

Using health chatbots for behavior change: a mapping study

Juanan Pereira · Óscar Díaz

Received: date / Accepted: date

Abstract This study conducts a mapping study to survey the landscape of health chatbots along three research questions: What illnesses are chatbots tackling? What patient competences are chatbots aimed at? Which chatbot technical enablers are of most interest in the health domain? We identify 30 articles related to health chatbots from 2014 to 2018. We analyze the selected articles qualitatively and extract a triplet $\langle technicalEnablers, competence, illness \rangle$ for each of them. This data serves to provide a first overview of chatbot-mediated behavior change on the health domain. Main insights include: nutritional disorders and neurological disorders as the main illness areas being tackled; “affect” as the human competence most pursued by chatbots to attain change behavior; and “personalization” and “consumability” as the most appreciated technical enablers. On the other hand, main limitations include lack of adherence to good practices to case-study reporting, and a deeper look at the broader sociological implications brought by this technology.

Keywords Chatbots · Mobile Agents · Mobile Healthcare · Instant Messaging · Software Agents

1 Introduction

Chatbots are machine agents that serve as natural language user interfaces for data and service providers [15]. Many chatbots are being devised towards

Juanan Pereira
ONEKIN Research Group
University of the Basque Country, UPV/EHU
E-mail: juanan.pereira@ehu.eus

Óscar Díaz
ONEKIN Research Group
University of the Basque Country, UPV/EHU
E-mail: oscar.diaz@ehu.eus

helping patients with symptom-based diagnosis, receiving instant feedback regarding general health questions. Chatbots can learn from previous interactions in order to increase the accuracy of their disease recognition. The vision is for chatbots to help people in less time and for less money than it would take to see a medical professional. Early experiences on the use of this technology began around 2014, and it is currently experimenting a great interest among both the medical and the computing communities. This work surveys the available knowledge about the usage of health chatbots.

Chatbots bring several benefits: anonymity, asynchronicity, personalization, scalability, authentication or consumability to name a few [34]. These benefits spark the interest for this technology to face a rapidly aging population and stringent demands of chronic illness attention [6]. From this perspective, chatbots align with previous IT technologies to support “mobile health”. Specifically, Embodied Conversational Agents (ECAs) (i.e. a sort of human-like avatars that are usually oriented to desktop computers and not mobile devices) have been assessed to support different kind of services, namely, interviews [52, 41], counseling [53, 45], chronic health conditions monitoring [48, 10] or medication adherence [49, 5]. These approaches are generally not usable in mobile phones, mostly due to browser-plugin requirements or assumption of large-screen availability [37, 36, 20, 6, 8]. On the other hand, Conversational Agents (i.e. those that use any unconstrained natural language input) have been recently surveyed [40]. Here, the emphasis is on the capabilities offered by Natural Language Processing in a healthcare setting, no matter whether these capabilities are offered through chatbots, ECAs or Smart Conversational Interfaces such as Apple’s Siri or Amazon’s Alexa.

Unlike previous studies, our focus is not so much about the means but the ends of the conversation. If chatbots are conversational interfaces, then what are that conversations aimed at? No matter the illness, chatbots engage in a conversation to track, educate, encourage or prevent some behavior on the patient. The term “behavior (change) centered paradigm” is being coined to stress that the final aim is to assist users in changing their behavior rather than the more traditional approach of assisting users in conducting a task [6]. Specifically, we focus on *using chatbots for behavior change for healthy purposes*. This gives rise to three research questions:

- RQ1 (“*for healthy purposes*”): What illnesses are chatbots tackling?
- RQ2 (“*for behavior change*”): What patient competences are chatbots aimed at?
- RQ3 (“*using chatbots*”): Which chatbot enablers are of most interest in the health domain?

This paper presents the results of a Systematic Mapping Study (SLR) to address the previous questions. Our work differs from previous surveys in both the platform (i.e. chatbots rather than ECAs) and the target audience (i.e. patients rather than clinicians), no matter the sophistication of the chatbot abilities (e.g. we do not exclude studies where user input occurred by clicking

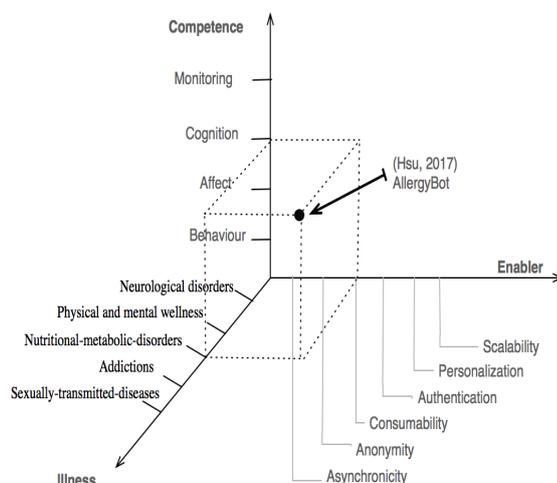


Fig. 1 The chatbot space along three dimensions: illness, patient competence and technical enabler

rather than voice or natural language recognition). Our interest in chatbots comes from having developed health chatbots ourselves[46].

2 The chatbot space

Chatbots allow for direct user engagement through text messaging. This engagement is sought for different purposes: marketing, banking, booking, etc. We focus on “*using chatbots for behavior change for healthy purposes*”. That is, primary studies (PSs) are regarded as instantiations of the previous statement along a triplet $\langle illness, competences, technicalEnablers \rangle$ where “*illness*” serves to pigeonhole PSs along the Curlie’s classification¹; “*competences*” denotes the desired patient behavior; and finally, “*technicalEnablers*” refers to the benefits brought by chatbot technology which are mentioned in the primary study. This conforms a three-dimension space to arrange chatbot proposals (see Fig. 1). The rest of this section introduces the main values for each dimension.

2.1 Technical Enablers

This subsection highlights what we value most from the chatbot technology in a health setting. Other domains (e.g. banking, retailing) might certainly value most other features.

¹ <http://www.curlie.org/Health/>

Asynchronicity. Chatbots provide the perfect blend between immediacy (prompt answer, i.e. a reactive behavior) and asynchronicity (notifications and reminders, i.e. a proactive behavior). When combined with social media, chatbots offer a powerful blend to not only reach, but also engage patients in illness prevention and care, specially youngsters due to their familiarity with the media.

Consumability. Some scenarios require for chatbots to be readily “consumable”, e.g. when craving appears in addiction treatment. Consumability is a description of customers’ end-to-end experience with technology solutions, and includes not only the use of the tool but also the extent to which installing, configuring, and administering the tool is perceived as easy². Consumability wise, chatbots outperform previous technologies in different aspects: installation (limited to add the chatbot to the list of contacts), platform independence (e.g. Instant Messaging (IM) apps are available for all major platforms -Android, iOS, Windows Mobile, Linux, Windows and macOS) and learning effort (i.e. chatbots follow a conversational interface, with multi-modal input -text, button-oriented-navigation or voice- without needing to learn new GUI). We also include as part of consumability, ubiquity. Basically, chatbots are instantly available in the same way that other IM contacts are, with no installation hassle. This might result in a competitive advantage compare with mobile apps [1]. Chatbots can also tap into the myriad of smartphone sensors to collect data in a transparent way (i.e. no consumption effort).

Anonymity. When it comes to sensitive healthcare issues, the possibility of interacting anonymously becomes a main enabler. Patients might feel less shame and open when interacting with computers, and show positive sentiment towards the software agents, feeling more private and anonymous in comparison with speaking to real humans [23].

Authentication. The process of verifying the identity of patients can be facilitated through built-in smartphone mechanisms. Chatbots can be secured using many of the same security strategies used for other mobile technologies: login credentials, two-factor authentication (i.e. the patient is required to verify their identity through two separate channels), biometric authentication (e.g. retina scan, fingerprint), etc.

Personalization. Meeting patients’ idiosyncrasies more effectively yields increasing user satisfaction which, in turn, leads to better treatment engagement. Smartphone sensors account for a transparent mechanism to collect patients’ behavior which can later feed AI algorithms. GPS and accelerometer data can serve as indicators of physical activity, facial recognition can help recognize user physical state, or in conjunction with a smartphone-based heart rhythm monitor, provide counseling to patients with heart-based condition. This promotes a more engaging and personalized conversation experience.

Scalability. Chatbots have the potential to target large audiences in a cost-effective way [37].

² While usability addresses a user’s ability to use a product, consumability is a higher-level concept that incorporates all the other aspects of the customer’s experience with the product <https://en.wikipedia.org/wiki/Consumability>

2.2 Behavior change

Previous technical enablers aim at providing an engaging experience, but what is the ultimate purpose? If chatbots are conversational interfaces, what is that conversation aimed at? In a health setting, the final aim is to facilitate behavior change through enhancing human competences [6]. This subsection revises these competences along those provided by Brinkman [6].

Monitoring. Awareness and tracking of bad habits is certainly a first step in changing behavior. Chatbots can benefit from external data capture through a wealth-sensor infrastructure: blood pressure, stress level, weight or amount of physical activity, could all be monitored to encourage healthy behavior.

Cognition. The next step for change is to elaborate on the causes and potential diagnosis. Chatbots can play a role on empowering patients to understand implications of health conditions. In this way, the hope is to support engagement through understanding.

Affect. Sustained engagement might be needed beyond the success of a first encounter with the chatbot. Empathy, understanding, acknowledging people's emotional state are key for sustained patient involvement. Combining personality and emotional aspects in the dialogues, for instance introducing social dialogues - small-talk or chit-chat sentences - can improve patients' satisfaction and engagement with the bot. Developing rapport with chatbots has also been in the radar towards promoting a sense of self-efficacy that ends up with patients persevering in the related behavior [25].

Behavior. Chatbots do not stop at tracking and informing patients. They can go a step further by influencing their behaviors. Chatbots might help through reminders (e.g. take medication, do exercise), gamification (e.g. badgers, pair competition) or removing potential barriers (e.g. leaving free agenda evening slots for jogging).

3 Methodology

To tackle our research questions, we systematically assessed existing evidence related to chatbot usage using SLR guidelines [47]. An SLR facilitates identifying and collecting key papers in a specific area of interest, and evaluating and interpreting the reporting discussions and findings. An SLR comprises a defined review protocol, search strategy, explicit inclusion and exclusion criteria, and specified information that will be retrieved from primary studies. This section summarizes this approach for our purposes

3.1 Search Strategy

To help build the search terms and identify potential overlapping, a set of key papers were identified, mainly those surveys related to nearby areas. On these grounds, the search string was conformed along the PICOC (Population, Intervention, Comparison, Outcome, Context) criteria:

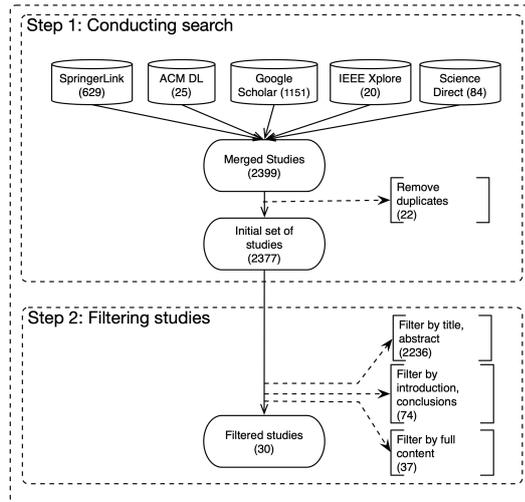


Fig. 2 Filtering studies.

- Population: patients (rather than health practitioners)
- Intervention: the use of chatbot for healthcare interventions
- Comparison: we do not compare different strategies for healthcare interventions but assess the area as a whole,
- Outcomes: characterization of promoting patients’ competences in terms of monitoring, cognition, affect, and behavior,
- Context: any context in which patients interact with chatbots.

The identified keywords were: *‘patient’*, *‘chatbot’* and *‘healthcare’*. Next, synonyms should be found. Along the guidelines of Petersen’s [47], we conducted an informal literature search in order to identify keywords and to find a balance between hits and noise. We noticed that the terms *‘chatbots’*, *‘conversational agents’* and *‘virtual agents’* tend to be used interchangeably. Hence, we included those terms. This resulted in the following search string:

((“conversational agent” OR “virtual agent” OR “chatbot”) AND (“health”)
AND (“patient”))*

However, the population filter (i.e. *‘patient’*) was too restrictive for two repositories, i.e. IEEE Xplore and ACM Digital Library. Hence, we decided to remove *‘patient’* from the search query, and support it as an exclusion criterion.

This query was matched against the title, the abstract and the keywords. We restricted the search to studies published from 2014 up to December 2018. Five electronic databases were consulted: IEEE Xplore, ACM Digital Library, Springer Link, Science Direct and Google Scholar. Fig. 2 outlines the resulting numbers. In total, we obtained 2377 primary studies in this first step.

3.2 Filter Strategy

Studies were excluded if they met any of the following criteria:

- EC1 (*'The study was published before 2014'*). Chatbots started to be created around 2014, when Telegram added support for chatbots development [38].
- EC2 (*'The study is not centric to mobile-based chatbots'*). Embodied Conversational Agents (i.e. a sort of human-like avatars that are usually oriented to desktop computers and not mobile devices) are left outside this study (see [29] for a review).
- EC3 (*'The study does not target patients'*). We removed articles that do not directly target patients (e.g. caregivers or healthcare professionals)

Filtering was conducted by one author, and if in doubt, the second author also intervened. This happened eleven times. No quality assessment was conducted except that of being peer-reviewed papers.

3.3 Data extraction and classification

This includes two main steps. First, identifying relevant topic key-wording that help answering the research questions. This yields the classification schema. Our classification schema includes three main dimensions: $\langle illness, competences, technicalEnablers \rangle$ (see Section 2). Second, data synthesis. Each PSs were positioned in the "chatbot space" along the aforementioned dimensions. A PSs is characterized along a single illness but it might tackle more than one competence and technical enabler. Illness are explicitly mentioned in the PS text. However, competences and enablers were obtained using an integrated approach to thematic analysis that aims at quantifying (according to predetermined categories) content in a systematic and reliable manner [14]. Two investigators reviewed details extracted from the set of PSs in order to identify the illness, the aimed patient competences, and the underlined technical enablers brought by chatbots. Tables 1 and 2 gather these 30 triplets that provide the raw data in which to base the findings.

4 FINDINGS

This section goes back to the research questions. References refer to the primary studies (PSs) being considered.

4.1 RQ1: What illnesses are chatbot tackling?

We classified Primary Studies (PSs) along the Curlie's classification for illness typification (see Fig. 3)



Fig. 3 Tree-map depicting illnesses categories tackled by revised chatbots. PS numbers are shown in brackets.

Neurological disorders (6 PSs). These chatbots are oriented towards assisting patients with psychiatric diagnosis illnesses, like dementia [3, 13], Alzheimer [25], insomnia [4], depression[50] or general psychiatric counseling [45].

Mental and physical wellness (8 PSs). Promoting healthy habits [42, 44, 9, 31, 24] and physical exercises [8, 20], along with delivering positive psychology and mental well-being techniques is the aim of 8 PSs.

Nutritional-metabolic-disorders (6 PSs). This basically covers obesity [11, 27] -specially child obesity [36, 22]-, allergies [26] and diabetes [10].

Addictions (3 PSs). This includes drug consuming cessation interventions with a special focus on alcoholism [41, 17] and smoking [16].

Sexually-transmitted-diseases (3 PSs). This includes HIV/AIDS [7, 53] and syphilis [35].

Others (3). Chatbot experiences also include treating aphasia [54], diagnosing rare diseases [39], and detecting heart condition and cardiovascular disorder problems [32].

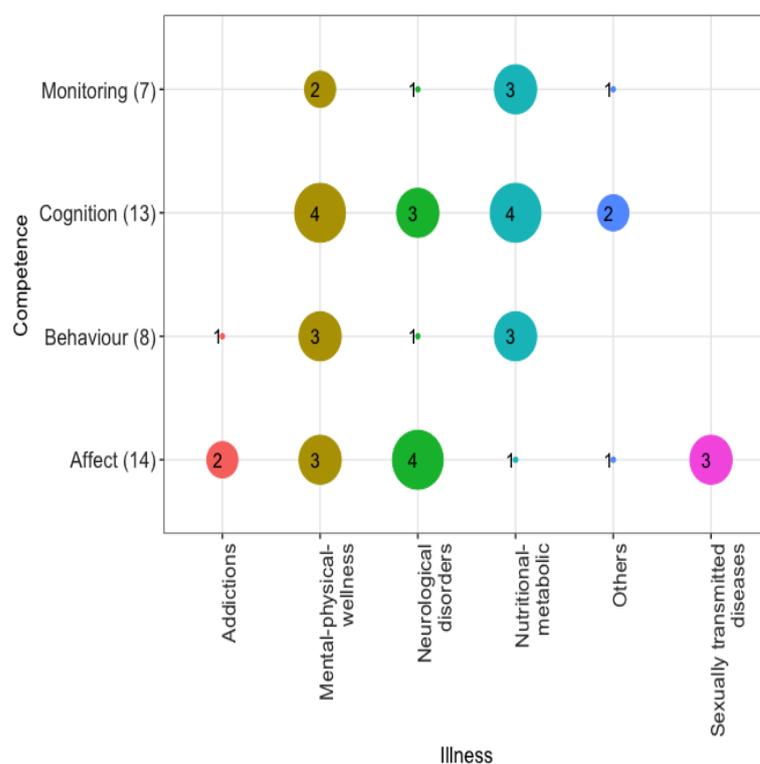


Fig. 4 Bubble chart mapping illness across human competences.

4.2 What patient competences are chatbots aimed at?

This moves us to the human competences chatbots look to promote (see Fig. 4). Evidences are collected from PS quotes which highlight the interest of the human competences. Hence, PSs account for one illness but potentially several competences.

Monitoring (7 PSs). Chatbots can help by tracking different bio-signals: physical activities [20, 8], heart rhythm [32], blood pressure [11], body temperature [27], facial-expressions [31], voice intonation [45], or sensed monitoring chat-based answer feedback [36].

Cognition (13 PSs). Unlike Expert Systems, chatbots target patients, no clinicians. This restricts cognition complexity to indicate simple causal relationships that help patients understand and support them in their decision making. In this way, chatbots become health coaches that back and advise users in different settings: their physical activities [8, 37]; counseling about healthy eating behavior [20, 27, 24]; facilitating patients to provide allergy-related information about visited restaurants and offering appropriate feedback [26]; making patients aware of depression episodes in cancer [50] and diabetes scenarios [10]; fighting back Alzheimer symptoms through quizzes (e.g.,

showing a clock and asking for the time) [25]; improving their adherence to insomnia therapy treatment [4]; providing advice in either general terms (i.e. first-aids like using Apple Siri, Google Now, Samsung Voice or Microsoft Cortana [44]) or illness specific (e.g., atrial fibrillation [32]); or keeping long-term recurrent interviews with the patient for rare disease diagnosis [39].

Affect & Attitude (14 PSs). Brinkman highlights how “attitudinal change, for example, self-efficacy, or attitude towards healthy living, can be a key enabler that could lead up to change in people’s behavior” [6]. Here, chatbots can be engineered to look as “a friendly advisor or mentor to the user rather than a therapist or health care professional” [17]. This is more so when handling sensitive health problems, like alcohol-related issues [17], syphilis [35], HIV/AIDS [53, 7], or aphasia where patients with difficulties to speak can practice without the fear of shame and embarrassment [54]. Here, chatbots provide a secure, anonymous setting, where patients can develop a sense of rapport without fear of stigmatization and discrimination.

In this setting, chatbots can analyze patient mood, and reinforce attitudinal changes, accordingly. First, mood analysis. It can be conducted in different ways: capturing participants’ facial expressions through the camera [11]; analyzing sentiment of text messages [31, 3]); or using prosodic and statistical features extracted through voice signal [50]. Once mood is being analyzed, chatbots might resort to distinct strategies to reinforce attitudinal changes: positive psychology in the context of mental well-being (e.g. practicing kindness, expressing gratitude [42]); telling jokes in the context of dementia [13]; motivational interviewing (i.e. aiming at finding the intrinsic motivation to change patients’ lifestyle) [41, 28]; or praising for the efforts of insomnia patients [4].

Behavior (8 PSs). Chatbots can influence patients’ behavior through distraction and encouragement. The former is illustrated by [16] where smokers are helped quitting by distracting them during the craving phases (e.g. sending multimedia content). In a similar vein, Cameron et al. report on a bot that starts conversations when users are experiencing stress in the workplace [9]. As for encouragement, different strategies are being attempted: meal-portion recommendation messaging at lunch time [22]; gamification techniques [21]; resorting to social-influence strategies (e.g. shared decision making [4]; promoting self-efficacy by tracking patients’ own challenges (e.g. counting steps as daily goals [36]).

Fig. 4 showcases how chatbots address the four competences. Interestingly, *affect & attitude* seem to be the most popular competences being tackled, hence moving chatbot technology a step further from Expert Systems, i.e. addressing not only cognition but also attempting to engage and improve the treatment adherence of patients.

4.3 Which chatbot enablers are of most interest?

This subsection cross-matches aims and enablers to uncover what technological features might deserve more focus, health wise. It also unveils enablers not yet so much used to account for illness treatments (see Figure 5). Evidences are collected from PS quotes which highlight the interest of the enabler. Hence, PSs account for one illness but potentially several enablers.

Scalability (3 PSs). This enabler permits to handle large patient bases in a cost-effective way. Interesting enough, only three studies highlight this enabler as key in their setting. In [36], authors refer to near 13000 conversational turns during 4 months, something nearly impossible to obtain without automating the conversations. Likewise, [16] sets scalability as a main non-functional requirement to handle near 7000 participants with distraction tricks during craving times while quitting smoking. Finally, [28] highlights the value of chatbots as an efficient tool to address a global public health challenge, the lack of enough human psychotherapists to cope with the increasing mental health problems among the population.

Personalization (13 PSs). This enabler facilitates tailoring to patients' medical history and specifics. This comes in handy in nutritional metabolic disorders and physical wellness. The former is illustrated for allergies [26] and diabetes [10] where chatbot behavior is adapted to the patients' profile. [27] and [24] also leverage this enabler to provide personalized diet and healthy eating tips to their patients. As for physical wellness, personalization allows for custom health screening [9], custom motivational support [20, 42] and custom guidance assistance [8]. In [22], the chatbot explicitly requests users about message appropriateness, and change its behavior accordingly.

In addition, personalization techniques can also help simulate affect. Discussing about addictions or at-risk behaviors can be emotional engaging (e.g. shame, discouragement, fear, anger) [41, 17]. For succeeding in the therapy efforts, affinity and empathy are a must. Chatbots can *simulate* emotion by resorting to computer vision to capture emotion, and react accordingly, e.g. introducing motivational comments in the conversation [16]. This explains that 3 out of 3 addiction-related PSs value personalization capabilities.

Consumability (11 PSs). Easy consumption becomes key to promote and engage patients in the persistent use of chatbots. This is the case in health screening [9], motivation support [20] and guidance assistance [8]. In addition, this enabler is also largely noted in the area of neurological disorders. Capturing participants' facial expressions or voice analysis are available to tune chatbot interactions [31]. This comes in handy to tackle mental disorders, usually aging-associated [3], where patients usually are more apt to use their voice than text commands. For patients with neurological conditions (dementia [13], depression [50] or Alzheimer [25]), chatbots' usability and accessibility features are highly regarded. Finally, ubiquity turns to be a must in different scenarios: when trying to detect early symptoms of cardiovascular disorders [32], when patients with allergies want to check if there are nearby suitable restaurants [26], or for rapid intervention in insomnia episodes [4].

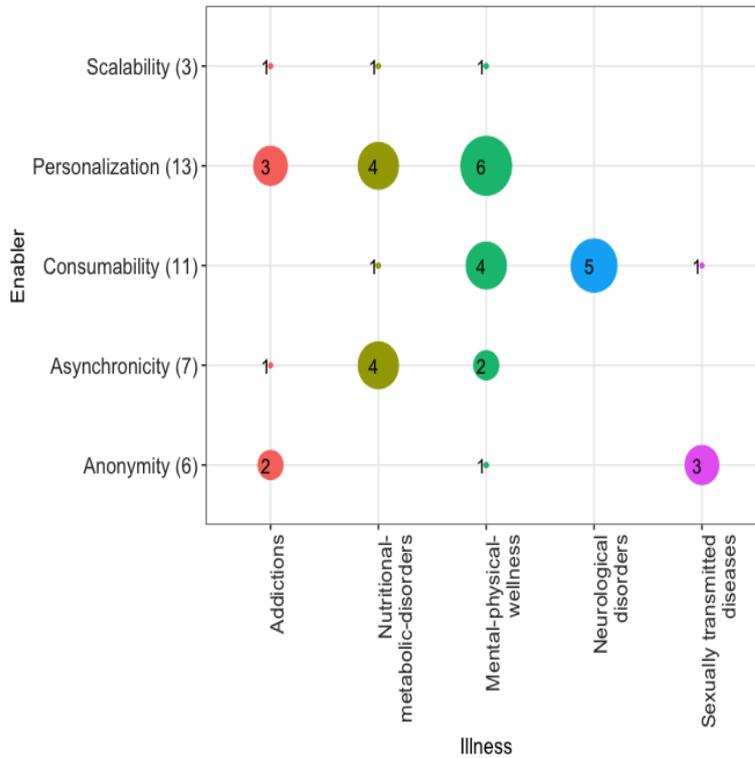


Fig. 5 Bubble chart mapping illness across technical enablers.

Asynchronicity (7 PSs). This enabler promotes “on-site openness” where patients expose their hesitance at the time craving appears [16]. This enabler is specially appreciated in three scenarios: remembering assigned exercises [36, 42]; praising for achievements [11]; and encouraging adherence to the treatment [22].

Anonymity (6 PSs). As expected, anonymity is a key enabler for sexually transmitted diseases. Potential carriers of sexually transmitted infection diseases tend to avoid early consultation of medical advice due to the fear of stigmatization [7, 53, 35]. Similar rationales hold for addictions [41], sensitive mental health issues [43], and aphasia where patients might unashamedly practice with the bot [54].

As a byproduct of this technical dimension, we were also interested in knowing the extent to which developers resorted to chatbot platforms or rather they had to elaborate on their own. Figures follow: custom developments (57.7%), Facebook Messenger (15.4%), Telegram (11.5%) and not-specified (15.4%). These figures seem counter-intuitive since platforms speed up chatbot development in contrast with cumbersome custom development. The rationales behind this option might be in the inability of chatbot platform to use custom UI widgets for communication, or more importantly, the inability to directly access

smartphone’s built-in or external -wearable- sensors (like accelerometer, gyroscope, or possible health related sensors like ECG, heart rate, breathing and activity monitors, spectrophotometer, or any other biometric sensor). When developers do not need these features, they tend to use Facebook Messenger and Telegram platforms for their bots. But custom development has also its downsides. In addition to costly development, consumability decreases since patients are forced to install the custom app. A compromise could be reached if IM platforms would allow developers to access - with users’ permission- any data available for the smartphone (as it is the case of apps).

5 Discussion

5.1 Implications for Research and Practice

To our knowledge, this is the first mapping study of chatbot-mediated behavior change on the health domain. Our findings show an increasing interest in this kind of technology to engage patients. However, the novelty of the area makes most reported efforts be devoted to develop the chatbot from scratch rather than conducting more systematic and ample studies. At this respect, the large number of *custom* developments points to the opportunity of dedicated platforms that tackle the specifics of the health domain. The youth of the domain is also noted by the lack of rigor when reporting experiences. Good practices as those presented in [30] would help compare chatbots that address similar illness or tap into the same enablers. Similar initiatives are also available in the health domain, e.g. the Consolidated Standards of Reporting Trials of electronic and mobile health applications and online telehealth (Consort-eHealth) [18].

Primary studies also fall short of reporting the role played by *supporters* during the chatbot intervention. In this setting, support is defined as contact with a human with the aim of increasing the patient’s ability to use the chatbot to obtain the intended treatment result. Indeed, integrating human support to promote engagement and provide technical and clinical troubleshooting is being reported as more efficacious than self-guided, self-help interventions [2]. This plays in favor of chatbot. Unlike traditional clinical practice, chatbots bring about opportunities to gain information about people’s use of chatbot and to intervene accordingly. The point to note is that integrating human support would benefit from models grounded on psychology insights.

Chatbots should be aligned with other Behavioral Intervention Technologies (BITs) insofar as adopting an Efficiency Model of Support, i.e. a conceptual framework to capture the interplay between information and intervention in an efficient way [51]. Here, efficiency is defined as “the ratio of the outcome of an intervention relative to the human resources required to deliver it, since each decision corresponds to supporting that intervention (what, when, how much, who provides it) represents a trade-off between devoting additional resources and accruing additional benefits” [51]. According to this model, de-

cisions should be based on the consideration of why people may fail to benefit from BITs. This model introduces five categories of possible failures. *Usability failures* refer to the design, ease of learning, and ease of use of the technical features. *Engagement failures* occur when the person has the capacity to use a BIT, but does not, for example, because he or she lacks the intention or motivation. *Fit failures* occur when the assigned tool does not meet the patient's needs. *Knowledge failures* occur if a person uses a tool, but does so incorrectly (e.g. patients might need more instruction in order to understand the treatment's rationale), which would impact the effective practice. Lastly, *implementation failures* occur when a person uses the tool within a BIT, but fails to incorporate it into his or her life (i.e. no recurrent usage). Unfortunately, support intervention has been largely overlooked in the reviewed primary studies. Some exceptions include:

- [19], where the chatbot includes a human-based backup "feature" in the process. This permits to respond to patients in need, "specially when they feel trapped at certain point during their treatment"; and
- [53], where authors suggest for chatbots dealing with very sensitive health-conditions like AIDS, the need for users to "request to talk to a trained counselor" at any time, and that the bot should link the patient to a trained professional whenever "a mention was made of thoughts of self-harm or suicidal ideation" (p.81).

So far, most examples are predominantly self-help interventions. In these interventions, support is not a continuous, pro-active process. Rather, a supporting individual (e.g. a physician) may be aware that the patient is using the chatbot and receives timely information to be able to e.g. reinforce use of the chatbot, but the chatbot mostly operates separately from the support. This form of support is intended to address engagement failures. But this might not be enough. The advantage of the Efficiency Model of Support is to provide a framework where chatbot designers might reflect about eventual failures and whether support resources might payoff.

Finally, social-technical implications are hardly mentioned. Monitoring, counseling, preventing and detecting illnesses might well thrive if chatbots could harness full data for wearable sensors, improve their natural language processing skills, or enhance the emotion detection algorithms. These topics are evolving research areas that will be applied, without doubt, in future chatbots. Yet, it remains to be seen if the ethics of delegating human health problems to software programs are acceptable trade-offs. Indeed, when it comes to serious health concerns that require immediate actions, chatbots do not always handle it adequate or timely [44]. This calls for continuous evaluation of chatbot functioning, including problematic responses, privacy breaches, or patient harm, that might eventually leave chatbots in "quarantine". Notice has been given about the deep social implications of bringing automatization to previously predominantly social activities [12]. Given the sensibility of health matters, chatbot studies would benefit from including a social-technical anal-

ysis where potential biases and the impact on the patient-doctor relationship were discussed.

5.2 Threats to validity

Construct validity. This includes two major threats. The first threat is that the research questions may not provide complete coverage of chatbot-mediated change behavior in the health domain. We focused our study on arranging chatbot experiences along three main dimensions: illness, competences and technical enablers. Other dimensions might also be of importance to understand the adoption of chatbots in the health area (e.g. patient age, social setting, etc.). The second threat is about not including all the relevant works in the field. This threat was alleviated by combining several databases and manual searches to selected journals and conferences. We also checked out with related reviews, specifically, the one about Embodied Conversational Agents and the one on Conversational Agents [36, 40].

Internal validity. Limitations mainly come from the categories being used. Firstly, we resort to the taxonomy of [6] to categorize chatbot opportunities to enhance human competences. This is a new taxonomy that has not yet been widely used by academia. Likewise, categories for the technology enablers brought by chatbots, though inspired in other works [34, 46], reflect our understanding as health chatbot developers ourselves. Whereas the vocabulary of clinical medicine is being standardized over a long period of time, chatbots are a recent technology where a nascent terminology is emerging as chatbots become mainstream.

External validity. Our mapping study provides a global view about illnesses, patient competences and chatbot enablers based on 30 primary studies. Kitchenham et al. suggest that 10 papers or fewer are insufficient for an assessment of completeness for a mapping study, whereas 30 papers would be enough [33]. For a mapping study, the aim is to look at the high-level research trends in a broad topic area where completeness might not be so critical as long as the search strategy remains unbiased. In our case, we selected chatbot experiences published between 2014 to 2018. It is unlikely that excluded papers published before 2014 may affect the generalizability of the results. Year 2014 is the one Telegram added support for chatbots development [38], and in this way, recognizing chatbots becoming mainstream.

Conclusion validity. Bias in data collection could hinder reproducibility of our study. We mitigated this threat by establishing a protocol to extract the data of each paper equally. In addition, Table 1 and Table 2 summarize classification decisions for researchers to replicate the study.

6 Conclusion

We classified 30 peer-reviewed articles along the triplet $\langle illness, competence, technicalEnablers \rangle$. In so doing, we provide three initial insights about the

healthcare-chatbot space. First, the two most active health areas are mental & physical wellness and nutritional & metabolic disorders. Second, “affect” and “cognition” are the human competences most sought-after by chatbots to attain change behavior. Finally, “consumability” is the largely mentioned chatbot enabler.

The field is still in a nascent period of investigation but with a large potential to provide a cost-effective way to handle a larger aging population. Based on the surveyed papers, two main considerations arise. First, the need for experiences to be reported along good practices [18] so that they can be compared or replicated. Second, the convenience of including the broader sociological implications brought by chatbots. Though this study focuses on the patient, other main stakeholders are also impacted: caregivers, physicians or relatives should also be surveyed about how chatbots affect their relationship with patients. After all, chatbots are embodiments of behavior-change principles from psychological science.

Table 1 Primary Studies classified by addressed illness and main enablers

Reference	Illness type	Detail	Enablers
Callejas,2014	mental-physical-wellness	mental-physical-wellness	ia, ac, ub, pe
Vuuren,2014	language disorder	aphasia	mm
Griol,2015	neurological disorders	alzheimer	ge, mm
Lisetti,2015	addictions	alcoholism	pe, an
Atay,2016	neurological disorders	dementia	mm, ge
Beun,2016	neurological disorders	insomnia	ia, ac, ub
Caballero,2016	rare-diseases	rare-diseases	ub
Elmasri,2016	addictions	alcoholism	an, pe
Kimani,2016	cardiovascular disorders	atrial-fibrillation	co, ub
Miner,2016	mental-physical-wellness	physical-health	mm, an
Brixey,2017	sexually transmitted diseases	HIV-AIDS	an
Cameron,2017	mental-physical-wellness	stress	mm, pe, ge
Cruz-Sandoval,2017	neurological disorders	dementia	mm, ge
Dubosson,2017	addictions	smoking	sc, pe, as
Fadhil,2017	mental-physical-wellness	healthy-habits	pe, ge
Heerden,2017	sexually transmitted diseases	HIV-AIDS	an
Hoa,2017	mental-physical-wellness	mental-well-being	pe, as
Hsu,2017	nutritional-metabolic-disorders	food-allergies	pe, ub
Jeong,2017	mental-physical-wellness	mental-Well-being	mm, ub
Kowatsch,2017a	nutritional-metabolic-disorders	obesity	sc
Kowatsch,2017b	nutritional-metabolic-disorders	obesity	sc
Oh,2017	neurological disorders	obesity	mm, ub
Cheng,2018	nutritional-metabolic-disorders	diabetes	mm, ge, as, ub, pe
Chung,2018	nutritional-metabolic-disorders	obesity	pe, as
Fernandez-Luque,2018	nutritional-metabolic-disorders	obesity	as, pe
Gabrielli,2018	mental-physical-wellness	healthy-habits	ub, pe
Huang,2018	nutritional-metabolic-disorders	ub, as, pe	obesity
Inkster,2018	mental-physical-wellness	mental-well-being	ub, as, pe, sc
Kobori,2018	sexually transmitted infections	syphilis	an
Roniotis,2018	neurological disorders	depression	mm, ub, ge

Enablers: Accessibility (ac), Anonymity (an), Asynchronicity (as), Gentle learning curve (ge), Installation (in), Instant availability (ia), Multi-modal-input (mm), Platform independence (pi), Personalization (pe), Scalability (sc), Ubiquity (u).

Table 2 Primary Studies, by purpose and leveraged competences

Reference	Purpose	Competences
Callejas,2014	Healthy aging (virtual coach)	m, c, a
Vuuren,2014	Therapy for persons with aphasia (PWA)	b
Griol,2015	Helping People with Alzheimer	c
Lisetti,2015	Motivational interventions (BMI) for behavior change in obesity, to alcohol and drug use	a,b
Atay,2016	How to use chatbots for engaging older community group members (including those with dementia)	a,b
Beun,2016	Enhancing adherence in Behavior Change Insomnia therapy	c,a
Caballero,2016	Diagnosing rare diseases	c
Elmasri,2016	Mental health intervention on 'alcohol drinking habits assessment'	c,a,b
Kimani,2016	Counselling to patients with Atrial Fibrillation	m,c
Miner,2016	Use of voice-based chatbots for responding to Mental Health, Interpersonal Violence, and Physical Health questions	c
Brixey,2017	Sexual health information on HIV/AIDS	c
Cameron,2017	Mental health counselling	c,b
Cruz-Sandoval,2017	Support People with Dementia	a
Dubosson,2017	Support smokers that want to cease	b
Fadhil,2017	Promote healthy and sustainable eating behavior	m,c,a,b
Heerden,2017	Guide users through an AIDS/HIV counselling and testing session	c
Hoa,2017	Conversational agent for promoting mental well-being	a
Hsu,2017	Inform patients with food allergies about nearby suitable restaurants	m,c,b
Jeong,2017	Mental healthcare (depression)	m,a
Kowatsch,2017a	Intervention for the treatment of childhood obesity	m,b
Kowatsch,2017b	Design and Evaluation of a chat app for behavioral intervention in adolescents' obesity	
Oh,2017	Psychiatric Counseling in Mental Healthcare Service Based on Emotional Dialogue Analysis	m,a
Cheng,2018	Type 2 diabetes (The agent can monitor patient's depression level)	m,c
Chung,2018	Healthy habits and obesity diagnose, monitoring, and prevention	m,c,b
Fernandez-Luque,2018	Reminders, recommendations for mothers of children with obesity and overweight	b
Gabrielli,2018	Guide users to make informal and enhanced health decision making, specifically around food choices	c
Huang,2018	Encourage users to reach a weight-loss target, offering personalized diet, eating tips and advice	m,c
Inkster,2018	Encourage users with depression to build emotional resilience skills	a
Kobori,2018	Screening of Sexually Transmitted Infections (syphilis diagnosis)	c
Roniotis,2018	Detecting depression	c

Competences: Monitoring (m), Cognition (c), Affect (a), Behaviour (b).

7 Compliance with Ethical Standards

Conflict of Interest Juanan Pereira declares that he has no conflict of interest. Óscar Díaz declares that he has no conflict of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

References

1. Abashev, A., Grigoryev, R., Grigorian, K., Boyko, V.: Programming tools for messenger-based chatbot system organization: Implication for outpatient and translational medicines. *BioNanoScience* **7**(2), 403–407 (2017). DOI 10.1007/s12668-016-0376-9
2. Andersson, G., Cuijpers, P.: Internet-Based and Other Computerized Psychological Treatments for Adult Depression: A Meta-Analysis. *Cognitive Behaviour Therapy* **38**(4), 196–205 (2009). DOI 10.1080/16506070903318960

3. Atay, C., Ireland, D., Liddle, J., Wiles, J., Vogel, A., Angus, D., Bradford, D., Campbell, A., Rushin, O., Chenery, H.J.: Can a smartphone-based chatbot engage older community group members? The impact of specialised content. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association* **12**(7), P1005–P1006 (2016). DOI 10.1016/j.jalz.2016.06.2070
4. Beun, R.J., Brinkman, W.P., Fitrianie, S., Griffioen-Both, F., Horsch, C., Lancee, J., Spruit, S.: Improving adherence in automated e-coaching. In: *International Conference on Persuasive Technology*, pp. 276–287. Springer (2016). DOI 10.1007/978-3-319-31510-2_24
5. Bickmore, T.W., Puskar, K., Schlenk, E.A., Pfeifer, L.M., Sereika, S.M.: Maintaining reality: Relational agents for antipsychotic medication adherence. *Interacting with Computers* **22**(4), 276–288 (2010). DOI 10.1016/j.intcom.2010.02.001
6. Brinkman, P.: Virtual Health Agents for Behavior Change: Research Perspectives and Directions. In: *Proceedings of the Workshop on Graphical and Robotic Embodied Agents for Therapeutic Systems (GREATS16) Held during the International Conference on Intelligent Virtual Agents (IVA16)* (2016)
7. Brixey, J., Hoegen, R., Lan, W., Rusow, J., Singla, K., Yin, X., Artstein, R., Leuski, A.: SHIHbot: A Facebook chatbot for Sexual Health Information on HIV/AIDS. In: *Proceedings of the 18th Annual SIGdial Meeting on Discourse and Dialogue*, pp. 370–373 (2017). DOI 10.18653/v1/W17-5544
8. Callejas, Z., Griol, D., McTear, M.F., López-Cózar, R.: A Virtual Coach for Active Ageing Based on Sentient Computing and m-health. In: *International Workshop on Ambient Assisted Living*, pp. 59–66. Springer (2014). DOI 10.1007/978-3-319-13105-4_10
9. Cameron, G., Cameron, D., Megaw, G., Bond, R., Mulvenna, M., O'Neill, S., Armour, C., McTear, M.: Towards a chatbot for digital counselling. *Journal of Medical Internet Research* **4**(1), e3 (2017)
10. Cheng, A., Raghavaraaju, V., Kanugo, J., Handrianto, Y.P., Shang, Y.: Development and evaluation of a healthy coping voice interface application using the Google home for elderly patients with type 2 diabetes. In: *Consumer Communications & Networking Conference (CCNC), 2018 15th IEEE Annual*, pp. 1–5. IEEE (2018). DOI 10.1109/CCNC.2018.8319283
11. Chung, K., Park, R.C.: Chatbot-based healthcare service with a knowledge base for cloud computing. *Cluster Computing* pp. 1–13 (2018). DOI 10.1007/s10586-018-2334-5
12. Crawford, E.: Bots Are Awesome! Humans? Not so Much. <https://chatbotsmagazine.com/bots-are-awesome-humans-not-so-much-7b2d62630668> (2016)
13. Cruz-Sandoval, D., Favela, J.: Semi-autonomous Conversational Robot to Deal with Problematic Behaviors from People with Dementia. In: *International Conference on Ubiquitous Computing and Ambient Intelligence*, pp. 677–688. Springer (2017). DOI 10.1007/978-3-319-67585-5_66
14. Cruzes, D.S., Dyba, T.: Recommended steps for thematic synthesis in software engineering. In: *2011 International Symposium on Empirical Software Engineering and Measurement*, pp. 275–284. IEEE (2011). DOI 10.1109/ESEM.2011.36
15. Dale, R.: The return of the chatbots. *Natural Language Engineering* **22**(5), 811–817 (2016). DOI 10.1017/S1351324916000243
16. Dubosson, F., Schaer, R., Savioz, R., Schumacher, M.: Going beyond the relapse peak on social network smoking cessation programmes: ChatBot opportunities. *Swiss medical informatics* **33** (2017)
17. Elmasri, D., Maeder, A.: A conversational agent for an online mental health intervention. In: *International Conference on Brain and Health Informatics*, pp. 243–251. Springer (2016). DOI 10.1007/978-3-319-47103-7_24
18. Eysenbach, G., Group, C.E.: CONSORT-EHEALTH: Improving and standardizing evaluation reports of Web-based and mobile health interventions. *Journal of medical Internet research* **13**(4) (2011). DOI 10.2196/jmir.1923
19. Fadhil, A.: A Conversational Interface to Improve Medication Adherence: Towards AI Support in Patient's Treatment. arXiv preprint arXiv:1803.09844 (2018)

20. Fadhil, A., Gabrielli, S.: Addressing challenges in promoting healthy lifestyles: The ai-chatbot approach. In: Proceedings of the 11th EAI International Conference on Pervasive Computing Technologies for Healthcare, pp. 261–265. ACM (2017). DOI 10.1145/3154862.3154914
21. Fadhil, A., Villaforita, A.: An adaptive learning with gamification & conversational UIs: The rise of CiboPoliBot. In: Adjunct Publication of the 25th Conference on User Modeling, Adaptation and Personalization, pp. 408–412. ACM (2017). DOI 10.1145/3099023.3099112
22. Fernandez-Luque, L., Lattab, A., Hors, S., Ahmedna, M.: Implementation and feasibility study of a tailored health education bot in Telegram for mothers of children with obesity and overweight. In: Qatar Foundation Annual Research Conference Proceedings, vol. 2018, p. HBPD506. HBKU Press Qatar (2018)
23. Fogg, B.J.: Persuasive technology: Using computers to change what we think and do. *Ubiquity* **2002**(December), 5 (2002). DOI 10.1145/764008.763957
24. Gabrielli, S., Marie, K., Corte, C.D.: SLOWBot (Chatbot) Lifestyle Assistant. In: Proceedings of the 12th EAI International Conference on Pervasive Computing Technologies for Healthcare, PervasiveHealth '18, pp. 367–370. ACM, New York, NY, USA (2018). DOI 10.1145/3240925.3240953
25. Griol, D., Molina, J.M.: An Ambient Assisted Living Mobile Application for Helping People with Alzheimer. In: International Conference on Practical Applications of Agents and Multi-Agent Systems, pp. 3–14. Springer (2015). DOI 10.1007/978-3-319-19033-4_1
26. Hsu, P., Zhao, J., Liao, K., Liu, T., Wang, C.: AllergyBot: A Chatbot Technology Intervention for Young Adults with Food Allergies Dining Out. In: Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems, pp. 74–79. ACM (2017). DOI 10.1145/3027063.3049270
27. Huang, C., Yang, M., Huang, C., Chen, Y., Wu, M., Chen, K.: A Chatbot-supported Smart Wireless Interactive Healthcare System for Weight Control and Health Promotion. In: 2018 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), pp. 1791–1795 (2018). DOI 10.1109/IEEM.2018.8607399
28. Inkster, B., Sarda, S., Subramanian, V.: An Empathy-Driven, Conversational Artificial Intelligence Agent (Wysa) for Digital Mental Well-Being: Real-World Data Evaluation Mixed-Methods Study. *JMIR mHealth and uHealth* **6**(11), e12106 (2018). DOI 10.2196/mhealth.12106
29. Isern, D., Moreno, A.: A systematic literature review of agents applied in healthcare. *Journal of medical systems* **40**(2), 43 (2016). DOI 10.1007/s10916-015-0376-2
30. Ivarsson, M., Gorschek, T.: A method for evaluating rigor and industrial relevance of technology evaluations. *Empirical Software Engineering* **16**(3), 365–395 (2011). DOI 10.1007/s10664-010-9146-4
31. Jeong, S., Breazeal, C.: Toward Robotic Companions that Enhance Psychological Well-being with Smartphone Technology. In: Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction, pp. 345–346. ACM (2017). DOI 10.1145/3029798.3034803
32. Kimani, E., Bickmore, T., Trinh, H., Ring, L., Paasche-Orlow, M.K., Magnani, J.W.: A Smartphone-Based Virtual Agent for Atrial Fibrillation Education and Counseling. In: International Conference on Intelligent Virtual Agents, pp. 120–127. Springer (2016). DOI 10.1007/978-3-319-47665-0_11
33. Kitchenham, B.A., Budgen, D., Brereton, P.: Evidence-Based Software Engineering and Systematic Reviews, vol. 4. CRC press (2015)
34. Klopfenstein, L.C., Delpriori, S., Malatini, S., Bogliolo, A.: The Rise of Bots: A Survey of Conversational Interfaces, Patterns, and Paradigms. In: Proceedings of the 2017 Conference on Designing Interactive Systems, DIS '17, pp. 555–565. ACM, New York, NY, USA (2017). DOI 10.1145/3064663.3064672
35. Kobori, Y., Osaka, A., Soh, S., Okada, H.: Novel application for sexual transmitted infection screening with an ai chatbot. *The Journal of Urology* **199**(4, Supplement), e189–e190 (2018). DOI 10.1016/j.juro.2018.02.516
36. Kowatsch, T., Ni's sen, M., Shih, C.H.I., Rügger, D., Volland, D., Filler, A., Künzler, F., Barata, F., Hung, S., Büchter, D.: Text-based Healthcare Chatbots Supporting Patient and Health Professional Teams: Preliminary Results of a Randomized Controlled

- Trial on Childhood Obesity. In: *Persuasive Embodied Agents for Behavior Change (PEACH2017)*. ETH Zurich (2017)
37. Kowatsch, T., Volland, D., Shih, I., Rügger, D., Künzler, F., Barata, F., Filler, A., Büchter, D., Brogle, B., Heldt, K.: Design and Evaluation of a Mobile Chat App for the Open Source Behavioral Health Intervention Platform MobileCoach. In: *International Conference on Design Science Research in Information Systems*, pp. 485–489. Springer (2017). DOI 10.1007/978-3-319-59144-5_36
 38. Kozinakova, B.: Analysis of chatbot systems focusing on the elderly as users. Master Thesis, Politecnico de Milano (2017)
 39. Lambert, A.O.C., Montañez, C.H.T., Martinez, M.B., Funes-Gallanzi, M.: A Conversational Agent for Use in the Identification of Rare Diseases. In: *Applications for Future Internet*, pp. 128–139. Springer (2017)
 40. Laranjo, L., Dunn, A.G., Tong, H.L., Kocaballi, A.B., Chen, J., Bashir, R., Surian, D., Gallego, B., Magrabi, F., Lau, A.: Conversational agents in healthcare: A systematic review. *Journal of the American Medical Informatics Association* (2018). DOI 10.1093/jamia/ocy072
 41. Lisetti, C., Amini, R., Yasavur, U.: Now all together: Overview of virtual health assistants emulating face-to-face health interview experience. *KI-Künstliche Intelligenz* **29**(2), 161–172 (2015). DOI 10.1007/s13218-015-0357-0
 42. Ly, K.H., Ly, A.M., Andersson, G.: A fully automated conversational agent for promoting mental well-being: A pilot RCT using mixed methods. *Internet Interventions* **10**, 39–46 (2017). DOI 10.1016/j.invent.2017.10.002
 43. Miner, A., Chow, A., Adler, S., Zaitsev, I., Tero, P., Darcy, A., Paepcke, A.: Conversational Agents and Mental Health: Theory-Informed Assessment of Language and Affect. In: *Proceedings of the Fourth International Conference on Human Agent Interaction*, pp. 123–130. ACM (2016). DOI 10.1145/2974804.2974820
 44. Miner, A.S., Milstein, A., Schueller, S., Hegde, R., Mangurian, C., Linos, E.: Smartphone-based conversational agents and responses to questions about mental health, interpersonal violence, and physical health. *JAMA internal medicine* **176**(5), 619–625 (2016). DOI 10.1001/jamainternmed.2016.0400
 45. Oh, K.J., Lee, D., Ko, B., Choi, H.J.: A Chatbot for Psychiatric Counseling in Mental Healthcare Service Based on Emotional Dialogue Analysis and Sentence Generation. In: *Mobile Data Management (MDM), 2017 18th IEEE International Conference On*, pp. 371–375. IEEE (2017). DOI 10.1109/MDM.2017.64
 46. Pereira, J., Díaz, O.: Chatbot Dimensions that Matter: Lessons from the Trenches. In: *Web Engineering, Lecture Notes in Computer Science*, pp. 129–135. Springer, Cham (2018). DOI 10.1007/978-3-319-91662-0_9
 47. Petersen, K., Feldt, R., Mujtaba, S., Mattsson, M.: Systematic Mapping Studies in Software Engineering. In: *Proceedings of the 12th International Conference on Evaluation and Assessment in Software Engineering, EASE’08*, pp. 68–77. BCS Learning & Development Ltd., Swindon, UK (2008)
 48. Richards, D., Caldwell, P.: Improving health outcomes sooner rather than later via an interactive website virtual specialist. *IEEE Journal of Biomedical and Health Informatics* pp. 1–1 (2017). DOI 10.1109/JBHI.2017.2782210
 49. Richards, D., Caldwell, P.H.: Gamification to Improve Adherence to Clinical Treatment Advice. *Health Literacy: Breakthroughs in Research and Practice: Breakthroughs in Research and Practice* p. 80 (2017). DOI 10.4018/978-1-5225-1928-7.ch005
 50. Roniotis, A., Tsiknakis, M.: Detecting Depression Using Voice Signal Extracted by Chatbots: A Feasibility Study. In: *Interactivity, Game Creation, Design, Learning, and Innovation*, pp. 386–392. Springer (2017)
 51. Schueller, S.M., Tomasino, K.N., Mohr, D.C.: Integrating Human Support Into Behavioral Intervention Technologies: The Efficiency Model of Support. *Clinical Psychology: Science and Practice* **24**(1), 27–45 (2017). DOI 10.1111/cpsp.12173
 52. Stratou, G., Morency, L.P., DeVault, D., Hartholt, A., Fast, E., Lhommet, M., Lucas, G., Morbini, F., Georgila, K., Scherer, S., Gratch, J., Marsella, S., Traum, D., Rizzo, A.: A demonstration of the perception system in SimSensei, a virtual human application for healthcare interviews. In: *2015 International Conference on Affective Computing and Intelligent Interaction (ACII)*, pp. 787–789 (2015). DOI 10.1109/ACII.2015.7344661

-
53. van Heerden, A., Ntinga, X., Vilakazi, K.: The potential of conversational agents to provide a rapid HIV counseling and testing services. In: *The Frontiers and Advances in Data Science (FADS)*, 2017 International Conference On, pp. 80–85. IEEE (2017). DOI 10.1109/FADS.2017.8253198
 54. Van Vuuren, S., Cherney, L.R.: A virtual therapist for speech and language therapy. In: *International Conference on Intelligent Virtual Agents*, pp. 438–448. Springer (2014). DOI 10.1007/978-3-319-09767-1_55