POLITECNICO
MILANO 1863

SCUOLA DI INGEGNERIA INDUSTRIALE
E DELL'INFORMAZIONE

amazon alexa

CoTra: an Amazon Alexa Skill performing Cognitive Training for Stroke Patients

E-HEALTH METHODS & APPLICATIONS

Authors: Giulia Carpani, Stefano Vannoni, Orith Halfon, Filippo Castellani, Elisabetta Marini

Professors:	S. Ferrante, G.E. Caiani
Professor Assistant:	E. Tauro
Academic Year:	2022-23

Project Repo: GitHub

1. Introduction

Contextualizing our project.

In the next years, the elderly population, defined as the portion of the population aged 65 and over, will be growing worldwide. According to data from the WHO [1], the global population of older people is projected to more than double by 2050, reaching nearly 2.1 billion. This represents an increase of over 3% per year and is due in part to advances in healthcare and improved living standards, which have led to longer life expectancy.

A new kind of healthcare could be the solution for the upcoming issues that our society will face. A patient-centered approach is needed and the technologies available nowadays can play a crucial role.

Possibilities and limitations.

Alexa could be one of them. Like other intelligent personal assistants, it could be used for medical purposes, such as providing health information, reminding users to take medication, or even assisting with medical diagnoses. However, it is important to note that these technologies are still in the development stage and have not been widely adopted for medical use. It is also important for users to understand the limitations of these technologies and to consult with a healthcare professional for medical advice and treatment. These technologies could be used to implement brain games and other mental exercises to help users improve their memory, attention, and other cognitive skills.

As it is reported on the official website: *"Alexa provides a suite of tools that makes it easier and convenient for you to develop experience that will help your customers manage their health and wellness"* [2]. By taking advantage of these characteristics, we developed a **skill that performs a cognitive training for stroke patients**:

CoTra – COgnitive TRAining

2. Methods and Materials

For the realization of the skill, we made use of the Alexa Skills Kit (ASK) which is integrated with several Amazon Web Services (AWS), including DynamoDB, S3, and Lambda [3]. Along with this fundamental framework we used the Draw.io software for sketching out our Voice Interaction Model.

3. General outline of the process

The way we approached the project can be subdivided in three main phases:

1. Conduct scientific research: understanding scope and purpose for our skill.
2. Project and design the application.
3. Realization of the prototype.

3.1 Planning & Scientific research

In literature we were able to find dozens of cognitive tests used daily in clinical settings to evaluate the cognitive abilities of individuals with brain injuries or disorders.

Those can provide valuable information on a person's cognitive abilities and areas for improvement. Those are not simple exercises. In the vast majority of cases there is the need for a specialized therapist to conduct such tests, but computer-assisted cognitive rehabilitation is possible, and it has proven to significantly improve the global cognition of patients on large scale studies [4].

Feasibility assessment.

Now we had to pick one of such trainings and despite the apparent trivialness, we conducted a proper study of feasibility on the ones[1] that were reported here [5],which proved to be very effective for the recovery of stroke patients.

Finally, we have chosen to implement the Paced Auditory Serial Addition Test (PASAT).

3.2 Voice User Interface & User Experience Design (SSML, Internationalization, Reminders)

(Please, while reading, be welcome to take a look to the flow diagram reported down below)

Our main goal was to create a personalized, engaging and effective voice experience for the user.

To do so we took advantage of the tools that ASK provides to developers.

Personalized Experience

Our skill was provided with specific functionalities and expressions in order to make it seem as if it was "tailored" on the user. It "remembers" the name of the user but also

[1]In the article results for the following tests are presented: Paced Auditory Serial Addition Test (PASAT), n-Back Test, Categorization Working Memory Span Test and Test of Language Performance (P-WAB-1).

the past performance over the last training session. In such a way the skill, soon after the very first **experience**, should start **feeling more** like another **"human"** who is coaching the user.

SSML[2]

Having a "human-like" interaction can be particularly important for cognitive training skills, as it can help users feel more comfortable and engaged with the skill. An application that uses a friendly and supportive tone of voice, provides personalized feedback and encouragement, can **help users feel more motivated and engaged with the training tasks**.

On the other hand, a cognitive training skill that uses a robotic or monotone voice, and provides generic and impersonal feedback, can be less engaging and effective for users

Reminders

Using reminders to help patients remember to do their cognitive training could **improve adherence to their training routine**. Adherence refers to the extent to which a patient follows their prescribed treatment plan, and it is an important factor in the effectiveness of any medical treatment. Similarly, to what has been done in [5], our skill by default offers the possibility to the user to set a training reminder every three days, hence 2.5 trainings/week.

Internationalization

Internationalization, or "i18n" for short, is the process of designing and developing a product, such as a skill, in a way that allows it to be easily adapted and localized for different languages, cultures, and markets. Our skill was carefully designed also for this purpose. As a matter of fact, our cognitive training exercise does not need major adjustments to be translated to other languages, indeed it can be directly **proposed to a wide range of individuals**. For this purpose, our skill has been realized and tested both in English and in Italian.

3.3 Realizing the prototype

While the ASK Skill Developer Kit is a powerful and versatile toolkit, for some developers it may be challenging to use. In order to be used proficiently it needs time to master and its usability really depends on the experience and background of the user. Creating CoTra was no easy task, but we are confident that our project stands on solid basis and could be further expanded to new trainings. Our exercise implementation is going to be described in detail.

Before that, for completeness is worth saying that as a final step we decided to make use of DynamoDB Storage Service to keep track of user performances over past interaction with the skill. Some of the key benefits of using DynamoDB include high scalability, low latency and high availability (which are very important aspects when it comes to the possibility of publishing a skill).

4. PASAT

PASAT in Clinical Practice

The PASAT test involves presenting the person with a series of numbers that they need to sum. In clinical settings the numbers are presented orally, and the procedure is performed at a specific pace, usually at a fixed rate of one number per second. The person is then asked to provide the sum of each set of numbers, and their response is recorded and evaluated for a score.

The first time two random numbers are given to the patient, he shall report the sum of them and remember the second number. At each new step, the patient will sum it with the newly provided number.

PASAT using Alexa

We divided our PASAT in three different intents. `1) Say Two Numbers Intent`: after a brief explanation that briefly recalls the tutorial, Alexa will generate, only at the first round, two random numbers and will save the second one in a *session attribute*; from the second round on, it will communicate to the patient the correctness or incorrectness of the previous sum, and the new random number. In all rounds, this intent will make the right sum and saves it in a *session attribute* called `Sum`.

`2) Register A Number Intent`: this intent is triggered when the number is said by the user. Specifically, the Built-In AMAZON.Number slot has been used for this purpose; in the English language, the control over the fact that the slot was actually a number was done automatically. In Italian, instead, we performed this control through an `if-else` block due to a different behavior between Italian and English of the AMAZON.Number slot (likely an imperfection of the

[2] Speech Synthesis Markup Language (SSML) is a markup language used to annotate and control the speech generated by text-to-speech (TTS) systems.

NLP engine). Indeed, the input could be a string because the patient may be in a non-controlled environment in which he could be distracted by third parties.

Afterwards, the intent performs a check on the equality of the sum pronounced by the patient, saved in *session attribute* as `Special Number`, and the real one, already managed by the previous intent. With respect to the result of this control, the score of the patient is increased by one point just in case of success and then saved in a *session attribute* as `Correct Counter`. In addition to the score, `Question Counter` is saved as a *session attribute* to store the number of steps the patient goes through, up to 20 steps.

3) `Check Sum Intent`: this is triggered at every step, and it controls the value of `Question Counter`. If it is less than 20 it returns to point (1), which generates a single new random number. When `Question Counter` reaches 20, Alexa ends the training reporting the score and asks for an eventual reminder to be set.

It is important to note that every time the patient performs a step of the training, counters are persistently saved on an external database. This aspect is relevant since there is the possibility that the user stops autonomously at a certain point of the training, in this way there is always trace of the performed steps. Next time the training will start from the beginning with the counters set to zero, but Alexa will motivate the patient to reach the end of the session.

On the contrary, if the patient reaches the end of the exercise, the following time Alexa, reading the score of the last session, will encourage the user to get a better score, if it is not maximum, or will congratulate in the case of a perfect score. All these aspects, concerning the motivation of the patient by Alexa at the very beginning of every training, are managed by the Motivational Intent.

In addition, Alexa is set to be very close to the patient and, any time the user needs some "Help", Alexa will provide all the instructions to make a good training.

Limitations

Unfortunately using Alexa Natural Language Processing does not provide the possibility of pacing the conversation as it should happen in a clinical environment.

A solution for this problem might be easier than expected if in a future update Alexa will implement some functionalities to regulate the "expected-answer-time windows".

5. Conclusions

Our skill can proficiently execute the following tasks:

- Implement a cognitive training and keep track of results (saving data)
- Interact with the user in a friendly and easy way (SSML)
- Set reminders to improve adherence.
- Potentially interact with a wide variety of populations thanks to the internationalization.

6. Final considerations

Developing an Alexa skill can be a challenging but rewarding process, and we are sure our skill could be a valuable new instrument into the Alexa ecosystem once published.

Thinking about the future…

There are several ways in which our cognitive training skill could be evaluated for real-world applications.

User feedback, performance metrics, and clinical trials are some of the key methods for evaluating the effectiveness and usefulness of any new technology and can provide valuable insights into the skill's performance.

By using these methods, it would be possible to evaluate the skill's effectiveness and usefulness. Furthermore, this would allow our team to take informed decisions on how to improve it and refine it.

Possibly it would be interesting to investigate the possibility of using the skill in a clinical environment, helping the therapists to perform the training.

Is always worth noticing that this procedure must involve the therapists themselves to ensure acceptance in the clinical pathways.

7. Flow Diagram

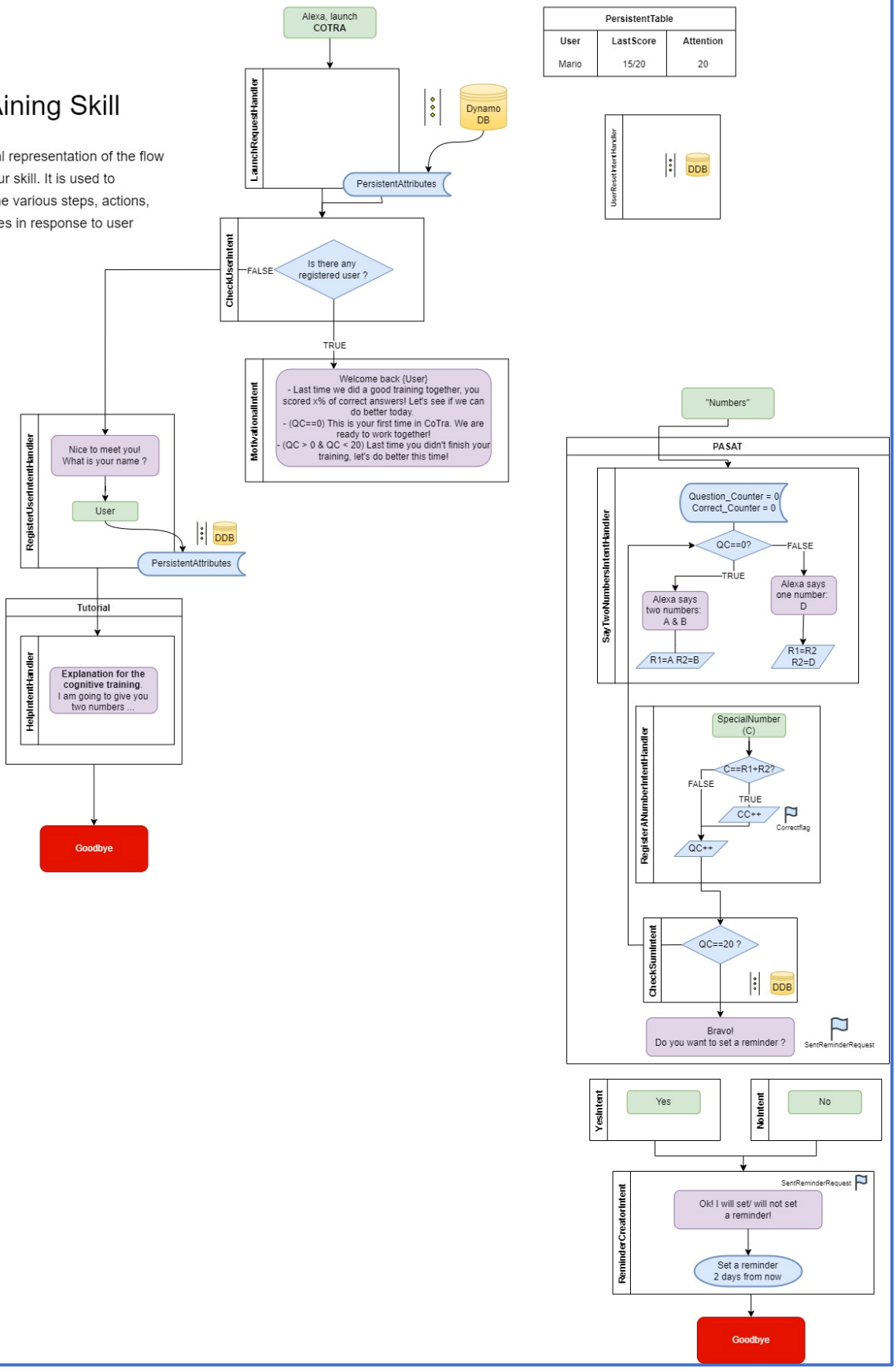
CoTra

Cognitive TRAining Skill

This Flow Diagram is a visual representation of the flow of interactions and actions our skill. It is used to understand and document the various steps, actions, and decisions that a skill takes in response to user input and requests.

LEGEND:
Color code is used to represent the actors in the conversation.

- Green = User
- Light Violet = Alexa
- Light Blue = actions taken by the Lambda functions that do not result in a specific sentence pronounced by Alexa.



Bibliography

[1] World Health Organization:

Ageing and Health

https://www.who.int/news-room/fact-sheets/detail/ageing-and-health

[2] Alexa Developer:

Alexa Health and Wellness Skills

https://developer.amazon.com/en-US/alexa/alexa-skills-kit/get-deeper/custom-skills/healthcare-skills

[3] Alexa Developer Console:

Amazon Alexa

https://developer.amazon.com/it-IT/alexa

[4] P. Nie *et al.*,

«The effects of computer-assisted cognitive rehabilitation on cognitive impairment after stroke: A systematic review and meta-analysis»

Journal of Clinical Nursing, vol. 31, fasc. 9–10, pp. 1136–1148, 2022, doi: 10.1111/jocn.16030.

[5] M. Nikravesh *et al.*,

«Working memory training in post-stroke aphasia: Near and far transfer effects»

Journal of Communication Disorders, vol. 89, p. 106077, gen. 2021, doi: 10.1016/j.jcomdis.2020.106077.