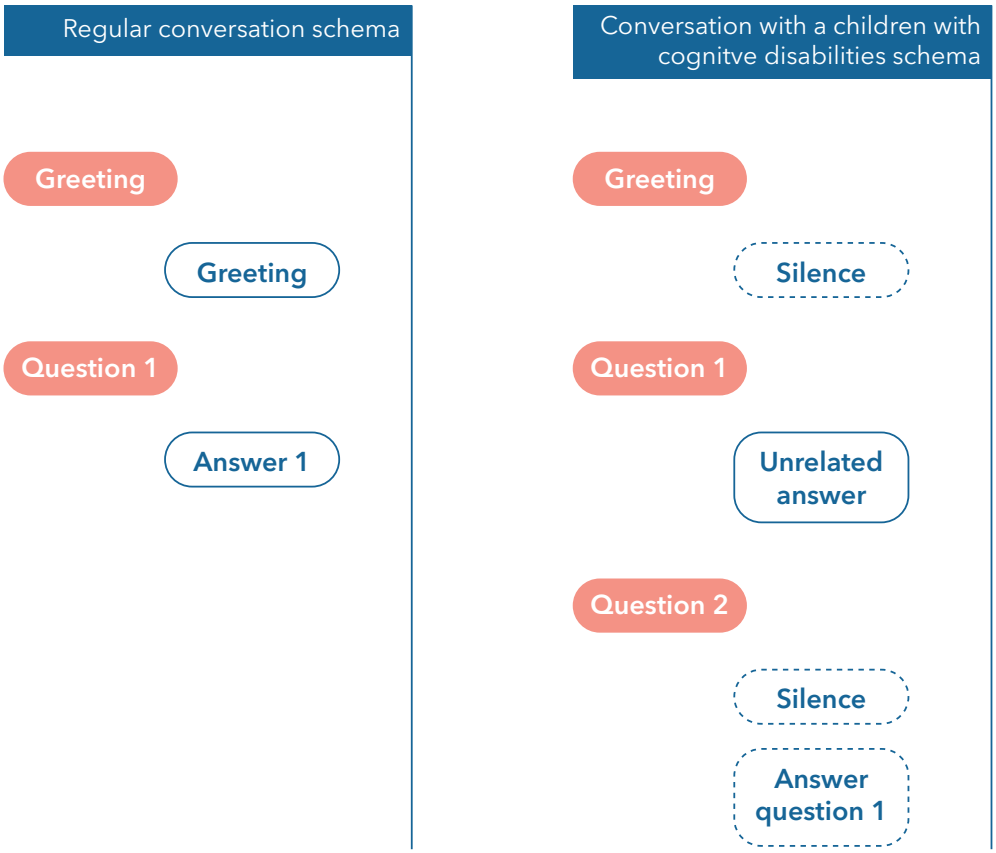
The Social Express is an application helping children to identify and manage emotions

3. Solution - UX Design

3.1. Exploration through Digital Ethnography

The first step of our design process was to try to understand the user. We based our study on the state of the art, the research papers and the studies on children with cognitive disabilities which allowed us to draw the big picture of their needs and the possible solutions that can be implemented in this project. We had to take into consideration how the child experiments daily activities and how they can impact his state of mind. We also did research about social interaction for people with special needs in order to try to anticipate their reactions throughout the conversation with the chatbot.

In order to understand the context and the user, we used digital ethnography. The idea was to adopt design methods cantered on children with special needs, instead of usual User Experience methods. Through research studies, therapist interview, parents' testimony and online videos, we tried to adapt regular conversation rules to children with cognitive disabilities codes. First, this programmed conversation needs to follow child's rhythm. In a normal discussion, the question and answer come one after another in a fluid way. The person answering to the question reply in a few seconds or express his or her intention to answer. A conversation with children with cognitive disabilities is more fragmented. The child imposes his own rhythm to the conversation and can answer in unpredictable way. He can ignore a question, easily loose interest in the conversation, take more time to reply or answer to a question asked a few minutes ago. During a discussion, it is recommended to wait 20 seconds before asking the question again, the child can be distracted and needs sometimes to focus on the conversation again.



Children can perceive a simple question such as “how are you?” as a stressful experience, so the chatbot should introduce itself before testing the child. However, it cannot notice the distress of its user that’s why the application should be used with the help of the parents. They have an important role in supporting the child in his answers and help him to stay focused throughout the story. The application is recommended to children aged 6 to 12 years old. In order to ensure child involvement throughout the interactive experience, the story and the activities should not be longer than 6 or 7 minutes. The child should be able to use the application at home, so that it can be easily integrated in the routine of the child, for example before bedtime.

Furthermore, each story should be verified and approved by a therapist. In fact, this application should be recommended by an expert to allow the child to continue his training between two sessions. A person with cognitive disabilities needs more time to assimilate a new concept. Emotional learning requires a lot of practice, when it is not intuitive and easy to understand for children with special needs, so we should provide a tool which lets them practice and repeat at their own pace, while being supportive and fun to use.

3.2. Persona & Scenario

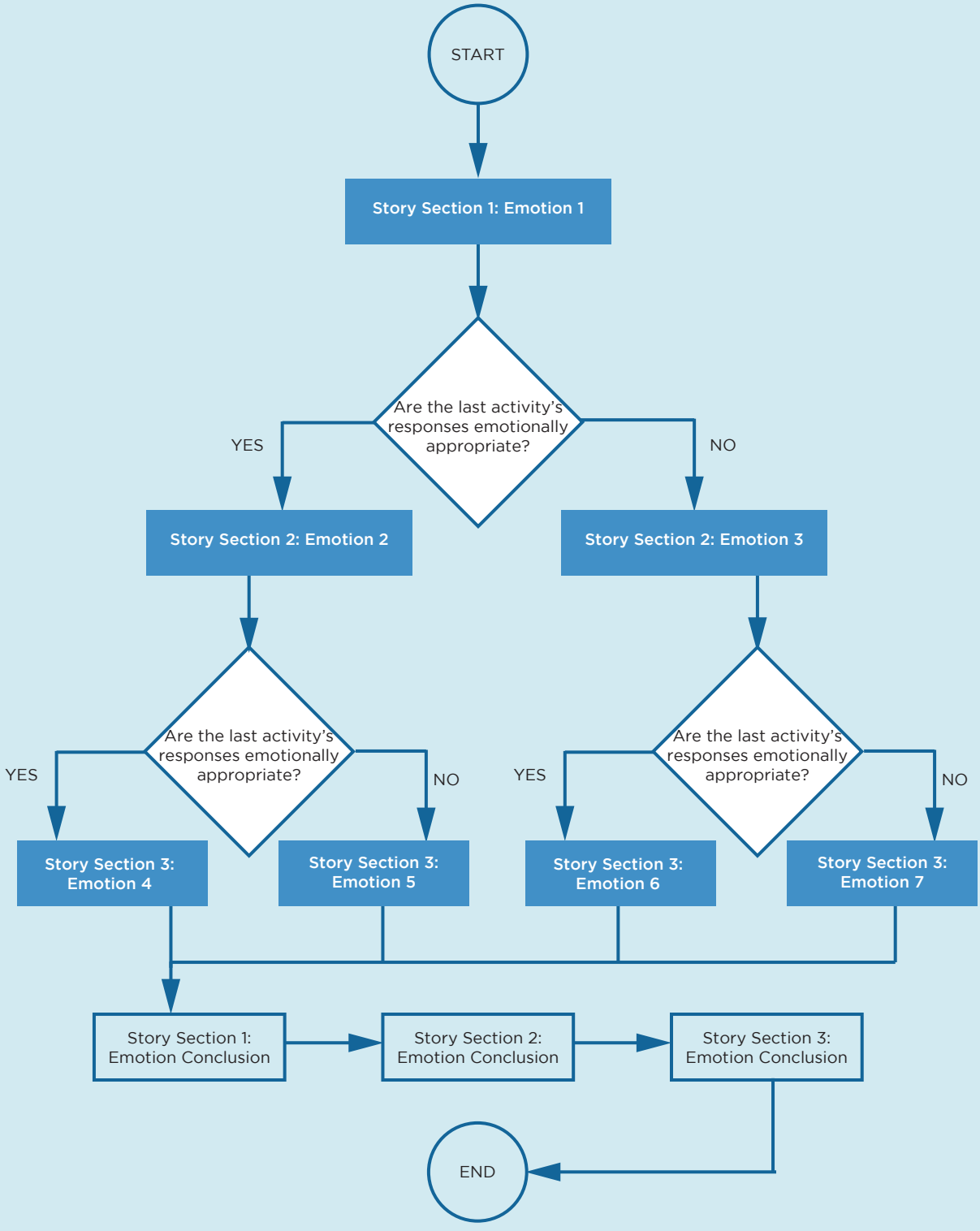
Based on the digital anthropology research, we drew the persona of our potential user. This profile doesn’t represent all the children with cognitive disabilities, but it is a fictive representation with key characteristics that are met among them. This persona depicts the profile of Nico, a 7-year-old child with a mild form of autism. He has already started his emotional learning with a therapist. Our application is recommended by the expert because Nico is unable to react appropriately whenever a situation occurs, being confused by his emotions. Just before going to bed, his parents are used to reading him a story, so it is easy to replace the book-based story time with the Emotional Storyteller Berno. The chatbot asks Nico about his day and proposes to listen to a narration. Nico listens to the interactive story, which sometimes reminds him of his day or past memories. He likes when he is asked to give his opinions on the characters’ feelings or actions. When the story ends, Berno talks to him about the emotions that were present in the stories and how to manage them. He may not always understand them at first, the more he practices, the better he will manage them. When new situations occur, he is able to associate them with events in the stories, so he understands better what is happening and how to manage the current emotions.

3.3. Structure of the Hyperstory

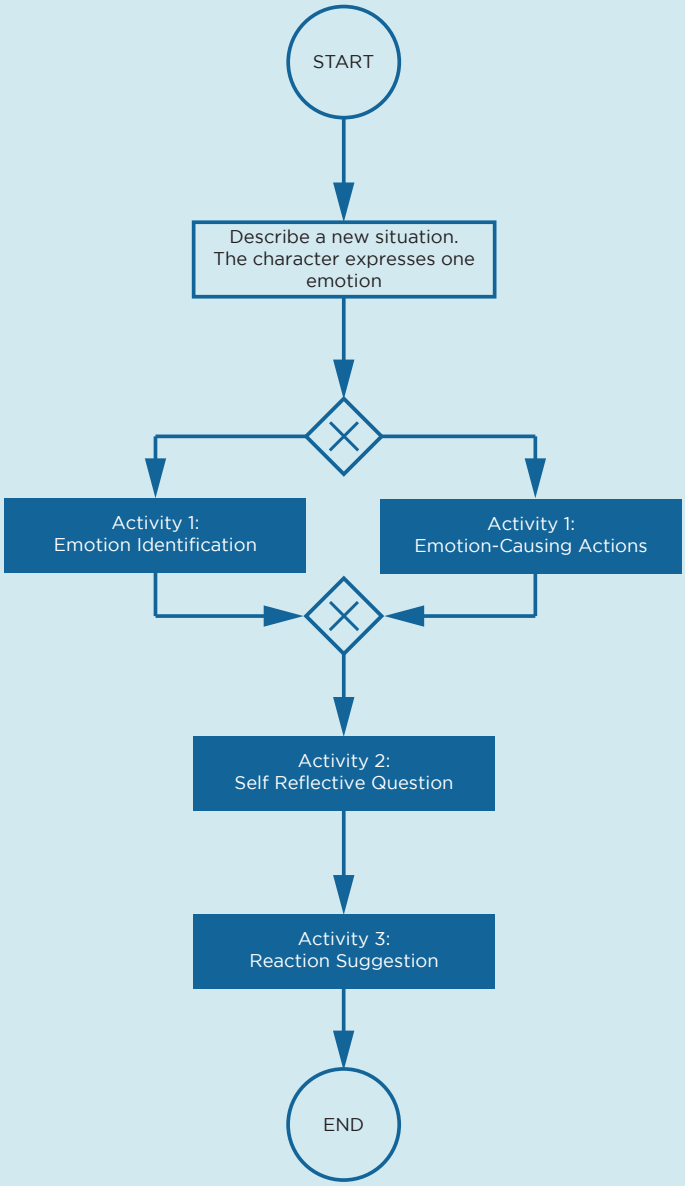
The stories provided by the chatbot are interactive, with multiple endings, which are influenced by the child’s decision and responses to certain activities. Thus, we have build them as hyperstories, which are composed of multiple story sections that have a corresponding emotion to focus on. The story sections are further split in activities.

Hyperstory

The hyperstory is a full narative with characters and actions, which is composed by a variety of story sections, that are reached by the child depending on his answers to previous sections. At the end of a story section, the hyperstory will continue with another section, that focuses on a different or similar emotion, based on the child's answers. After the last section, the hyperstory goes over the emotion that has been discussed in each section, so that the child can review the story and what he has learned.

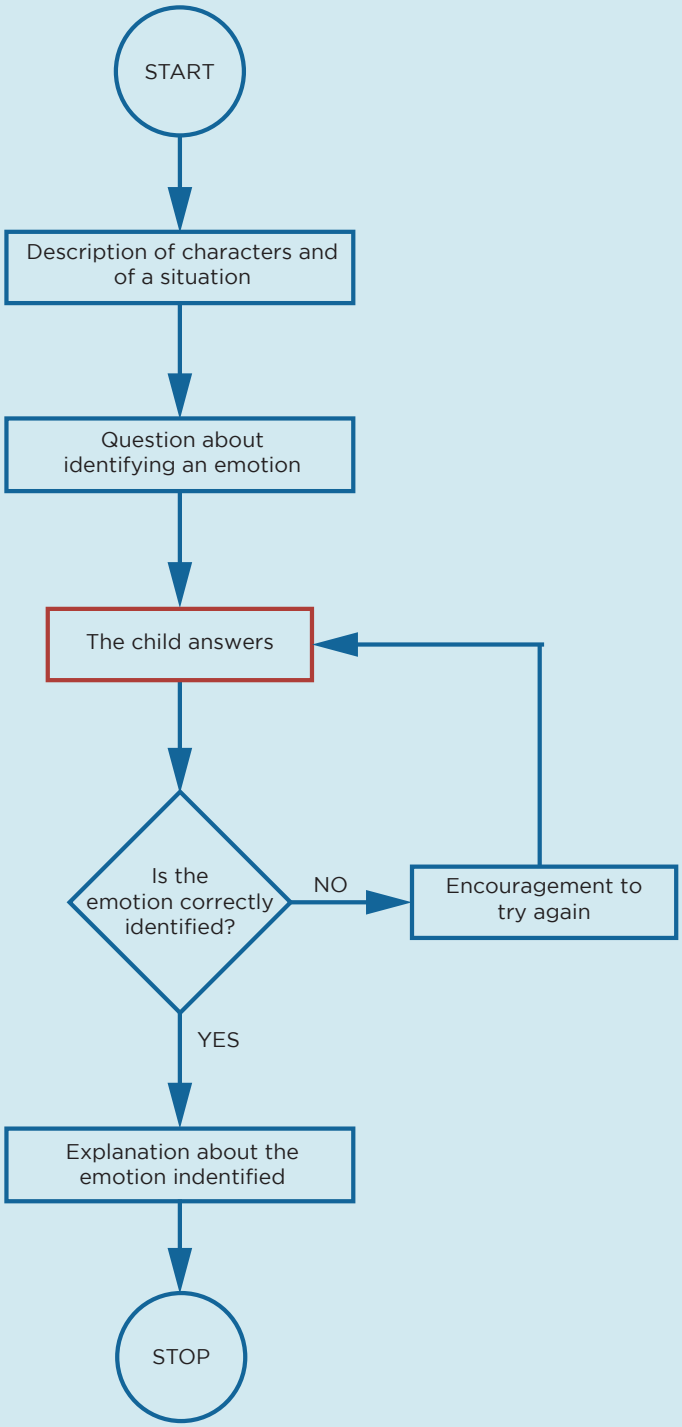


Story Section
A story section corresponds to an episode of a story, which is composed of a description of a situation and a set of activities. The activities that can be part of a story section are the Emotion Identification or the Emotion-Causing Actions as the first activity, followed by a Self Reflective Question, and ending with a Reaction Suggestion.



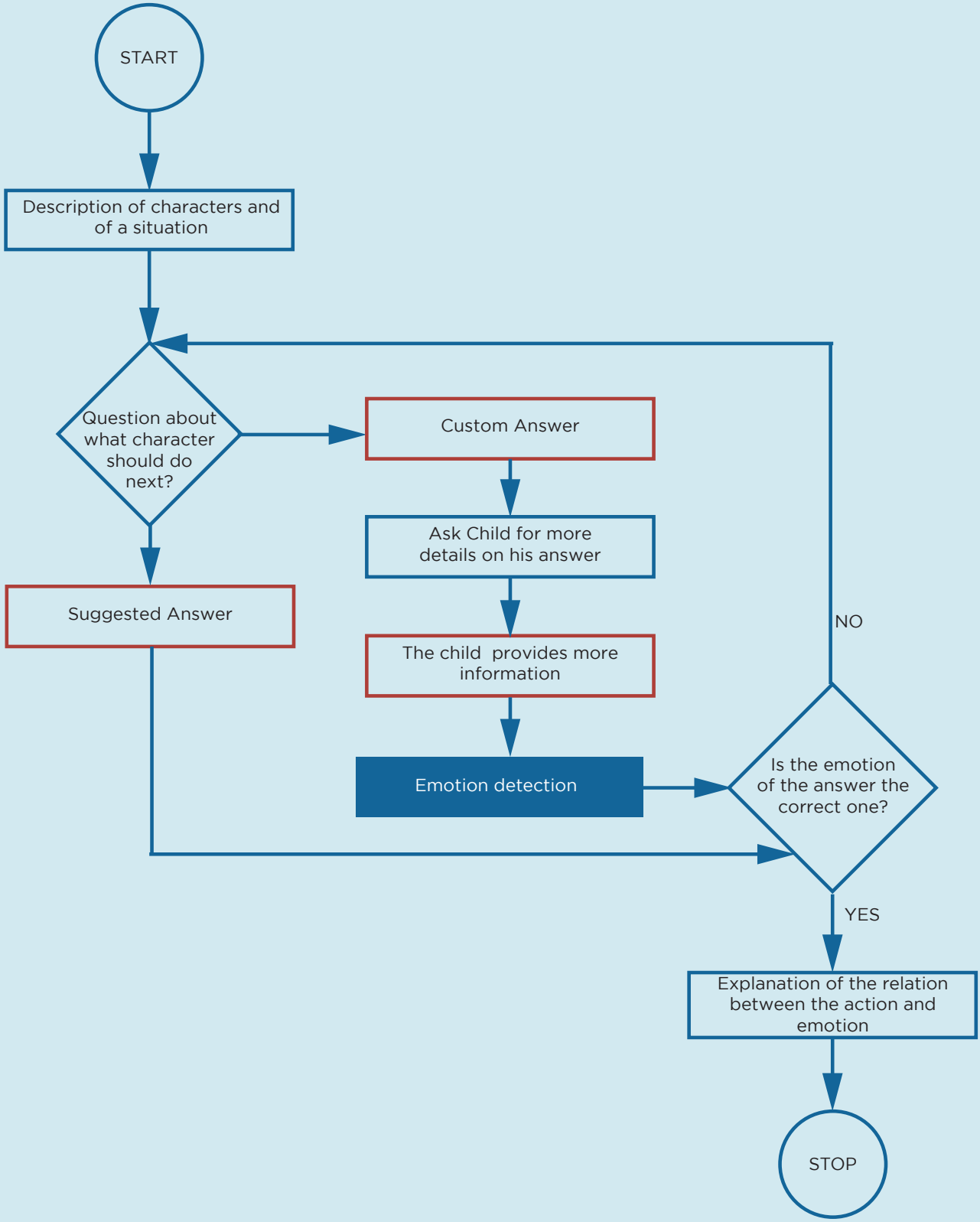
Emotion Identification Activity

This activity is one of the first activities in a story section, and its goal is to train the user over the identification of emotions from a situation, based on the textual description and images as visual clues. The chatbot presents the characters and the context they are in, then it asks the child a question about an emotion that can be deduced from the situation. If the child's answer is correct, the chatbot congratulates him and explains the emotion and what are the story elements that suggested it. If the answer is wrong, the child is encouraged to try to answer again.



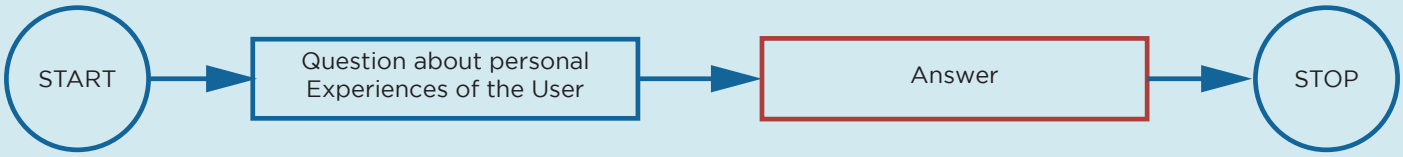
Emotion-Causing Actions Activity

This activity's role is to help the child learn better that actions may trigger certain emotions, such as playing with your sister may trigger happiness. The chatbot presents a situation and the characters involved, and the chatbot is asked what the character should do, in order to trigger a specific emotion. The child can give a suggested answer, that the chatbot already knows, or he can input a custom answer, to which he is asked to offer his reasoning. Then, the child's input is handled by the Emotion Detection module, which returns the emotion detected in the last 2 input messages. If the emotion is the expected one, the chatbot congratulates him and explains the relationship between the action and the specific emotion. If the answer is wrong, the child is encouraged to try again.



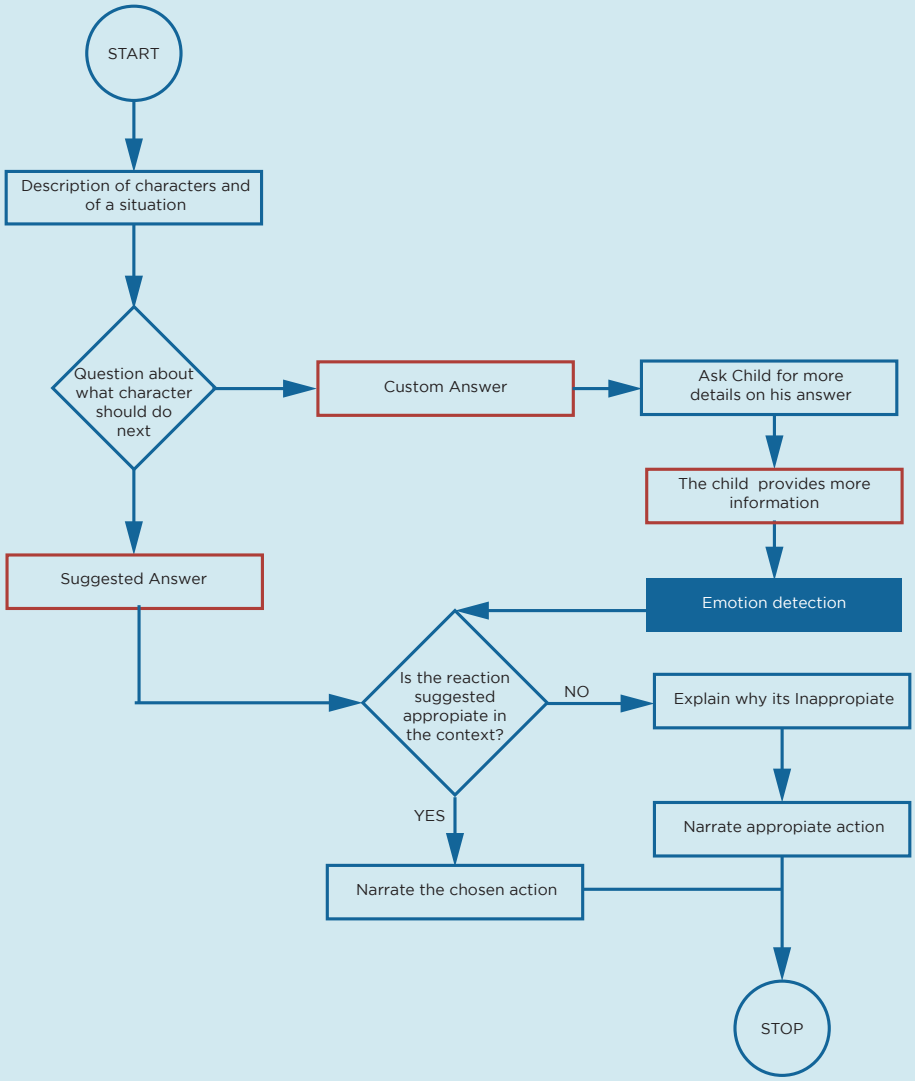
Self-Reflective Question Activity

The second activity in a story section that the child may encounter is the Self-Reflective Question, which is a question about the experience of the child about the emotion of the current story section, or about other story-related details. The goal of this activity is to encourage the child to help him identify personal experiences that are similar to the story, so that he can understand better the situations that he was in.

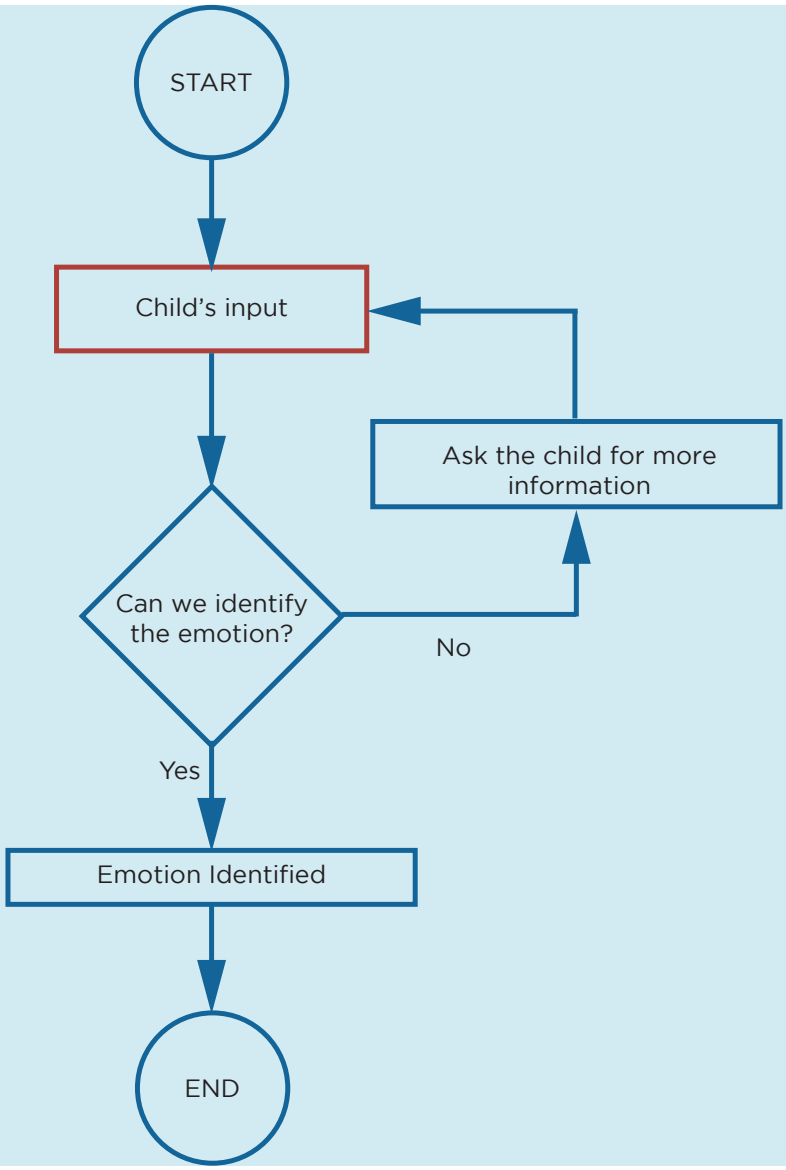


Reaction Suggestion Activity

The role of this activity is to help the child empathise with the characters in the story, and suggest actions that they should take next. It usually follows after a Self-Reflective Question, and after the situation and the characters have been already introduced. The chatbot asks what the character in the story should do next, and the child can give a suggested answer, that the chatbot already knows, or he can input a custom answer, to which he is asked to offer his reasoning. Then, the child's input is handled by the Emotion Detection module, which returns the emotion detected in the last 2 input messages. Then the emotion is checked agains the acceptable emotions for the specific context, and if it is among them, then it seems that the child has understood the situation and has suggested an action appropriate to the context. If not, then the chatbot explains to the child why his suggestion may not be appropriate and offers acceptable alternatives.



Emotion Detection Component
The Emotion Detection receives the child's last custom answer and reasoning (so 2 input messages) and attempts to idenfity an emotion in them. If an emotion has not been identified, we are asking again for more information from the child. If we manage to identify an emotion, the emotion is returned as a result.



3.4. User Journey

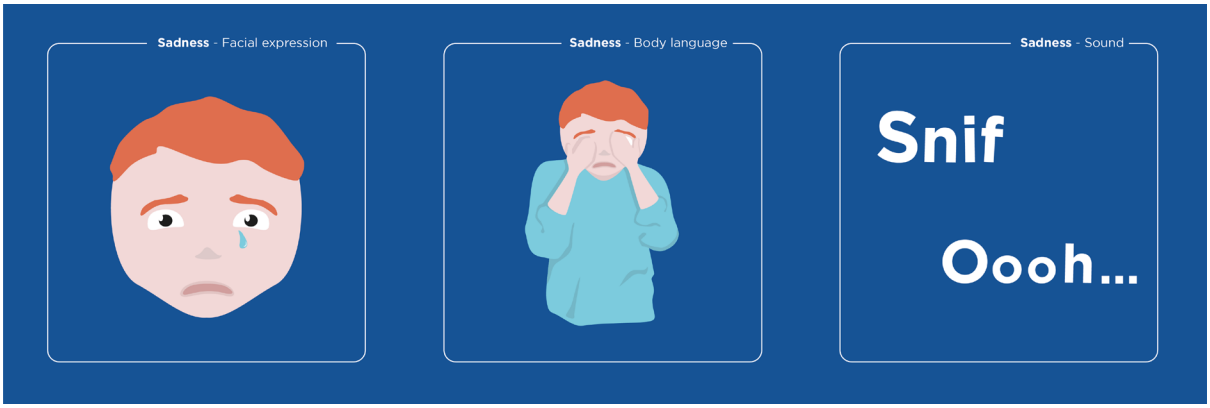
A child with cognitive disabilities can reach the same learning level as other children, but it will require more time. In fact, a simple activity needs to be divided into several steps. In order to support the child in his understanding of emotions, the chatbot breaks down the situation. If a child without disabilities could easily make this activity without any help, our user needs help to focus on important information. Once essential elements have been highlighted, the child has enough tools to understand the context and identify emotions.

Thanks to the flexibility of the chatbot, the parents can have a role in the storytelling. If necessary, they can take time to explain the situation to the child. When he is ready, he can press the button to speak to the chatbot, which remained silent waiting for the child's answer. Moreover, if the child needs more explanation, he can press the 'play' button of the dialog and the chatbot repeats its message.

Each session is intended to start with an open domain conversation. It has the advantage of both preparing the child to hear a story and attempting to detect an emotion from his messages. The questions asked to child are related to daily activities. For example, if the application is used at the end of the day, the chatbot will ask about what the child did during the day. By answering to the question, the user may express specific emotion, which the chatbot can detect and play a story according to it. Another example is that, during the initial conversation, the chatbot asks the child 'How are you feeling?'. This is a self-reflective question, which the child may not answer accurately at the beginning of his training, but after a few sessions he may express a feeling. At the end of the open domain conversation, the chatbot starts to narrate the story. The fictive scenario is composed by different types of activities and questions, some of them being open questions, yes-or-no question, or story-based activities.

At the beginning of the story, the child is asked to identify an emotion. If he is not able to answer, some proposals are displayed after 15 seconds of silence. He has to say how the character is feeling based on the situation which is described and illustrated. The story continues only when the child is right. Usually at the end of the story section, the child is asked to suggest and appropriate action for the character to take, and the chatbot will attempt to identify the emotion that results from his answer. Any detected emotion is considered correct, but the context of the situation will determine if the reaction or the action that the user chose is appropriate or inappropriate. The child's answer has an impact on the continuation of the story. If the child has been able to choose the appropriate reaction, it means he understood what the character felt, and the story continues with a different emotion. However, if the action he chose is not correlated to character's emotion, the chatbot will chose a path that focuses on the same emotion, or on a related emotion, in order to practice and repeat the notions more times. The use of multiple paths in the story helps the user to learn the consequences of his actions, as well as leaving open for redoing of the story without making it too repetitive

Throughout the story, the child learns to identify an emotion and how to react to it. He can remind himself of these fictive scenarios in real-life situations. At the end of the story, the chatbot makes a sum up of the emotions the child learned. In order to illustrate the learning of each story section, children are provided with a visual tool that presents characteristics of the emotion. The child can rely on facial expression, body language and the onomatopoeia of the emotion to identify it.



Example of the 'visual toolbox' about sadness

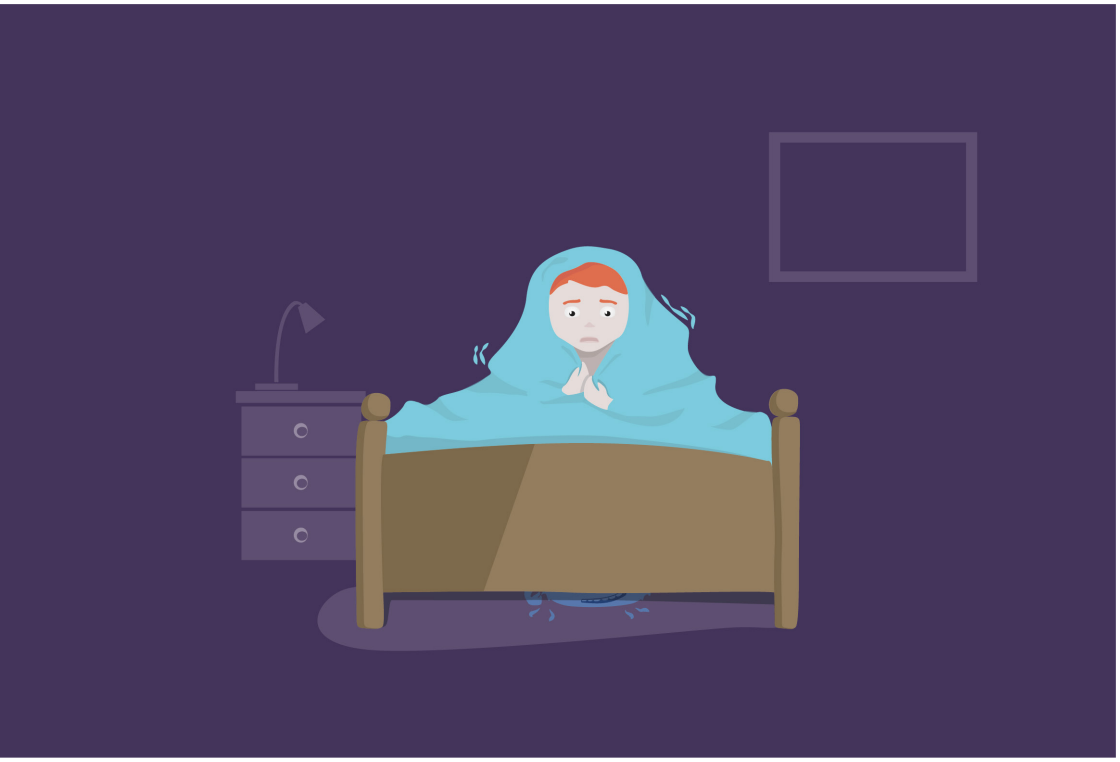
3.5. Design Choice / Interface

In order to improve child's involvement, we observed methods used by therapist with children with special needs. Abstract concepts such as emotions are more difficult to manage, that's why it is preferable to have a visual support. Usually a therapist uses gestures, images or even 3D object to materialise an idea, we inspired from this visualization methods to create the interface. Having a screen-based interface is essential for children. The application is a flexible and interactive tool which facilitate the engagement and the attention of the child throughout the session.

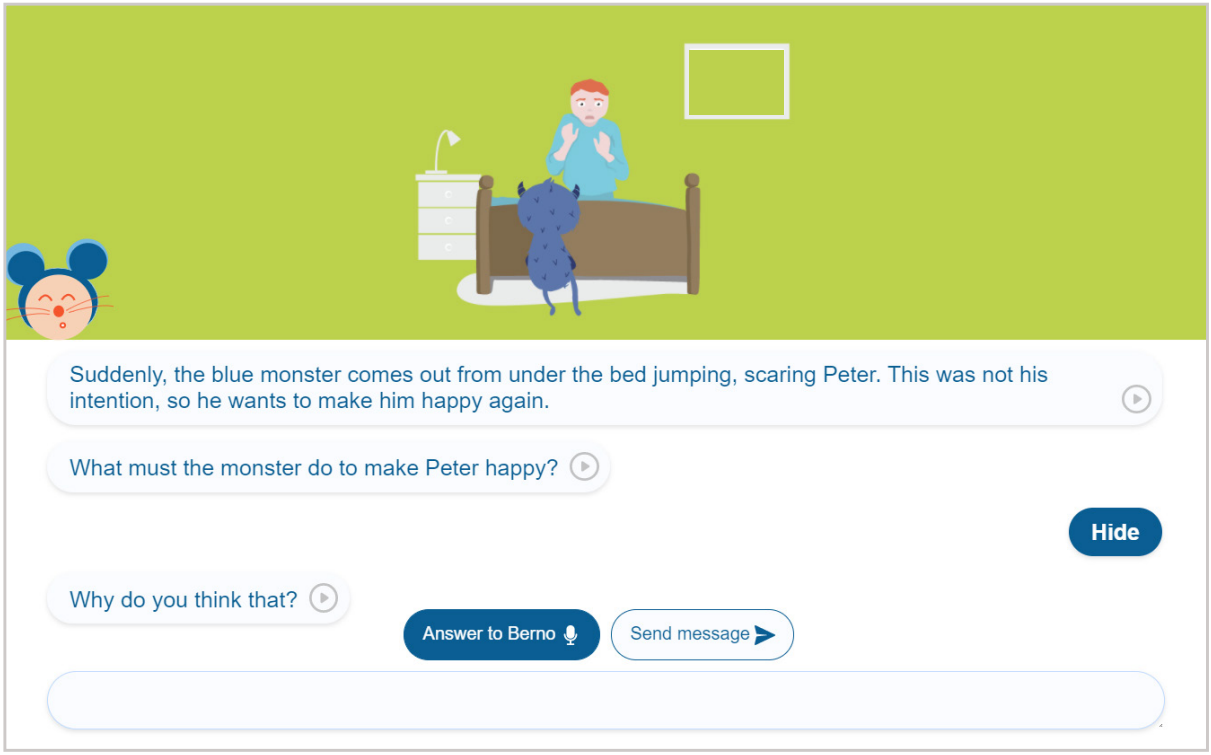
Furthermore, during a social interaction, sometimes children express shyness and they have difficulties to maintain a visual contact with their interlocutor. As they are sensitive to stimuli of the external world, they can lose interest in the conversation and focus their attention on, for example, an object around them. They have a fast visual flow and looking at the support of the conversation can help them to be more present. Embodying the chatbot in a visual interface catches child's attention and creates a support for the parents. It helps the child to visualise the emotion, which comes alive through the characters and the situation they encounter. Also, during a social interaction with a child with special needs, it is better to squat in order to be at the same height as him. The support of this interactive story is a tablet or a phone, the child has the ability to manipulate it. Having the control on the conversation may lead to decrease his shyness and be more engaged in the activity. In order to personify the chatbot, we used a human voice, which sounds more realistic and warmer than a synthetized one.

As our target group may be unable to read or write, the conversation needs to be voice-based. Using the same tools as a storyteller, the chatbot narrates the story out loud, add intonation and pauses to its speech. The child can rely on both the text of the story and the voice of the conversational agent. In order to simplify the interactions and make it more realistic, the child has the possibility to answer to the chatbot using the microphone, by pressing a button.

In terms of visual interface, the application is divided into two parts. The conversation with the chatbot is displayed in the lower part. The upper part is dedicated to the story, the illustration evolves according to the scenario and the emotion. In fact, we used one specific colour to highlight each feeling. When the story describes an emotion, the colour of the background of the drawing changes accordingly. This method can help the child to remind

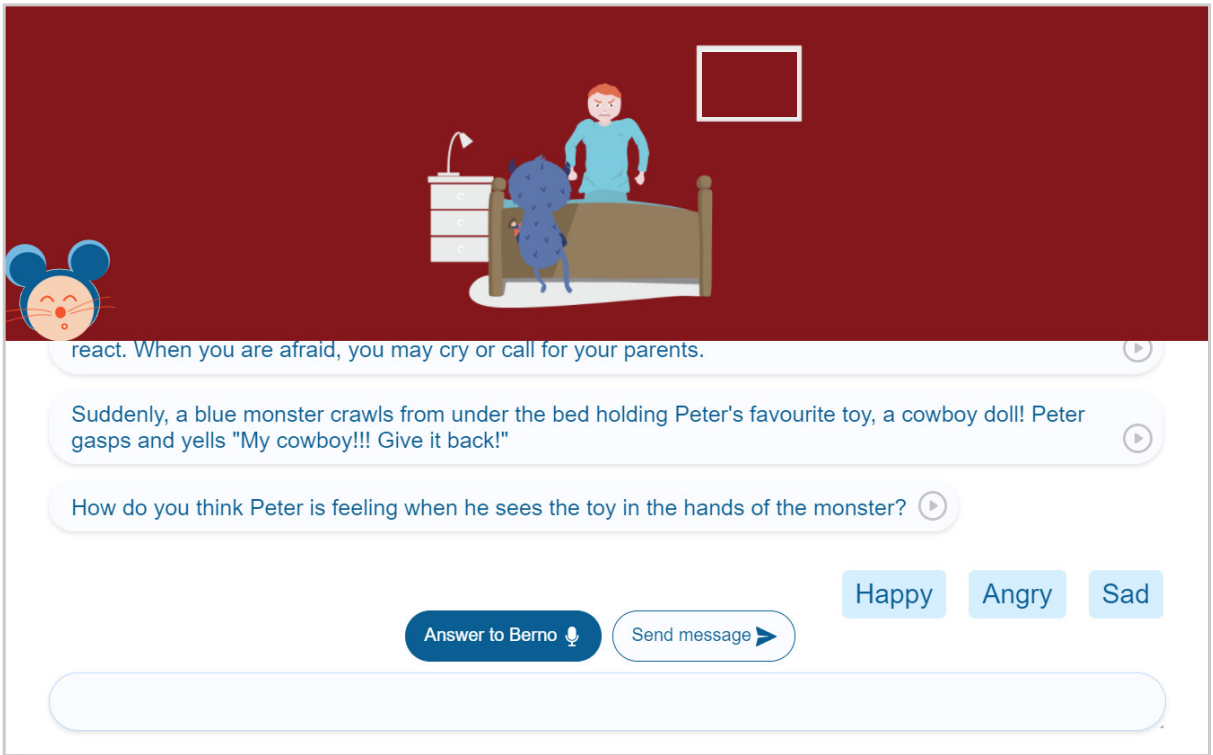


The colour of the background changes according to the emotion



Throughout the story, the chatbot offers different type of activity

him an emotion, by associating it a colour. The application is based on the learning of the six basic emotions; happiness, sadness, fear, anger, surprise, disgust, that's why we used six different colours.
The illustration themselves had been designed in order to emphasize the character's emotions. Children with cognitive disabilities have a fragmented perception of the world, they have difficulties to make the connection between several elements even in the same pictures. They can be distracted by the table or the lamp at the background and don't give attention to character's facial expression. That's why we decided to almost make the decoration disappear in favour of the characters themselves and the element of the context which can help the child understanding the emotion. Facial expressions (eyes, mouth, eyebrow) and body languages (gesture, position of the body, relation of the body with the objects) have brighter colours. Thus, visual interface is made to favour the story and give additional tools to the child.



Illustrations give tools to the child to better understand the emotion (facial expression, body language, context)

4. Implementation

In this section we are describing the technical implementation details of the emotional storyteller, such as the high level architecture of the client and the server, the database schema and the technologies adopted. The product does not need a special hardware infrastructure to be used. It can be accessed through a browser, on mobile or desktop devices.

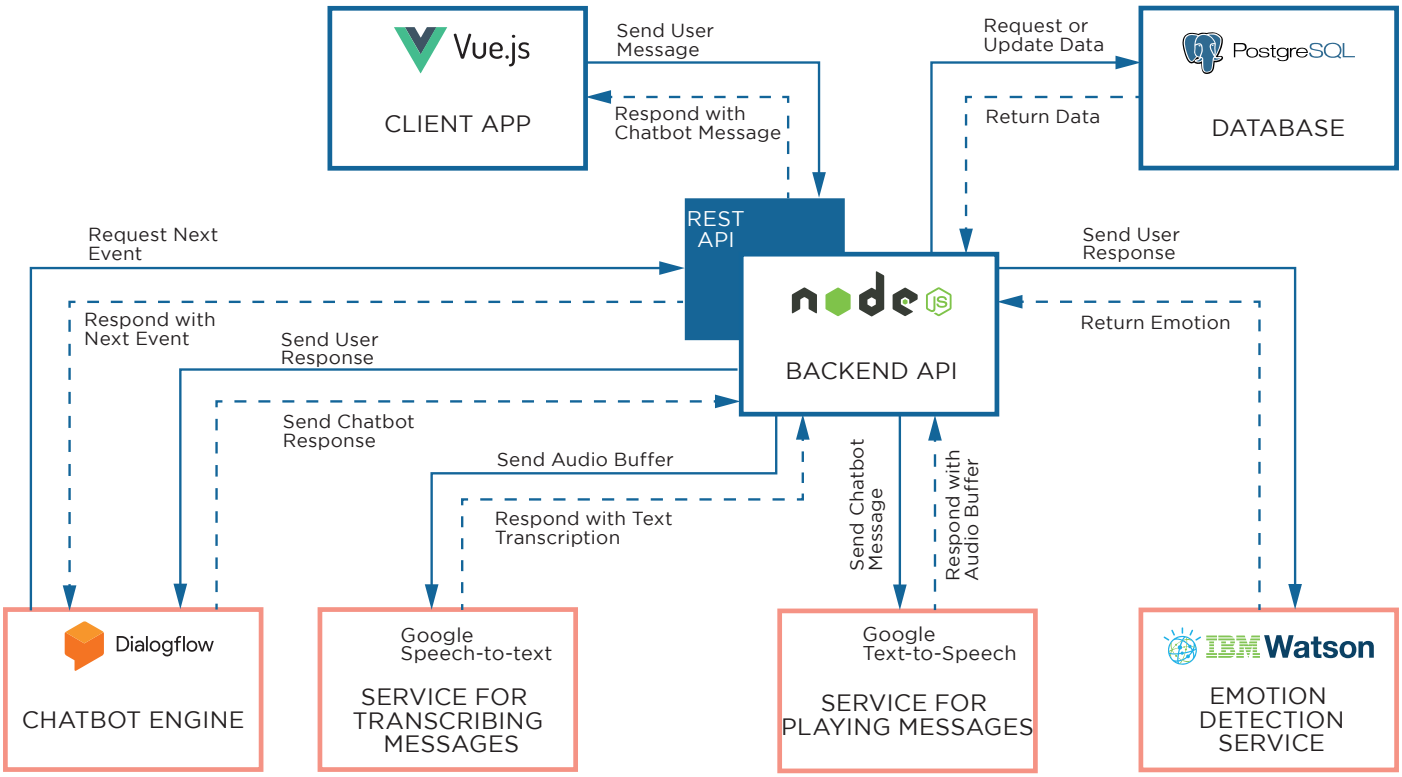
4.1. High Level Software Architecture

We have chosen to build the emotional storyteller as a cross-platform, mobile-friendly web application, so that it will be accessible from a wide variety of devices. We have built it using the 3-teir Architecture, which consists of a client application, a server application that holds the business logic, and a database.

The software respects the separation of concerns by handling the functionality in components that have specific roles, and which communicate with each other. The end user interacts with the client application (front-end), which sends HTTPS requests to a REST API that offers access to the server application's business logic. The server application (back-end) connects to the database, in order to execute various operations for saving or getting data, and makes use of external services such as DialogFlow through API calls or the use of SDK libraries.

The most important high level components that are involved in the functionality of the application are:

1. The Client Application – a JavaScript web application, built using the framework Vue.JS
2. The Server Application – a Node.JS application, written in JavaScript
3. The Database – a SQL, relational database, built in PostgreSQL



Architecture Schema

4. The external services:
- Google DialogFlow – used for creating and managing the chatbot interactions
- Google SpeechToText – for transcribing the audio input of the user
- Google TextToSpeech – for playing the messages that the chatbot is sending
- IBM Watson NaturalLanguageUnderstandingV1 - for detecting emotions in the user's input.

4.2. Client Application (Front-End)

The client application (front-end) is the a web site what allows users to interact with the emotional storyteller chatbot. It built using the Vue.JS framework, and written in JavaScript, HTML and CSS. Its architecture is modular, having components which interact with each other, or with independent services. The front-end application communicates with a back-end application through HTTPS requests towards a REST API, which handles the data manipulation and functionalities that provide the content displayed by the client app.
The index.html is a simple HTML template file that serves as the entrypoint to the application. In it, Vue.JS adds automatically the scripts that contain all the functionality, which are built and generated by Webpack, which is a module bundler that takes the various components and creates one file with all the functionality. The main.js file is the place where the Vue application is instantiated, along with the 3 main modules that we use:

- App – which is the main component that holds all the other components
- Vue Router – a module which is useful for navigation among different pages
- Vuex Store – a module that allows accessing and storing information accumulated during a session in a common place.

There are 3 categories, or layers, of functionality and components, that interact with each other:

1. Pages – the pages are loaded with the help of Vue Router. The application currently holds only 1 page, in which the chatbot conversation and stories are playing. The pages are Vue components that hold and organize other components;

- StoryPage.vue - the main page of the application, it holds the components for displaying the chatbot and main interaction tools: StoryCover, Conversation, and InputMessage.

2. Components - a common Vue component is an independent piece of functionality, with a well defined role, that has a visual template, a set of styling rules, and functionality. The components can interact with each other, and can be used to create other components. They are using services in order to extend their capabilities;

- StoryCover.vue - the header of the chatbot, which holds the character Berno, as well as the images of the story when a story is playing

- Conversation.vue - the component handles the display of the conversation, as well as the creation of new messages when the user sends a new message, or when the chatbot response has been received;

- Message.vue - this component is used by Conversation.vue for each message, and it has the functionality of re-playing a chatbot message, as well as displaying suggestions for the user, when appropriate;

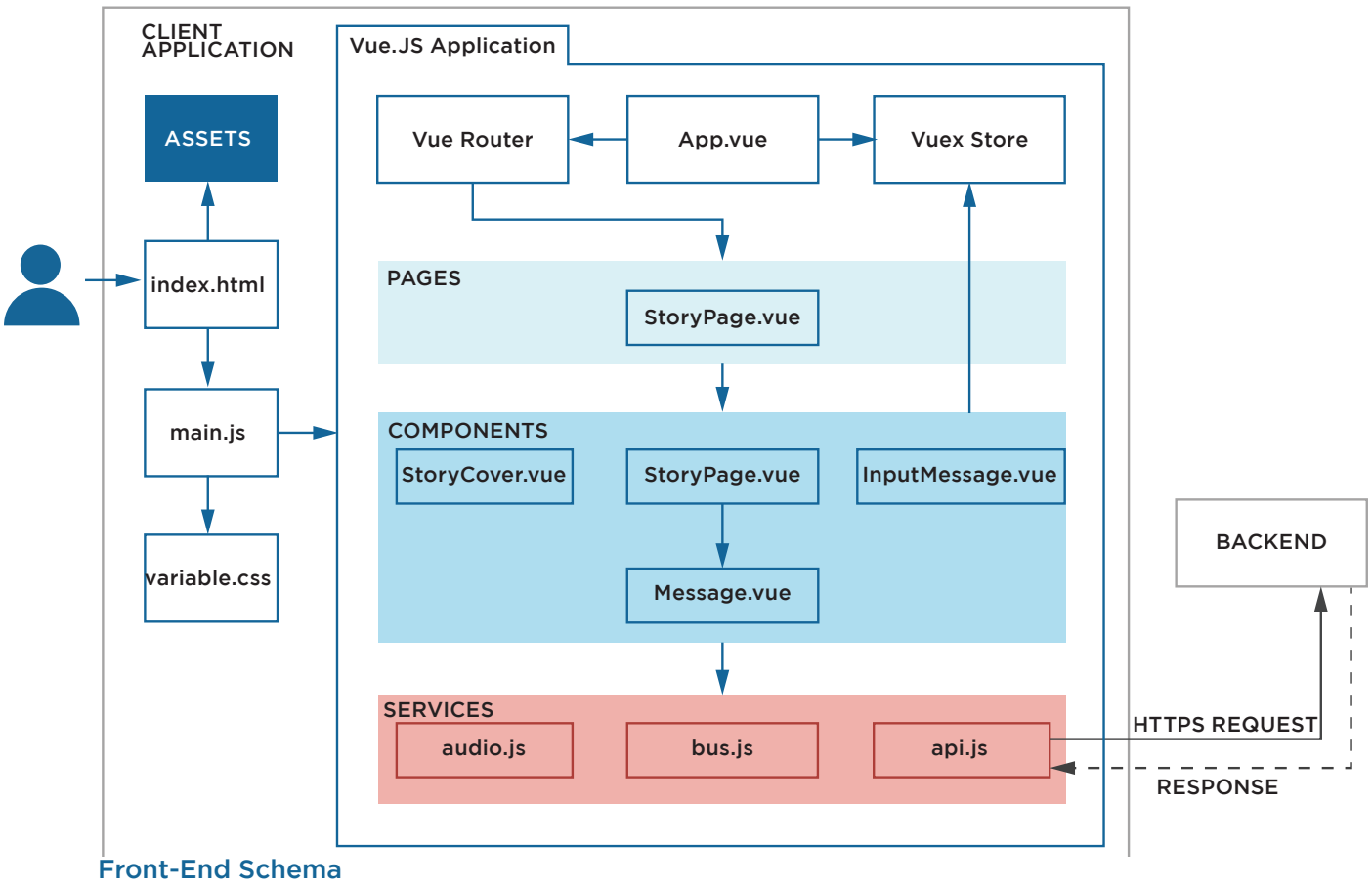
- InputMessage.vue - the component is used for triggering the capture of audio input or handling text input from the user, as well as requesting the API service to send it to the back-end.

3. Services – they are Vue modules that hold functionality which serves to all components, especially related to content management and data handling.

- audio.js - the module is in charge with the functionality of recording messages, as well as playing messages in buffer format;

- api.js - the module is the main component that makes requests to the back-end, among which we can mention requests for getting the audio buffer of a message from text, decoding a vocal message into text, or getting the next chatbot response based on the user's input;

- bus.js - the module has no outstading internal functionality; it is used for emitting and catchig custom events between components, in order to assure a loose coupling.



Front-End Schema

4.3. Server Application (Back-End)

The server application (backend) is a REST API built on Express.js framework based in Node.js. It handles all the frontend requests and Google's Dialogflow Webhooks via data manipulation either through request specific database calls or through external APIs for additional services. The APIs are comparmentalized to ensure that individual components complete requests with minimum dependence upon each other.

The components of the backend can be seperated as such:

-The main script intergrating all the functionalities and is responsible for handling any request that is received on the assigned port is managed by index.js.

-Each of the services listen for requests through their respective Routes that are imported into index.js through which they serve as a connection to the services script itself.

-All the Services are made modular and provided appropriate responses to all requests sent through the port. T access to necessary information from the database via data access objects

-The Data Access Objects are used to structure queries their respective tables in the database. The connection from the script to the database for posting the queries is handled trhough connection.js

-The database configuration and the tuples for the individual tables of the database are stored in the migrations, seed and knexfile.js

Services:

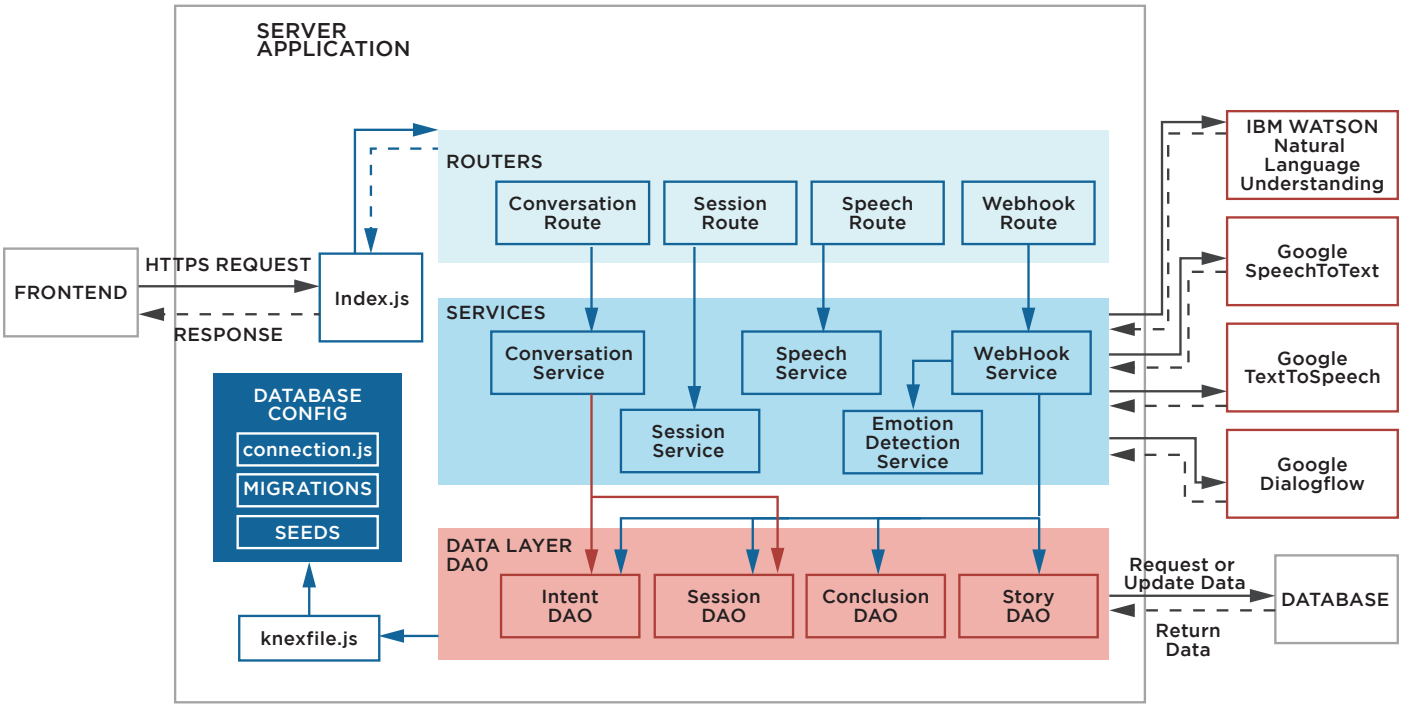
1. SessionID Generator - For every new connection setup in the frontend, a new and unique session ID is generated by sessionIDService.js for handling requests to dialogflow through the assignment of individual sessions for each user. This is done to maintain simultaneous chat sessions with dialogflow when more that one user is using the application. Thus multiple users are able to interact with the application with all the required session information stored under their respective ID in the database, providing reliable data management.

2. Conversation - This is the main service provided to the frontend of the application. It interacts with the chatbot that is created on dialogflow to provide responses to the user via conversationService.js. This is where the frontend provides the user queries to the chatbot and recieves its reply. Here a connections with dialogflow is esablished through a session ID to avoid conflicts that would arise from having multiple users. It also updates the database on the various emotions covered in the conversation so that suggestions can be provided to the user based on what they have observed. It also provides the user with suitable suggestions and relavent images URLs for certain chatbot responses

3. Webhook - It control the flow of the conversation between the user and the chatbot by switching between the various contexts from open domain conversations to stories and finally its conclusion with suggestions. It is also used to call the only service which is not accessable through the backend api which is emotion detection. This is used to trigger actions in dialogflow through events to alter the course of the conversation based on the emotion detected from the user's responses.

-Emotion Detection - IBM Watson's Natural Language Understanding API is called whenever the user's response needs to be analyzed to extract the emotional context of a user's reply to help dialogflow decide how the conversation is meant to continue or which story could the chatbot tell the user

4. Speech & Text - Transcribing and narration services are handled externally by using Googles Speech to Text and Text to Speech services via their API. For the transcribing service and audio or text data needs to be provided by the when calling the service. Audio data is encoded through a buffer to meet api the standard and is responded with its corresponding text whereas text data is provided with the required audio encoding and is provided with the corresponding audio.



Back-End Schema

4.4 Database Schema

The database configuration and setup is done in PostgreSQL using the knex package in the knexfile.js. All database interactions are handled using Data Access Object files with each table having its own DAO.

To handle the privacy of the user, no personal information of the user is stored in our database. The session information is also removed once the data has served its purpose, hence maintaining no history of the users interactions with the application

Story Table - It contains the various stories that are currently present within the chatbot. Each story must be have its respective name, an event key to call the initial section of that story and An initial emotional context for the application decide which story needs to be played depending on the conversation prior to it.

Intent Table - It encompasses all the data pertaining to the story. Each intent includes the intent ID and the intent's name to tolerate changes made to the chatbot to a certain degree, the suggestions that can be used by the user pertaining to the current chatbot response, URLs of images that visualize the current chatbot response and an event key incase of a emotion detected from the user response

Conclusion Table - It include information of conclusion and suggestions to be provided. These conclusion incrporate the name of the conlusion, the event key need to trigger the specific conclusion and the story section that triggers this specific conclusion

Session Table - It stores all the information necessary for the user session to continue smoothly. The session requires a session ID to identify and maintain its connection with dialogflow, the set of user responses needed to identify the emotion being expressed, the emotional sections visited by the user and the number of conclusions the user has been presented.

SESSIONS	
session_id(PK)	string
user_responses	string
sections	string
conclusion_no	integer

STORY	
id (PK)	integer
story_name	string
event_name	string
emotion	string

INTENT	
id	integer
intent_id(PK)	string
intent	string
suggestions	string
image_urls	string
joy	string
sadness	string
fear	string
anger	string
disgust	string

CONCLUSSION	
id(PK)	integer
conclusion_name	string
event_name	string
section_name	string

Database Schema

4.5 External Services

The application leverages external services through API calls for all of its complex functionalities that need to be addressed but require specialized algorithms and large datasets to optimize. These functionalities include:

1. Chatbot Creation and Deployment – The chatbot, Berno is created using Google's Dialogflow ES console. This is where all predicted user replies and their responses for open domain conversations or stories to teach emotional intelligence are incorporated. The chatbot is made accessible via the Dialogflow API which must be accessed with the project key and a session ID to activate the agent and enable a conversation with the chatbot.

2. Transcribing and Narration - For enabling the user to choose between the audio or text mode to interact with the chatbot, we make use of Google's Speech to Text and Text to Speech services via an API to convert the audio clip sent by recording the user's vocal response from the frontend and pass it to the service of meeting their encoding standards to use dialogflow to maintain a sincgle chatbot service. The Narration is requested by the frontend for each specific sentence that require multiple audio clips so that the user is able to select which part of the response he wishes to here again.

3. Semantic Parsing - In this service, the emotion detection of IBM Watson's Natural Language Understanding API which parses the provided text to extract features emotional relevance is used to understand the sentiment of the users response. This is provided through a score based system of 5 emotions: joy, sadness, anger disgust and fear.

4.6 Technologies

A number of frameworks and packages were used for the development of the application and its deployment. These tecnologies greatly simplify the process through the use of preexisting libraries that help reduce redundant code through function calls that drastically reduce the need for manual implementation of exhaustive structures to maintain the application. include:

- Vue.js - Used to display the relevant information in a manner that is effectively understood by the user. This information is obtained through API calls to obtain the relevant data and also provides the means for the user to interact with the data through application to provide the service to the user.

- Node.js with Express.js - Implements REST API through Express.js to listen to requests and setup routes to the various services provided by the server for the smooth operation of the application. These routes leads to services implemented using packages to node.js to manipulate available data or to call on external services.

- PostgreSQL with Knex - Creation and managent of table data in the database is handled using PostgresSQL. Query building for the access of the information present in the database by the backend services provided is performed using knex.

- Heroku - Provides a cloud service to host the server and provides a domain through which user can interact wth the application via the frontend. It also provide storage solutions for storing the database .

5. Value Proposition

Designing for children with special needs was a big challenge. We tried to understand their perception of the world and anticipate their reaction. The sensory perception includes all the information the brain receives from the five senses. A person with autism, for example, is not able make the different between to sensory stimuli. He will assimilate all the external information and won't make any selection. Thus, it is more difficult for him to understand a context and can lead to sensory overload. That's why it was not possible for us to base on regular conversation rules. We needed to integrate in our design process but also in the chatbot unpredictable reaction of children. We needed to anticipate these reactions and program an answer which can guide the user. This is the role of the proposals displayed after 15 seconds of silence. This detail makes the application more inclusive and accessible by children with special needs. Knowing that children with cognitive disabilities can be easily distracted, we tried to improve the engagement of the child through interactive activities and personal choices. However, involving a person using voice and visual support may not be sufficient for children with special needs.

A few of the technical challenges that we encountered are mainly concerning the time and resources available for designing and developing the aforementioned solution. The main difficulties include:

We are unable to predict responses provided by the target group without extensive testing. The users might respond to questions in an unpredicted manner which might throw off the chatbot designed using specific training phrases that were not formulated by professionals

The reactions to the stimulus that must be empathized from the characters are not analysed for validity and appropriateness, only the emotion that can be detected from the actions present in the response. This might lead to a misconstrued notion how to act given a certain emotion

The stories are not fine-tuned to the needs of the target group and may not properly optimized to get the emotional reaction we expect the user to have

An open domain conversation with children having cognitive disabilities is a delicate process which requires tremendous care and attention to details that cannot be imitated easily using a chatbot

The application would only improve as we test the product on more users and diversify our target group by taking into considerations any factors that might bring up the feeling of the conversation being unnatural. By adding more stories into the application, we would be able to address a wider variety of emotions and help our users in building up a toolkit that they can use out in the real world. Designing more activities that are unique and engaging will help retain our user's attention for longer periods of time and have them return to the application on a more regular basis.

Most of the current application solutions to teach emotional intelligence are targeted at the general public or all children from the ages of 6 to 18. Many of them do not emphasise the need to act appropriately, nor provide the user any suggestions for controlling their action in the application, apart for binary choices. One of our most similar competitors is 'Kid in Story Book Maker' which helps guardians make personalized stories for the children to teach them about emotions and social skill, providing the care takers the tools to create the stories.

Unlike other product on the market, our solution helps the child to better empathize with the characters as they are able to influence their actions by putting themselves in the characters shoes. The advantages of our technology are its ability to influence and adapt the story based on the emotion that is detected in the child's responses. After analysing the user response, the chatbot chooses the story path according to it: if the response was not appropriate given the context, it may mean that he did not understand the context or

the emotion well, so the story continues with sections about similar emotions. In this way, the child reinforces his understandings or difficult emotions, and improves his reactions to them. On the other hand, if the child adopts an appropriate reaction to the story context, the chatbot will continue the story with a different emotion. Thus, the aim of this application is to offer a personalized experience and to adapt to the rhythm of each child. Finally, the emotion detection is enhanced by the personalized activities. Indeed, through children's answer or the chosen proposal, the following questions will be different.

We had the opportunity to discuss and ask for advice from Mrs. Anita Durai, teacher at Sunshine School For Differently Abled Children, from Mumbai, India. She helped us understand what the capabilities and limitations of children with cognitive impairments are, and that emotional learning has to be accompanied with information about appropriate behaviour and actions with respect to emotions. After finalizing the product, we have requested her opinion on it.

"Its straightforward interface with a simplistic story really shines through as it is effective for children with short attention spans. The ability to allow children to choose their preference over the speaking and typing makes this application more reachable for those suffering with cognitive impairments." - Mrs. Anita Durai, teacher at Sunshine School For Differently Abled Children

6. Future work

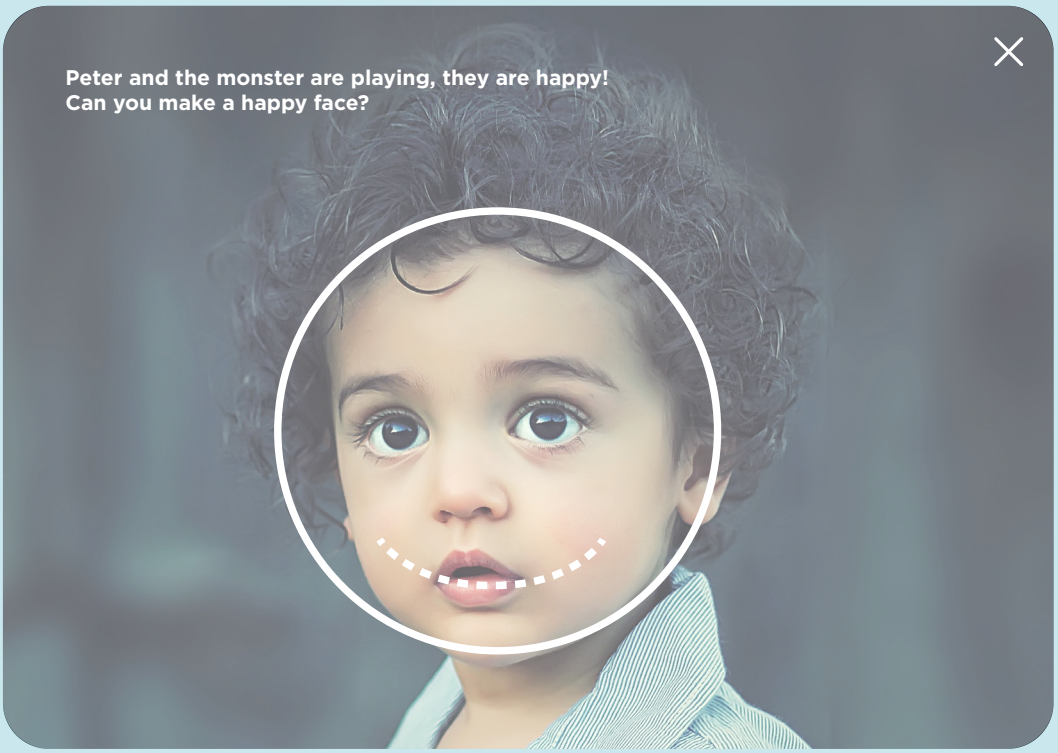
The current application could be improved on a design and technical point of view. In the short term, we could organise empirical evaluation in order to test the design with the child but also the whole structure of the story. Moreover, we could develop additional activities to enrich the story. In order to increase the engagement of the child, the activity could be more diverse and immersive. The education of children with special needs organised following a specific schema. The first level of emotion recognition is to be able to identify it on a drawing (a simplified representation), then on a photo (a realistic representation). The third level is to be able to recognize other people's feeling based on a video, which corresponds to real-life situation. Finally, the child needs to imitate the emotion to validate his learning of the emotion. In a future work, we could apply these four levels to the interactive activities of the story. Because children learn better by doing than listening, we would like to develop an activity able to validate an imitation or disapproved it. During the story, the camera of the device will switch on. The child, guided by the drawing of a happy, will have to smile in order to reproduce this emotion. The advantage of this activity is that making the child smile will automatically have a positive impact on child's mood.

Regarding the technical aspect of the project, the evaluation of user responses can be done not only using emotion detection from text, but also its intensity to reduce the number of false positive results applicable. Emotion detection could be made more effective by using facial recognition and speech analysis, so we could better determine the user's true feelings to personalize the application's experience. During a social interaction, the intonation, speed and tone of the voice give precious information to the listener. Emotion using face recognition would allow us to measure listener's response in real-time and throughout the story. Facial expression and body language are spontaneous and can reveal important information about child's feelings. For instance, blinking can inform us about the attention of the listener. Infact the decrease of number of blinking means that the person is focused and don't want to lose any information.

In the long term, we would like to give a role to the therapist. The expert could reinforce the personal experience of the service, by selecting or changing stories according to his or her patient's needs. More stories and better open domain conversations can be added and optimized to better suit the needs of our target group, which can be made to affect the experience of the application to better personalize and improve the emotional education of the child. This can be done through extracting information such as the emotion a child finds hardest to identify or respond to and provide more instances for the child to improve upon. The addition of different voices for different characters can also increase engagement of the child towards the stories and help immerse themselves into its world. This can make it easier for the children to empathize with characters.

An authentication system to track progress and provide timely reminders to regularly use the app can also ensure that users make steady progress in their education with recaps to help them better retain these suggestions so that they can use them when they get the chance.

Future activity :
Imitating an emotion



Future activity :
Drawing an emotion

7. Bibliography

Marine Cesmat - L'anomalie de discrimination des émotions du visage sous-tendrait le déficit social caractéristique de l'autiste. (The facial emotion discrimination anomaly is believed to underlie the social deficit characteristic of the autistic) Médecine humaine et pathologie - HAL (2014) - 135 pages - ffdumas-01499508f

Robert D. Friedberg - Storytelling and cognitive therapy with children - Journal of cognitive psychotherapy (1994) - 11 pages - DOI : 10.1891/0889-8391.8.3.209

Cathy Pearl, Head of Conversation Design Outreach at Google - Designing Voice User Interfaces (2018)

Ruba AL, Pollak SD (2020) Children's emotion inferences from masked faces: Implications for social interactions during COVID19. PLoS ONE 15(12): e0243708

Kuan Tian Ying - Personalised Avatar on Social Stories and Digital Storytelling: Fostering Positive Behavioural Skills for Children with Autism Spectrum Disorder (2016)

Sandra Costa, Alberto Brunete, Byung-Chull, Nikolaos Mavridis - Emotional Storytelling using virtual and Robotic Agents (2016)

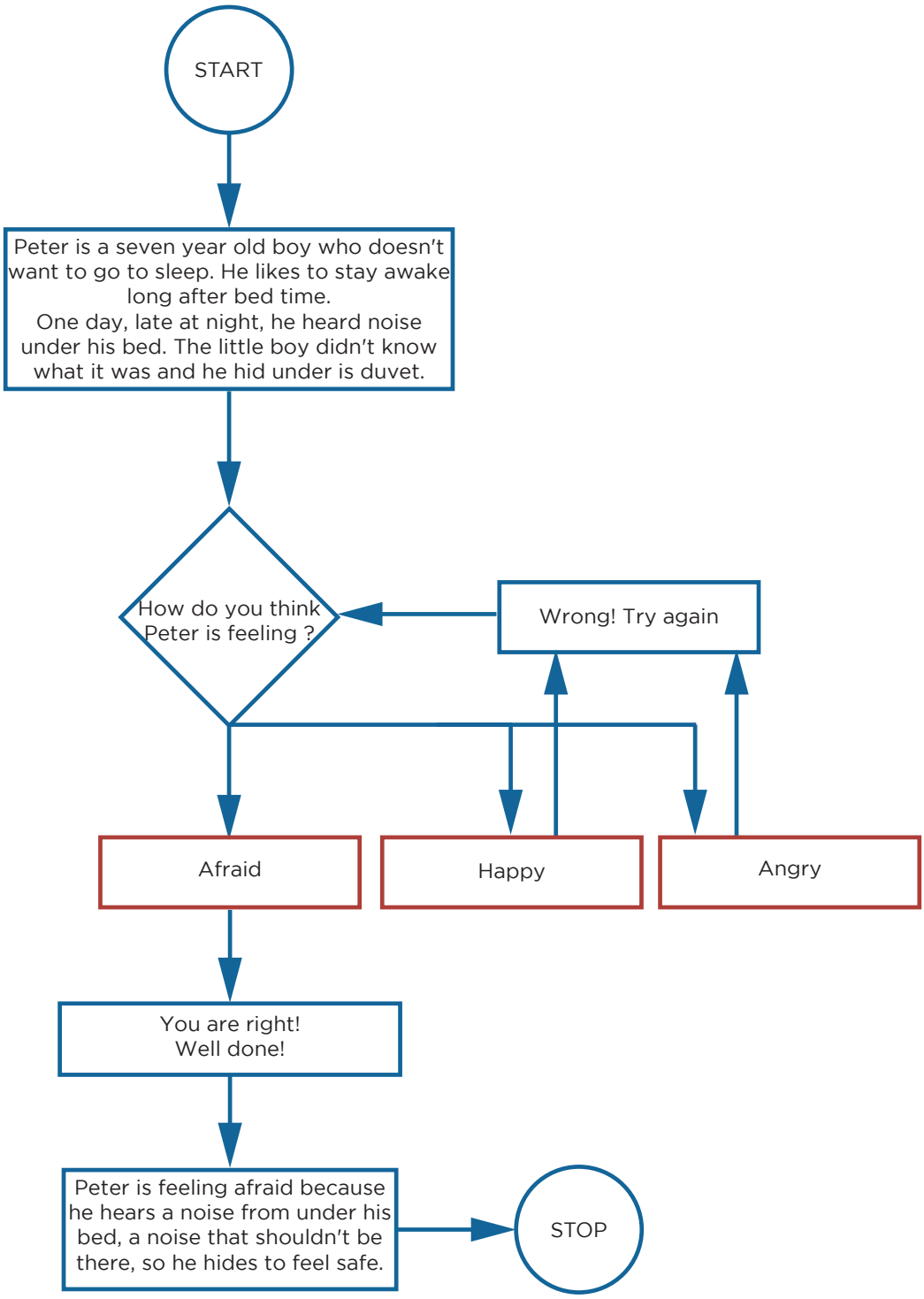
Kiran Ijaz, Anton Bogdanovych, Simeon Simo, Enhancing the Believability of Embodied Conversational Agents through Environment, Self and Interaction Awareness, Thirty-Fourth Australasian Computer Science Conference, (2011)

Sangil Vadivel, V.M. Praseetha, Deep Learning Models for Speech Emotion Recognition, Journal of Computer Science (2018), DOI: 10.3844

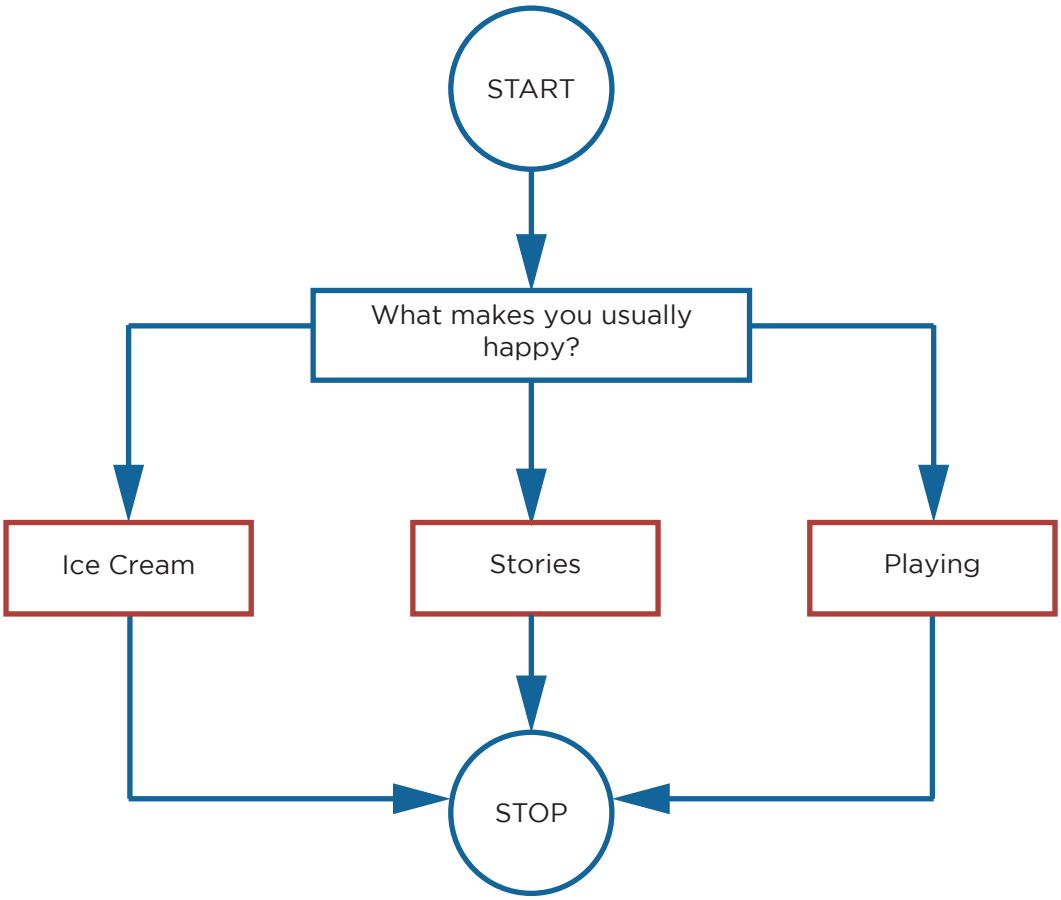
Sandeep Pandey, H S Shekhawat, S R M Prasanna, Speech Emotion Recognition Using Deep Learning Techniques: A Review, 29th International Conference Radioelektronika, (2019), DOI: 10.1109

8. Annexes

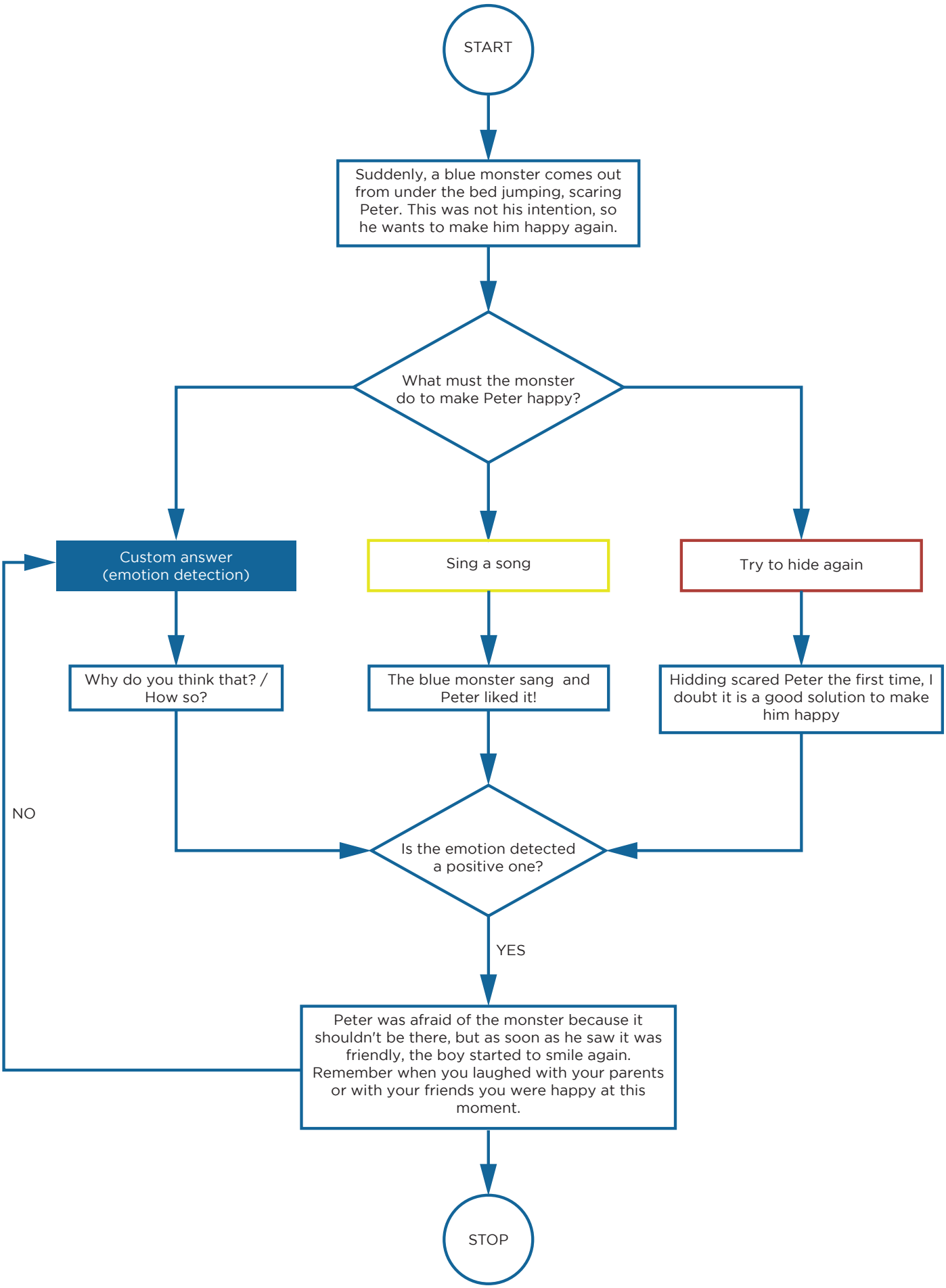
URL: «Berno - the emotional storyteller application is available for testing at the following link: https://emotionalstoryteller.herokuapp.com/



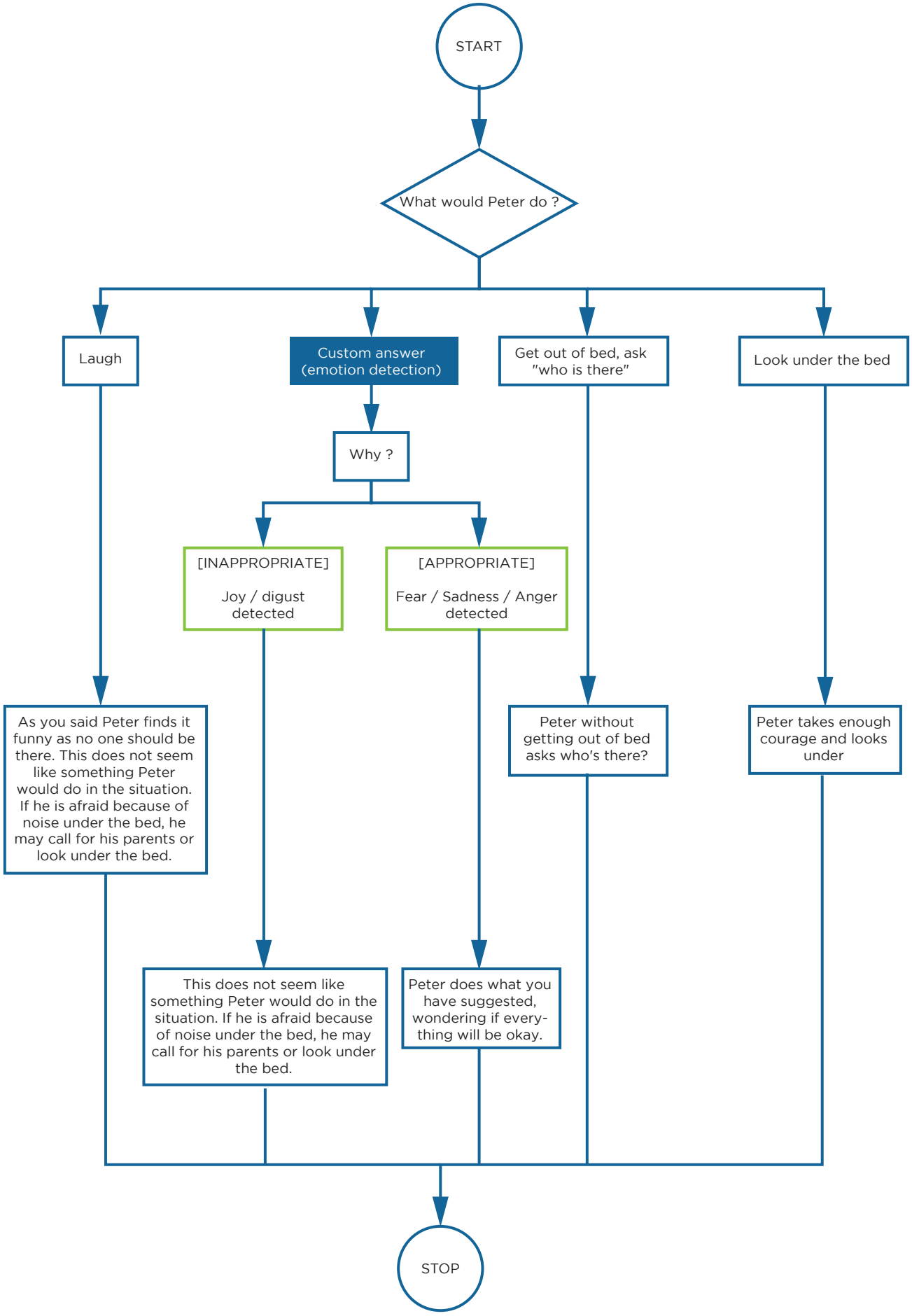
Annex 1 - Emotion Identifying Activity Example



Annex 2 - Self Reflective Question Example



Annex 3 - Reaction Suggestion Activity Example



Annex 4 - Emotion-Causing Actions Activity Example