Co-funded by the European Union

Specialized and updated training on supporting advance technologies for early childhood education and care professionals and graduates

MODULE VII.2

Early care and application of smart resources: Intelligent personal assistants

Lecturer:

Dr. D. Raúl Marticorena Sánchez
Área de Lenguajes y Sistemas Informáticos
Departamento de Ingeniería Informática
UNIVERSIDAD DE BURGOS

e-EarlyCare-T

“Specialized and updated training on supporting advance technologies for early childhood education and care professionals and graduates”, eEarlyCare-T, reference 2021-1-ES01-KA220-SCH-000032661, is co-financed by the European Union's Erasmus+ programme, line KA220 Strategic Partnerships Scholar associations. The content of the publication is the sole responsibility of the authors. Neither the European Commission nor the Spanish Service for the Internationalization of Education (SEPIE) is responsible for the use that may be made of the information disseminated herein”.

Specialized and updated training on supporting advance technologies for early childhood education and care professionals and graduates

Índice de Contenidos

I. Introduction

This module introduces the concept and use of bots or intelligent personal assistants. Given the technological advances in artificial intelligence, networks and cloud computing, the use of conversational assistants that simulate people, helping us in everyday tasks such as managing the agenda, shopping, etc., is emerging.

Specifically, the fundamental concepts of bots will be reviewed, as well as their more particular applications in healthcare, finally pointing out their future applications in early care.

II. Objectives

The basic objectives of the following sub-module are as follows:

- Introduce its history and the concept of bot or intelligent personal assistant.
- Explain the different characteristics of bots.
- To detail the aspects or components of a bot and current solutions.
- To review their possible practical applications in healthcare.

III. Specific contents of the subject

3.1. Bots or intelligent personal assistants

Nowadays, the use of **bots** or **Intelligent Personal Assistants** (**IPA**) is becoming **widespread** in **all areas**, providing multiple types of services, from resolving doubts, searches, recommendation services, agenda management, booking and purchasing tickets, etc.

3.1.1. Foundations and historical precedents

Its theoretical foundation is based on the **Turing Test**, developed by Alan Turing in the 1950's. In simple terms, this test proposes that the computer must show "intelligent" behavior, such that it could deceive another human interlocutor in a conversation by pretending to be another human being. Once this is achieved, the Turing Test would be passed.

In 1966, Joseph Weizenbaum developed ELIZA for the IBM 7094. It was a psychotherapy bot that dealt with patients about their problems, generating strong emotional reactions, even though they were aware that they were dealing with a bot.

The program analyzed keywords and based its response on them, and was probably the first program to pass the Turing Test.

A later step in this branch was the appearance of PARRY, a conversational agent simulating a paranoid patient with schizophrenia (Colby, 1975). These are the first examples of valid application to health issues, surpassing the Turing Test.

The proliferation and development of conversational agents receives a boost that is reflected in the creation of **competitions**. The Loebner Award is established as an annual platform for conversational bot competition. ALICE (Artificial Linguistic Internet Computer Entity) gains great attention in 1995, winning the award three times (2000, 2001 and 2004).

In 2010, Siri emerged as Apple's **commercial** solution integrated into its mobile devices, followed by Google Now in 2012, Alexa (Amazon) in 2014 and Cortana (Microsoft) in 2014. In 2016, a final hatching occurs with the integration of bots in social networks such as Facebook with its own messaging platform, which would be joined by API.ai, LinkedIn, Viber, etc. developments.

In more recent times, Amazon created its own competition - "Alexa Prize" in 2017 - with an approach similar to the Turing Test. The challenge is to create a "social bot" that converses coherently and engages a human, conversing on a topic for 20 minutes. The current benchmark is four-time Loebner Prize winner Mitsuku (Prize L, 2019).

In recent times, the application of these conversational bots in the health field has gained a lot of weight with situations such as the COVID-19 pandemic. As an example, in 2020 WhatsApp agreed with the World Health Organization (WHO) to complete a chatbot service to answer questions related to COVID-19. Although not exempt this type of solutions from risks and malicious uses, which have forced organizations such as UNICEF to define guidelines for good use and implementation (UNICEF, 2022a; UNICEF 2022b).

3.2. Definitions

A bot exposes software services through a conversational interface. These bots can be referred to as chatbots, conversational agents, conversational interfaces, intelligent personal assistants, and in many other ways, depending also on the interface used with the user (Shevat, 2017).

The concept of "**conversation**" is vital. Often in these bots, conversations are considered a single interaction (question/answer) (e.g. setting an alarm, checking the temperature or weather forecast, making a call, etc.) when in the real world, humans understand conversation to be a longer exchange of questions/answers that are related

to each other. Although bots can also have long conversations on one topic, with several question/answer exchanges.

In the following Figure 1**¡Error! No se encuentra el origen de la referencia.**, an example of a simple conversation with a text-based chatbot, usually referred to as chatbots, is shown as an example of the expected operation. On the left are the automatically generated interactions of the software. On the right are the questions asked by a person. Without precise information, it should be indistinguishable whether the bot messages are generated by a program or by another person.

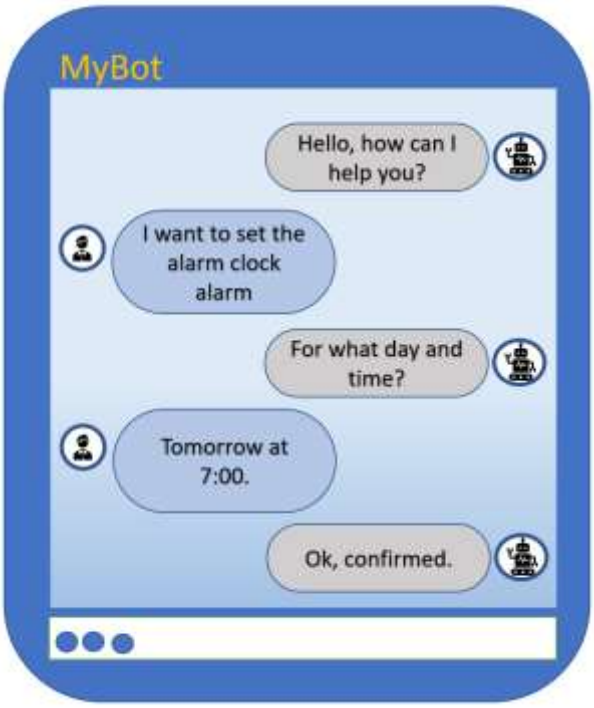

Figure 1. Example of interaction with a bot (Image: own elaboration)

The bot must be distinguished from the service it provides. The bot is only a means or interface to the service. Although this interface can be more or less complex.

Advantages of its use:

- Increased user engagement (loyalty) by making it easier for users to perform their tasks.
- Ease of use vs. web interfaces and mobile apps.

Disadvantages of its use:

- They are not the solution to every type of problem posed (at present).
- Fear of some risk of loss of privacy.
- With voice-based interfaces, there is a certain amount of social shaming.

3.2.1. Most common use cases

In practice, the most common successful use cases are:

- Productivity and training (coaching): focused on remembering tasks to perform, and management of personal or group tasks to complete. Also to help to follow diets, manage expenses, perform sports activities, etc.
- Alerts and notifications: replacing the use of email and notification apps. They can work with groups in a more productive way.
- Router to humans: finally redirects to a human interlocutor, but assigning the best person for the resolution, by means of a guided conversation.
- Customer services and answers to frequently asked questions: in support of the most common and recurring questions.
- Third-party integration: to integrate third-party services into the current product.
- Games and entertainment: with the basic objective to entertain and amuse.

Starting from these cases, we will identify different typologies.

3.3. Typology

By target:

- **Personal bot / private *bot*:** serves as a personal assistant, in a one-to-one conversation (e.g. setting an appointment in my personal calendar).
- **Team *bot*:** assists a group of people to achieve an objective (e.g. setting a meeting date and time for a group).

In practice, the personal model is more widespread for simplicity, and even in home bots, with devices such as Amazon Echo or Google Home, all users in the same household are treated without distinction, as if they were the same person.

By scope:

- **Domain-specific:** exposes a single service (product, brand or target).
- **Super bot**: exposes multiple services at the same time.

In this last category we find solutions such as Google Assistant, Amazon Alexa, Apple Siri, etc. that encapsulate several services. They basically group different functionalities in a modular way, so that they can be augmented transparently to the user. These modular functionalities are often referred to as skills.

By aim:

- **Business**: facilitate a business task or process. The goal is to solve an objective. Task and workflow oriented.
- **Consumer**: entertain while facilitating a commercial interaction. Aimed at a better and entertaining user experience.

By access:

- **Text**: the conversation is based on text entered by keyboard and display of the response on screen. They are often referred to as chatbots (e.g. with web platforms that include them such as Slack, Facebook Telegram, WhatsApp, WeChat, etc.).
- **Voice**: the conversation is based on the use of audio to ask and return the answer without the need to physically interact with the devices (e.g. Amazon Alexa, Microsoft Cortana, Apple Siri, Google Assistat as the de facto standard).
- **Multimodal**: combine both elements discretely - text or voice - and may additionally require touch interaction on screens or the combined use of other devices or artifacts (e.g. cameras, watches, devices, wearables, etc.).

In this case, the access modes are not exclusive, and you can have a bot that can be accessed in all of the ways mentioned above.

By integration:

- **Legacy system**: service existing software systems by offering new ways of interacting with pre-existing services..
- **New *bots***: interfaces to new services or products created from scratch.

3.3.1. Evaluation and selection criteria

When choosing an applied technological solution, certain characteristics must be taken into account:

- Target audience.
- Business vs. consumer.
- Form of interaction (text vs. voice vs. multimodal).
- Devices required to interact.
- Associated costs of software hosting and hardware purchase.

Depending on the previous criteria, with more social or economic issues in mind, not all bot solutions are valid.

3.4. Generic aspects

You should be aware that behind a bot there is basically a **software** running on hardware platforms that host such software.

Although in each technological platform the **vocabulary used** may vary, a summary of the most common vocabulary with the basic concepts in a bot is presented in Table 1 below (Note: the vocabulary of DialogFlow and Alexa are used as references).

Table 1. Basic vocabulary in the construction of a bot.

Element	Definition	Use
Intent	Objective (or intention) that a customer has when asking a question..	Intents are defined as an agglutinator of the different actions. When a question is asked, the **NLU** system searches for the closest intent.
Utterance	Literal phrase entered by the user.	A set of alternative sentences that are equivalent to resolve an intent are defined. The **NLU** system interprets and resolves the intent match.
Entity	Type of data that can be extracted from the user's message or utterance.	They are used as variables that can be defined and take different values, in order to perform customized actions based on those values.
Context	Similar to a context in a real conversation, defining variables that determine the evolution or path of the conversation.	They are used to define and establish more advanced conversations, where there may be different paths in the conversation.
Fallback	Default intent when the input has not been recognized.	In the event that the chatbot is unable to recognize the user's input, a typical response action should be set. Ideally the number of times a fallback is executed will be reduced as the chatbot is trained and improved.
Event	They trigger the execution of an intent automatically without requiring user input.	It allows the automation of actions, such as launching an initial question when the user enters the website containing the chatbot.

In Figure 2, we can see how training sentences would be defined for an intent to communicate the due dates of an academic assignment:

Figure 2. Definition of questions for an intent in DialogFlow

Once these training sentences have been set, the possible set of response sentences must also be set. In some cases, several equivalent responses are given, so that one is chosen randomly, giving the bot more variety. Or, as in the example in Figure 3, use more complex formats (i.e. JSON notation) to generate rich media content.

Figure 3. Enriched response for an intent in DialogFlow

As a result of the query to the bot, Figure 4 would show a screen presentation like the one shown below, where in addition to showing a plain text, a multimedia element is added, with a button as a hyperlink, which, once clicked, would open the content in a web browser:

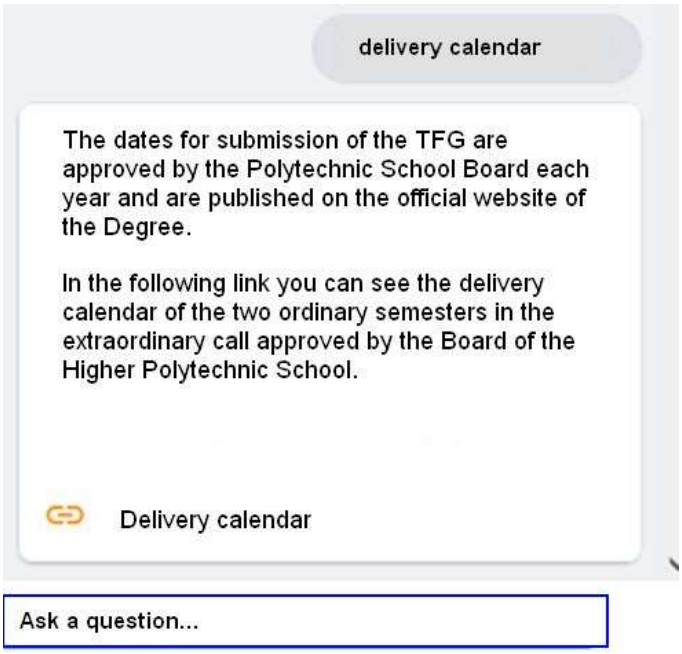

Figure 4. Chatbot display with rich content

Having presented the vocabulary and a basic example, other features are detailed:

- **User interaction style**: personality or tone with which it wants to communicate. The bot, as with people, must have its own recognisable personality, coherent over time.
- **Artificial intelligence**: fundamental basis of a bot that includes several elements.
    - Natural Language Processing (**NLP**): processes the sentence by analysing its structure.
    - Natural Language Understanding (**NLU**): extracts the meaning of the sentence, once its structure has been previously analysed.
    - Text To Speech (**TTS**): converts a text file to its corresponding sound.
    - Speech To Text (**STT**): it translates sound by generating the corresponding text.
    - Prediction models: models based on artificial intelligence techniques to predict the most appropriate response given a sentence. It associates the user's sentence (utterance) with the intention and the corresponding response.
    - Image recognition in multimodal systems: the analysis of the image and its recognition can make it possible to know what is in front of the camera (person) and what their state of mind is.
    - Conversation management: the system must manage the flow of questions and answers in an appropriate way, remembering the previous context, in order to reach the end of the conversation successfully.

    - Sentiment analysis: phrases can implicitly carry connotations that can vary their interpretation and the flow of conversation. For this purpose, sentiment analysis techniques are applied, weighing the positive, negative or neutral components of the sentence, affecting the subsequent response.
- **Conversation management:** how the conversation is managed, from the beginning to the achievement of the set objective.
    - Onboarding: how access to the bot is prepared, informing of its objectives and purpose, modes of interaction, functionality provided and how to get help.
    - Functional scripting: management of conversation flows or stories, trying to mitigate failure (loss of the optimal path or "happy path" in the conversation).
    - Feedback and error handling: enabling continuous improvement and a good user experience.
    - Help and support: in case of loss of the optimal conversation flow or "happy path".
- **Enriched interactions**: in bot using text-based interfaces, responses can be enriched with multimedia or web elements to help simplify the cognitive load of the conversation.
    - Files: uploading or downloading of files (e.g. work documents, shopping lists, etc.)
    - Audio
    - Video
    - Images, maps o charts.
    - Buttons: allows you to pre-set answers, speeding up the conversation.
    - Templates (platform-dependent)
    - Links
    - Text formatted with colours, styles, etc.
    - Emojis or reactions
    - Persistent menus
    - Typing indicators: to pretend that the bot is typing as if you were a human.
    - Commands: shortcuts with short commands to invoke actions in a similar way to text consoles.
    - Webviews: open a mini-view of a portion of a web page.
- **Context and memory**: remembering the previous context of the conversation, or recent conversations, which can completely change the flow of the current conversation. Humans do this naturally, but in the case of bots it is one of the most complex issues.
- **Discovery and installation**: making it easy to locate, download, install, and facilitate initial interaction.
- **Engagement methods:** methods to get the user to become a regular customer of the bot.

  - Notification of new content.
  - Guiding and assisting invocation: providing easy methods of initiating the conversation by teaching the user the process.
  - Subscription: allowing the user to define their interests, filter what they want, etc.
- **Monetisation:** generating revenue or other benefits from the use of the bot either directly or indirectly.

3.4.1. Conversation management: onboarding

A first issue is to prepare the user's **first contact** with the bot well, as a bad experience may mean that the user will not use it again.

The recommendations or heuristics to be followed are:

- Declarar el propósito y contexto de la conversación: haciéndolo claro al usuario o grupo de usuarios. Presentar e introducir claramente qué hace el *bot*.
- State the purpose and context of the conversation: making it clear to the user or group of users. Introduce and introduce clearly what the bot does.
- Teach the user how to use the bot: explain how to interact with the bot, such as whether there is a way to activate it (e.g. wake word with voice-based interfaces), main functions, keywords or commands, etc.
- Configuration: request additional information from the user if required for the operation of the bot (e.g. preferences or customisation for the user).
- Prompt the user to get value from the bot: encourage or direct the first questions to start the conversation. Even providing an example of a simulated conversation.
- Establish the tone and personality of the bot: for consistency, the bot should maintain the same tone throughout the conversation. Ideally, it should have a "personality".
- Make the bot's entry into group conversations explicit: similar to introducing a new staff member, the bot should be introduced to everyone in an appropriate way. The bot should introduce itself to all members, announcing that it has joined.

3.4.2. Functional scripting

The conversation flow can be basically **task- or topic-driven**. In the first case, the aim is to find the optimal set of conversational interactions to find the goal (e.g. to book a ticket) or to complete a precise **task**.

A common way is to model the set of states and transitions that occur in the conversation. These are also called command-and-control systems.

Issues to consider in this type of bots are:

- In the responses, the possible set of responses can be closed in order to limit the conversation and achieve the optimal or **"happy path"**.
- If at some point the user does not give an expected input, a **divergent flow** is produced and **error handling** will have to be carried out.
- To extract information that may come in disorder, **entity extraction** is carried out. It consists of extracting elements (entities) from the question that have their own semantics to resolve the conversation flow.
- **Intent mapping and conversational control**: association between the different user inputs and the actions/responses to be given. While in classic graphical interfaces the user is given a closed set of menus, buttons, etc., in a bot it is offered in a different way. Either implicitly in the question itself or by closing a possible set of valid commands in the answer.
    - The possible actions that a user may require as a result of one or a set of possible inputs are called **intents**.
- **Shorthand**: if a context and memory of previous actions is maintained, questions can be shortened, using knowledge of the previous state.
- **Stories/flows**: allow branching and grouping of more common or repeating conversations (e.g. "user wants an unavailable product").
- **Conversation funnels**: as a user asks a question, the possible path of the conversation narrows. It is convenient to define these funnels to reduce the possible set of paths.

In **topic-driven flows**, the flow is less directed and even "circular" in character. Different aspects of a topic are talked about and discussed. They generally include more interactions than a task-driven bot, and their additional goal is to achieve a certain user engagement (e.g. a bot discussing a movie or series).

They are rather more complex to implement given their more abstract characteristics, applying the concepts already seen in task-driven bots (i.e. course correction, entity extraction, etc.).

3.4.3. Entity extraction

In a sentence (utterance), some of the terms used may have a certain special semantics, which can also condition the conversation later on. In the vocabulary of bots, this is called an **entity**, and it is very important that the bot is able to extract and remember them properly.

For example, let's say we ask:

"What is the temperature today in Madrid?"

In addition to answering the question, when analysing the sentence, entities of different nature can be detected, a concept to be solved such as **temperature**, a **date** with value "today" and a **city** or destination with value "Madrid". If subsequently the question were to be asked:

"And the forecast for the next few days?"

In this case the bot, if it has extracted the previous entities, no longer has to ask the user requesting additional information, since the question is no longer ambiguous, given that the initial time frame of reference was "today" and the city "Madrid", and we were asking for the concept of "temperature". Conversely, if the entities had not been extracted and stored in context, the bot would not have accurate information (i.e. it would have no memory of the thread of conversation), and would not know what we are talking about (e.g. forecast of what? on what dates? what place are we talking about? etc.).

Other typical examples are remembering or maintaining entities with requested quantities, favourite colours, ages, names, etc. in such a way that the conversation flows, remembering previous choices already made in the conversation thread.

3.4.4. Context and memory

In practice, many bots stick to the **question/answer** paradigm, for example in a simple FAQ resolution. For each new question, it starts from a new context, not remembering the previous conversation. However, for other types of conversations, it is necessary to remember the previously exchanged messages, which constitute a context.

Context

To apply the context, the intent and the set of entities previously associated in the conversation must be analysed. The associated variables can be global or long-term (e.g. relating to the user and previous conversations) or local to the conversation or short-term (e.g. day and time of booking, colour chosen, etc. in the current conversation).

When we move in the conversation to another intent, local variables may or may not be forgotten, depending on their usefulness, but global variables should not be forgotten.

Further inference of context through the use of pronouns must be resolved through **NLU** and is outside the scope of this module. Another way to infer context is through rich content that helps to capture intents.

Memory

It is a more general concept, relating to remembering intentions (intents) and entities, or even entire conversations from the more distant past, and is still under investigation.

3.4.5. Error handling

Mistakes in conversation are far more common than one would like. Both in real life, as well as what particularly occurs when conversing with a bot.

If you look at the logs of all conversations over time, you can see the variety of possible entries that do not lead to a successful conversation. Be it simple unintelligible texts, meaningless questions or utterances, sentences that are not connected to the previous conversation, even insults and the use of offensive, out-of-place language.

In addition, many of the phrases that should be recognised as valid in the early stages of bot configuration may not have been included to achieve the correct intent. In the face of possible errors and divergences in conversations, different measures can be taken:

- **Path correction**: this consists of redirecting the conversation to the optimal path (happy path), answering that the previous question cannot be resolved, but giving options to redirect the conversation with options that are contemplated. Or by indicating that a note is taken, and that it cannot be resolved now, but it can be resolved in the future.
- **Human intervention**: it is a common solution that when the conversation diverges so much from the path to success, a real person is redirected to continue the conversation and resolve it.
- **Restart the conversation**: this is the easiest solution, but it can be very annoying, generating a bad user experience.
- **Redirect to another bot**: not widely used at the moment, but given the modular architecture with modular applications (skills) of the predominant systems on the market, it will be common for bots to specialise and redirect requests from one to another (even raising the dilemma of the conversation between bots).

Other considerations in error handling are the consistency of responses throughout the conversation, showing a certain "character" or "personality" of the bot, which must be maintained in a coherent manner. In the same way that a human being would respond to the appearance of problems in a conversation, and not by generating numerical code or messages that cannot be interpreted by a person.

This process of error management and correction is not immediate, and must occur gradually, learning from mistakes made in the early stages. By reviewing the intentions that have failed, it will be possible to deduce and improve what goes wrong,

adjusting and refining the set of training sentences associated with an intent. The process of growth and improvement of a bot is continuous, as is the learning process of a person.

3.5. Voice-based assistants

Advances in speech recognition (Speech to Text or **STT**) and text to speech (Text to Speech or **TTS**) over the last two decades have given a definite boost to the incorporation of speech-based interfaces as an additional element of bots. Figure 5 shows a typical early-generation assistant with integrated microphone and speaker, without a display.

Figure 5. Example of a conversational device for the home [Image: Unsplash licence - https://unsplash.com/)

Commercial examples of large companies supporting voice-based assistants that are common in everyday life are Amazon Alexa, Apple Siri, Google Assistant or Microsoft Cortana. However, we must differentiate between the concept of voice assistants (i.e. software generally integrated in the cloud), such as those mentioned above, and their corresponding "smart speakers" (or even more complex and multimodal devices) such as Amazon Echo, Google Home or Harman Kardon Invoke (for Cortana), which are still the physical means of access to the conversational agent (i.e. hardware).

From a practical point of view, the introduction of this voice-based interaction, both for asking and answering through this combination of software and hardware, adds some additional issues to consider.

Initially there are certain **advantages** in its use over the keyboard-based text input solution:

- Faster question issuance.
- Hands-free", leaving the user free to perform other actions while using the bot (and in a safer way).

- Intuitive: interaction with speech is very natural.
- Empathy: the inclusion of tone, volume, intonation and speed of speech add information that helps to better interpret the answer and avoid misunderstandings.

Additionally, the reduction of screens on wearable devices (e.g. phones, watches, bracelets, rings, etc.) invites the use of these interfaces.

On the other hand, they also offer certain disadvantages to consider in their use:

- Use in public spaces: speaking or raising your voice in public spaces can give a strange feeling to the other people around you. The problem is exacerbated if several people do this at the same time.
- Feeling uncomfortable talking to a computer.
- Ingrained habit of typing to interact with devices.
- Privacy: if you want to discuss sensitive issues (e.g. health) or listen in on private matters (e.g. reading messaging), you don't want anyone else to hear your conversation.

With regard to error handling, these interfaces raise additional issues to be addressed such as:

- No speech (i.e. the user does not ask a question for a certain period of time). The conversation can be terminated or the user can be asked again.
- Recognition problems (i.e. the sentence is not recognised even though it has been heard). Can be asked again.
- Intent handling problems (i.e. the sentence is recognised, but there is no appropriate programmed response or the wrong response is given) already present in text-based bots. Their resolution is more complex.

When designing purely voice-based interfaces, some additional issues have to be taken into account. From a conversation initiation point of view, one does not want the device to be continuously listening, for privacy reasons, and therefore there is the concept of a **wake word** that must be spoken by the user to explicitly initiate a conversation by activating the microphone.

On the other hand, responses should **not be excessively long**, as the cognitive load may exceed the user's capacity. In multimodal systems this is additionally solved by relying on the presentation of information on the screen, but this is not always possible and must be considered when giving certain answers.

3.6. Technology solutions for personal assistants

The following is a review of current solutions, both from the four major technology leaders, in terms of the use of **conversational bots** as well as other platforms that allow the integration of bots. We address both text and voice solutions (with integrated microphone and speaker), and multimodal solutions combining images, hyperlinks, etc., provided that the associated device also has a smaller or larger screen.

Alexa Amazon

Based on Amazon's online ecosystem, it is one of the benchmarks in recent years, having one of the largest cloud computing infrastructures behind it. It is an agglutinator of functionalities called skills. By adding more or fewer skills to Alexa, the user obtains more or fewer functionalities and types of conversations. With a focus on home automation and control of devices in the home, it is one of the current benchmarks.

Associated physical devices: Amazon Echo, Amazon Dot or Amazon Echo Show, etc.

Apple Siri

First precedent in conversational agents, integrated into iPhones. However, its more closed and high-cost philosophy, common in Apple products, hinders integration with third parties and its wider diffusion.

Associated physical devices: iPhone.

Google Assistant

It is Google's solution for conversational agents, building on all its previous experience in text-based search engines such as Google. The support of Android devices, which are currently very widespread, gives a strong boost to this solution, also associated with lower costs. It follows a similar skills architecture to Amazon Alexa. Also, with a certain orientation towards home automation and control of devices in the home.

Associated physical devices: Google Nest, Nest Mini, Nest Audio, Nest Hub Max, etc.

Microsoft Cortana

Microsoft product to compete directly with Amazon Alexa and Google Asssistant. Its possible abandonment in the next few years suggests that it will not be better supported in the future. Its fundamental (although not exclusive) link to Microsoft platforms (e.g. Windows 10 on computers) and Microsoft's lack of success on mobiles has meant that it is less widely used. Future integration with more successful Microsoft 365 products such as Outlook or Teams is expected.

*Associated physical device*s: Harman Kardon Invoke and Surface Headphones.

Table 2. Conversational agents shows a summary of the four assistants that currently constitute the de facto standard.

Table 2. Conversational agents

Assistant	***Company***	***Starting date***	***Referring device***	***Wake word***
Alexa	Amazon	November 2014	Echo	"Alexa"
Siri	Apple	October 2011	iPhone	"Siri"
Assistant	Google	May 2016	Nest	"Ok Google" "Hey Google"
Cortana	Microsoft	January 2015	PC Windows 10	"Hey Cortana"

Along with the four big dominators, there are other platforms that allow the development and integration of bots, usually referred to in this branch as chatbots, and in some cases also partially linked to these four companies. Table 3 below lists, but is not exhaustive, some of the best-known examples.

Table 3. Chatbots development platforms

Product	Description	Features
Amazon Lex	Amazon product for the development of chatbots.	It uses the same **NLU** engine as Alexa, but incurs costs after the second year of implementation, due to the use of Amazon's platform.
Chatcompose	Chatbots platform for marketing and support.	Offers a live chat option, allowing the inclusion of human agents in the conversion. Limited number of chatbots in the free version.
Chatfuel	Integration of chatbots.	Only Facebook Messenger and Instagram integration is available.
DialogFlow	Google product for the development of chatbots.	Simplicity in its use and free of charge. Unlimited chatbots and with an intuitive graphical interface for creation.
Microsoft Bot Framewework	Also known as Azure Bot Service. Microsoft solution for the creation	Offers integration into the Microsoft ecosystem with

	and integration of chatbots.	Office and Teams. With certain limitations in the free version.
Rasa	*Open source framework for machine learning and chatbots creation.*	Developed in Python, with a high learning curve, without cloud hosting and in its free version without graphical interface.
Watson Assistant	IBM product for the development of chatbots.	Focus on more complex developments, with more complex conversational models.

3.7. Practical applications in health

Based on the study by (Car et al., 2020), the objective of its application is **to improve accessibility, personalisation and efficiency in patient care through bots.** Their inclusion as an emerging technology should always be carried out with this objective in mind, to improve treatment, and not simply for the sake of introducing an additional technological element. This paper reviews works on databases such as MEDLINE, EMBASE, PubMed, Scopus and Cochrane Central, focusing on the use of terms such as chatbots, conversational agents, conversational **AI**, etc.

While there is already precedent in the literature for the use of text messages and SMSs as a tool for mental and physical health treatment (Hall et al., 2015; Rathbone & Prescott, 2017), these have been replaced by more modern web-based solutions, integrated into mobile apps or with more advanced voice interfaces. In particular, **mobile phones (smart phones)**, as ubiquitous devices in widespread use, have massively become the object of study for their application in therapeutic intervention.

On the other hand, advances in artificial intelligence are giving a boost to much more advanced bots. Applied to coaching (virtual coaches), personalised personal agents or specific applications for the control of behavioural changes (Car et al., 2020).

In health, the main lines of application in recent years are:

- Treatment and monitoring.
- Support to health services.
- Patient education.

In such contexts, jobs tend to be more oriented towards **topic-oriented conversations** rather than tasks, and work on the patient rather than the professional. Applications with a text interface tend to be more common than the use of voice. Only in the particular context of older people, due to sight and mobility problems, the use of voice is more widespread.

On the other hand, its use on the web favours its use and reduces abandonment, compared to its use with mobile apps, which makes incorporation or onboarding more difficult, with the exception of widespread messaging apps such as Facebook, Messenger, Telegram, Whatsapp, etc.

In (Car et al., 2020), although recognising the wide range of fields in which it is applied in health, some particular fields are highlighted, such as:

- Mental health (Abd-Alrazaq et al., 2020; Bérubé et al., 2021; Piette et al., 2013).
- Neurodegeneration (Li et al., 2020; Rahman et al., 2021).
- Obesity and diabetes (Steinberg et al., 2014).
- Sexual health (Bauermeister et al., 2017).

Although other lines of application are also in their infancy:

- Primary care (Lee et al., 2021; Fan et al., 2021; Schario et al., 2022).
- Cardiology (Nahar & Lopez-Jimenez, 2022).
- Coaching for adolescents (Gabrielli et al., 2020).
- Dermatology.
- Disability (Masina et al., 2020).
- Nursing (education) (Shorey et al., 2019).
- Cardiovascular diseases (Kowalska et al., 2020).
- Kidney diseases (Fink et al., 2016).
- Pulmonary diseases (Gross et al., 2020; Kim et al., 2021).
- Geriatric (Gudala et al., 2022; Bennion et al., 2020).
- Stress management (Mauriello et al., 2021).
- Obstetrics (Chung et al., 2021).
- Oncology (Bibault et al., 2019; Greer at al., 2019; Chaix et al., 2019; Greer et al., 2019; Hong et al., 2021).
- Orthopaedics (Bian et al., 2020).
- Paediatrics (Wong et al., 2021; Espinoza et al., 2020).
- Vaccination (Ferrand et al., 2020; Wijesundara et al., 2020).

From the perspective of early childhood education it has been approached with the use of *PopBots* (Crompton et al., 2018; Williams et al., 2019). But from a point of view closer to industrial robotics with which they interact in a constructivist way, with Lego elements, sensors, motors, tablets and mobile apps, not in a conversational and patient-oriented sense. However, in the more specific area of **early care**, the application of this type of technology is still scarce, more from the point of view of **assisting the therapist**, and therefore there is a whole line of research open in the future on the use of conversational bots in this field.

Summary

The module sets out with the basic elements to understand what botsare (as intelligent personal assistants) and how they work at an abstract level. It also shows their possible use and usefulness in the field of health, and points to their future applications in **early care**.

Glosary

AI: Artificial Intelligence

IPA: Intelligent Personal Assistant

NLP: Natural Language Processing

NLU: Natural Language Understanding

TTS: Text To Speech

STT: Speech To Text

Bibliography

Basic bibliography

Colby, K. M. (1975). Artificial Paranoia: Computer Simulation of Paranoid Processes. Elmsford, N.Y.;Pergamon Press.

Pearl, C. (2016). Designing voice user interfaces: Principles of conversational experiences. O'Reilly Media.

Shevat, A. (2017). Designing bots: Creating conversational experiences. O'Reilly Media.

Complementary bibliography

Abd-Alrazaq, A. A., Rababeh, A., Alajlani, M., Bewick, B. M., & Househ, M. (2020). Effectiveness and safety of using chatbots to improve mental health: Systematic review and meta-analysis. In *Journal of Medical Internet Research* (Vol. 22, Issue 7, p. e16021). JMIR Publications Inc. https://doi.org/10.2196/16021

Bauermeister, J., Giguere, R., Leu, C. S., Febo, I., Cranston, R., Mayer, K., & Carballo-Diéguez, A. (2017). Interactive voice response system: Data considerations and lessons learned during a rectal microbicide placebo adherence trial

for young men who have sex with men. *Journal of Medical Internet Research, 19* (6), e7682. https://doi.org/10.2196/jmir.7682

Bérubé, C., Schachner, T., Keller, R., Fleisch, E., Wangenheim, F. v., Barata, F., & Kowatsch, T. (2021). Voice-based conversational agents for the prevention and management of chronic and mental health conditions: Systematic literature review. In *Journal of Medical Internet Research* (Vol. 23, Issue 3, p. e25933). JMIR Publications Inc. https://doi.org/10.2196/25933

Bian, Y., Xiang, Y., Tong, B., Feng, B., & Weng, X. (2020). Artificial intelligence-assisted system in postoperative follow-up of orthopedic patients: Exploratory quantitative and qualitative study. *Journal of Medical Internet Research*, *22* (5), e16896. https://doi.org/10.2196/16896

Bennion, M. R., Hardy, G. E., Moore, R. K., Kellett, S., & Millings, A. (2020). Usability, acceptability, and effectiveness of web-based conversational agents to facilitate problem solving in older adults: Controlled study. *Journal of Medical Internet Research, 22* (5), e16794. https://doi.org/10.2196/16794

Bérubé, C., Schachner, T., Keller, R., Fleisch, E., Wangenheim, F. v., Barata, F., & Kowatsch, T. (2021). Voice-based conversational agents for the prevention and management of chronic and mental health conditions: Systematic literature review. In Journal of Medical Internet Research (Vol. 23, Issue 3, p. e25933). JMIR Publications Inc. https://doi.org/10.2196/25933

Bibault, J. E., Chaix, B., Guillemassé, A., Cousin, S., Escande, A., Perrin, M., Pienkowski, A., Delamon, G., Nectoux, P., & Brouard, B. (2019). A chatbot versus physicians to provide information for patients with breast cancer: Blind, randomized controlled noninferiority trial. *Journal of Medical Internet Research*, *21* (11), e15787. https://doi.org/10.2196/15787

Car, L. T., Dhinagaran, D. A., Kyaw, B. M., Kowatsch, T., Joty, S., Theng, Y. L., & Atun, R. (2020). Conversational agents in health care: Scoping review and conceptual analysis. In *Journal of Medical Internet Research* (Vol. 22, Issue 8, p. e17158). JMIR Publications Inc. https://doi.org/10.2196/17158

Chaix, B., Bibault, J. E., Pienkowski, A., Delamon, G., Guillemassé, A., Nectoux, P., & Brouard, B. (2019). When chatbots meet patients: One-year prospective study of conversations between patients with breast cancer and a chatbot. *JMIR Cancer*, 5(1), e12856. https://doi.org/10.2196/12856

Chung, K., Cho, H. Y., & Park, J. Y. (2021). A chatbot for perinatal women's and partners' obstetric and mental health care: development and usability evaluation study. JMIR Medical Informatics, 9 (3), e18607. https://doi.org/10.2196/18607

Crompton, H., Gregory, K., & Burke, D. (2018). Humanoid robots supporting children's learning in an early childhood setting. *British Journal of Educational Technology, 49* (5), 911–927. https://doi.org/10.1111/bjet.12654

Espinoza, J., Crown, K., & Kulkarni, O. (2020). A guide to chatbots for COVID-19 screening at pediatric health care facilities. *JMIR Public Health and Surveillance*, *6* (2), e18808. https://doi.org/10.2196/18808

Fan, X., Chao. D., Zhang, Z., Wang, D., Li, X., y Tian, F. (2021). *Utilization of Self-Diagnosis Health Chatbots in Real-World Settings: Case Study*. J Med Internet Res, 23(1), e19928. https://doi.org/10.2196/19928

Ferrand, J., Hockensmith, R., Houghton, R. F., & Walsh-Buhi, E. R. (2020). Evaluating smart assistant responses for accuracy and misinformation regarding human papillomavirus vaccination: Content analysis study. Journal of Medical Internet Research, 22 (8), e19018. https://doi.org/10.2196/19018

Fink, J. C., Doerfler, R. M., Yoffe, M. R., Diamantidis, C. J., Blumenthal, J. B., Siddiqui, T., Gardner, J. F., Snitker, S., & Zhan, M. (2016). Patient-Reported Safety Events in Chronic Kidney Disease Recorded With an Interactive Voice-Inquiry Dial-Response System: Monthly Report Analysis. *Journal of Medical Internet Research*, *18* (5), e5203. https://doi.org/10.2196/jmir.5203

Greer, S., Ramo, D., Chang, Y. J., Fu, M., Moskowitz, J., & Haritatos, J. (2019). Use of the chatbot "vivibot" to deliver positive psychology skills and promote well-being among young people after cancer treatment: Randomized controlled feasibility trial. JMIR MHealth and UHealth, 7(10), e15018. https://doi.org/10.2196/15018

Gabrielli, S., Rizzi, S., Carbone, S., & Donisi, V. (2020). A chatbot-based coaching intervention for adolescents to promote life skills: Pilot study. *JMIR Human Factors*, *7* (1), e16762. https://doi.org/10.2196/16762

Gross, C., Kohlbrenner, D., Clarenbach, C. F., Ivankay, A., Brunschwiler, T., Nordmann, Y., & v Wangenheim, F. (2020). A Telemonitoring and Hybrid Virtual Coaching Solution "CAir" for Patients with Chronic Obstructive Pulmonary Disease: Protocol for a Randomized Controlled Trial. *JMIR Research Protocols*, *9* (10), e20412. https://doi.org/10.2196/20412

Gudala, M., Ross, M. E. T., Mogalla, S., Lyons, M., Ramaswamy, P., & Roberts, K. (2022). Benefits of, Barriers to, and Needs for an Artificial Intelligence-Powered Medication Information Voice Chatbot for Older Adults: Interview Study With Geriatrics Experts. *JMIR Aging*, 5 (2), e32169. https://doi.org/10.2196/32169

Gupta, J., Singh, V., y Kumar, I. (2021). "Florence- A Health Care Chatbot," 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS) (p. 504-508). https://doi.org/10.1109/ICACCS51430.2021.9442006

Hall, A. K., Cole-Lewis, H., & Bernhardt, J. M. (2015). Mobile Text Messaging for Health: A Systematic Review of Reviews. Annual Review of Public Health, 36(1), 393–415. https://doi.org/10.1146/annurev-publhealth-031914-122855

Hong, G., Folcarelli, A., Less, J., Wang, C., Erbasi, N., & Lin, S. (2021). Voice assistants and cancer screening: A comparison of alexa, siri, google assistant, and cortana. *Annals of Family Medicine*, *19* (5), 447–449. https://doi.org/10.1370/AFM.2713

Kim, A. J., Yang, J., Jang, Y., & Baek, J. S. (2021). Acceptance of an informational antituberculosis chatbot among korean adults: Mixed methods research. *JMIR MHealth and UHealth*, *9* (11), e26424. https://doi.org/10.2196/26424

Kowalska, M., Gładyś, A., Kalańska-Łukasik, B., Gruz-Kwapisz, M., Wojakowski, W., & Jadczyk, T. (2020). Readiness for voice technology in patients with cardiovascular diseases: Cross-sectional study. *Journal of Medical Internet Research*, *22* (12), e20456. https://doi.org/10.2196/20456

Lee, H., Kang, J., & Yeo, J. (2021). Medical specialty recommendations by an artificial intelligence chatbot on a smartphone: Development and deployment. *Journal of Medical Internet Research*, 23 (5), e27460. https://doi.org/10.2196/27460

Li, J., Maharjan, B., Xie, B., & Tao, C. (2020). A personalized voice-based diet assistant for caregivers of alzheimer disease and related dementias: System development and validation. *Journal of Medical Internet Research*, *22* (9), e19897. https://doi.org/10.2196/19897

Mauriello, M. L., Tantivasadakarn, N., Mora-Mendoza, M. A., Lincoln, E. T., Hon, G., Nowruzi, P., Simon, D., Hansen, L., Goenawan, N. H., Kim, J., Gowda, N., Jurafsky, D., & Paredes, P. E. (2021). A suite of mobile conversational agents for daily stress management (popbots): Mixed methods exploratory study. *JMIR Formative Research*, 5 (9), e25294. https://doi.org/10.2196/25294

Masina, F., Orso, V., Pluchino, P., Dainese, G., Volpato, S., Nelini, C., Mapelli, D., Spagnolli, A., & Gamberini, L. (2020). Investigating the accessibility of voice assistants with impaired users: Mixed methods study. *Journal of Medical Internet Research*, *22* (9), e18431. https://doi.org/10.2196/18431

Nahar, J. K., & Lopez-Jimenez, F. (2022). Utilizing Conversational Artificial Intelligence, Voice, and Phonocardiography Analytics in Heart Failure Care. In *Heart Failure Clinics* (Vol. 18, Issue 2, pp. 311–323). Elsevier Inc. https://doi.org/10.1016/j.hfc.2021.11.006

Piette, J. D., Sussman, J. B., Pfeiffer, P. N., Silveira, M. J., Singh, S., & Lavieri, M. S. (2013). Maximizing the value of mobile health monitoring by avoiding redundant patient reports: Prediction of depression-related symptoms and adherence problems in automated health assessment services. *Journal of Medical Internet Research*, *15* (7), e2582. https://doi.org/10.2196/jmir.2582

Prize L. Mitsuku Wins 2019 Loebner Prize and Best Overall Chatbot at AISB X. AISB – The Society for the Study of Artificial Intelligence and Simulation of Behaviour. 2019. URL: https://aisb.org.uk/mitsuku-wins-2019-loebner-prize-and-best-overall-chatbot-at-aisb-x/

Pumplun, L., Fecho, M., Wahl, N., Peters, F., y Buxmann, P. (2021). Adoption of Machine Learning Systems for Medical Diagnostics in Clinics: Qualitative Interview Study. J Med Internet Res, 23(10), e29301. https://doi.org/10.2196/29301

Rathbone, A. L., & Prescott, J. (2017). The use of mobile apps and SMS messaging as physical and mental health interventions: Systematic review. *In Journal of Medical Internet Research* (Vol. 19, Issue 8, p. e7740). JMIR Publications Inc. https://doi.org/10.2196/jmir.7740

Rahman, W., Lee, S., Islam, M. S., Antony, V. N., Ratnu, H., Ali, M. R., Mamun, A. al, Wagner, E., Jensen-Roberts, S., Waddell, E., Myers, T., Pawlik, M., Soto, J., Coffey, M., Sarkar, A., Schneider, R., Tarolli, C., Lizarraga, K., Adams, J., … Hoque, E. (2021). Detecting parkinson disease using a web-based speech task: Observational study. *Journal of Medical Internet Research*, *23* (10), e26305. https://doi.org/10.2196/26305

Schario, M. E., Bahner, C. A., Widenhofer, T. v., Rajaballey, J. I., & Thatcher, E. J. (2022). Chatbot-Assisted care management. *Professional Case Management*, *27* (1), 19–25. https://doi.org/10.1097/NCM.0000000000000504

Shan, Y., Ji, M., Xie, W., Qian, X., Li, R., Zhang, X., y Hao, T. (2022). Language Use in Conversational Agent–Based Health Communication: Systematic Review. J Med Internet Res, 24(7), e37403. https://doi.org/10.2196/37403

Shorey, S., Ang, E., Yap, J., Ng, E. D., Lau, S. T., & Chui, C. K. (2019). A virtual counseling application using artificial intelligence for communication skills training in nursing education: Development study. *Journal of Medical Internet Research*, *21* (10), e14658. https://doi.org/10.2196/14658

Steinberg, D. M., Levine, E. L., Lane, I., Askew, S., Foley, P. B., Puleo, E., & Bennett, G. G. (2014). Adherence to self-monitoring via interactive voice response technology in an ehealth intervention targeting weight gain prevention among black women: Randomized controlled trial. *Journal of Medical Internet Research*, *16* (4), e2996. https://doi.org/10.2196/jmir.2996

Wijesundara, J. G., Fukunaga, M. I., Ogarek, J., Barton, B., Fisher, L., Preusse, P., Sundaresan, D., Garber, L., Mazor, K. M., & Cutrona, S. L. (2020). Electronic health record portal messages and interactive voice response calls to improve rates of early season influenza vaccination: Randomized controlled trial. *Journal of Medical Internet Research*, *22* (9), e16373. https://doi.org/10.2196/16373

Williams, R., Park, H. W., Oh, L., & Breazeal, C. (2019). Popbots: Designing an artificial intelligence curriculum for early childhood education. *33rd AAAI Conference on Artificial Intelligence, AAAI 2019, 31st Innovative Applications of Artificial Intelligence Conference, IAAI 2019 and the 9th AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2019*, *33* (01), 9729–9736. https://doi.org/10.1609/aaai.v33i01.33019729

Wong, J., Foussat, A. C., Ting, S., Acerbi, E., van Elburg, R. M., & Chien, C. M. (2021). A chatbot to engage parents of preterm and term infants on parental stress, parental sleep, and infant feeding: Usability and feasibility study. *JMIR Pediatrics and Parenting*, 4 (4), e30169. https://doi.org/10.2196/30169

Web

Amazon (2022a) Amazon Alexa official site: What is Alexa? (n.d.). Amazon (Alexa). https://developer.amazon.com/en-US/alexa. Retrieved, July 17, 2022.

Amazon (2022b) Chatbot | Deep learning | Amazon Lex. (n.d.). Amazon Web Services, Inc. https://aws.amazon.com/lex/?nc1=h_ls . Retrieved, July 17, 2022.

Apple (2022) Siri. (n.d.). Apple . https://www.apple.com/siri/. Retrieved, July 17, 2022.

Chatcompose (2022) . ChatCompose - Chatbot Platform for Sales and Support Automation. https://www.chatcompose.com/en.html. Retrieved, July 17, 2022.

Chatfuel (2022, 11). Chatfuel | Customer support and sales automation. https://chatfuel.com/. Retrieved July 17, 2022.

IBM (2022) IBM Watson assistant - Virtual agent. (n.d.). IBM - United States. https://www.ibm.com/products/watson-assistant. Retrieved July 17, 2022.

Google (2022). Dialogflow. https://dialogflow.cloud.google.com/. Retrieved July 17, 2022.

Meta (2022) Wit.ai. https://wit.ai/ Retrieved July 17, 2022.

Microsoft (2022b) Microsoft bot framework. (n.d.). https://dev.botframework.com/ Retrieved, July 17, 2022.

Open source conversational AI. (2020, December 1). Rasa. https://rasa.com/ Retrieved, July 17, 2022.

UNICEF (2022a) Safer Chatbots | UNICEF. (n.d.). Retrieved July 14, 2022, from https://www.unicef.org/documents/safer-chatbots

UNICEF (2022b) Safer Chatbots Implementation Guide | UNICEF. (n.d.). Retrieved July 14, 2022, from https://www.unicef.org/documents/safer-chatbots-implementation-guide