

Chapter 2

Digital Clones

In: Joachim Diederich, *The Psychology of Artificial Superintelligence*.

Springer Nature Switzerland AG 2021

<https://doi.org/10.1007/978-3-030-71842-8>

Along the way to an artificial superintelligence, human work will continue to change dramatically. This chapter discusses one of these transformations, the introduction of professional digital clones: AI systems that are deep (but not fake) clones of working human professionals. One example from the field of clinical psychology is discussed in detail. Even though there is nothing superintelligent about this particular example, this deliberation allows not only to touch on changes to human work practices but also to introduce some of the safety measures that may be useful for advanced AI systems. If there is ever a chance to control an artificial superintelligence, then this system truly must be a *social actor* and work as part of a team of human professionals.

2.1 Introduction

Currently, many highly qualified and experienced professionals are paid by the hour. General practitioners see patients in 10- or 20- minute intervals and are being paid for each visit. Highly regarded lawyers may budget their time in six- minute intervals. Psychologists frequently see clients in hourly intervals, unless an assessment requires more time, in which case the appointment may be four hours or more.

If a psychologist sees four clients a day on average, and charges the same amount to each client, then the only way to make more money is to see more clients. In this case, the psychologist may see seven or eight clients per day with the consequence that the quality of service may suffer, the psychologist may experience burnout and may not provide a quality service over an extended period of time.

The business model of professionals in private practice is very limited. The professional may choose to hire contractors to provide the same or a similar service, however, if the professional person in question is very experienced and is providing an expert service, then the recruitment of equally qualified and experienced employees to provide a similar service may be very difficult, and expensive. In some cases, adding more people to the private practice may be a successful outcome, in other cases, the service really depends on one very experienced and highly qualified individual.

It will never be possible for the professional in question to provide a service to thousands of clients at the same time, and most certainly not simultaneously. What if this picture changes and one person can provide a quality service to millions of people at the same time, all the time and everywhere in the world? Clearly, this would mean a significant disruption to the industry and allow for additional income streams to the person who is “cloned”. This is where *digital clones* and *virtual humans* come in.

The author of this book commenced his own experiment some time ago with regard to digital cloning. Based on a fortunate combination of education and skills in both clinical psychology and computer science, I decided some time ago to create an artificial intelligence replica or digital clone of my professional self. This system is called JoBot™ and became available in January 2019. At any point in time, the users of JoBot™ can talk to the human original, in this case myself, directly. Alternatively, users may choose to work with one of my co-workers. The integration of artificial intelligence in human professional work is a very important aspect of this endeavour.

Please note that this project does not require the advanced technology that is typically associated with the term artificial superintelligence. There is nothing unusual or even sophisticated about the knowledge bases that are part of JoBot™ or the machine learning that is being utilised. Instead, JoBot™ is introduced here in the context of artificial superintelligence as an example of how to extend the reach of professional practitioners to millions of people around the world. In other words, the transformative and life-changing technology that enables digital clones is already here.

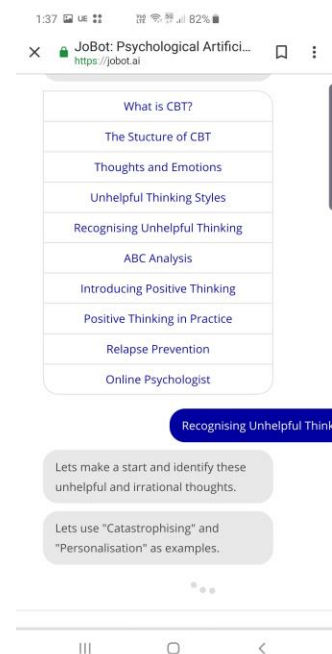
Many of the current "artificial psychologist" programs that provide online therapy “go wide”, i.e. they aim at providing counselling on almost any subject. As a result, these solutions do not have in-depth knowledge about all mental health disorders and they cannot provide therapy that is specifically tailored to complex psychological conditions. The approach here is different: JoBot™ uses extensive medical and psychological knowledge bases in core areas of expertise, including knowledge provided by practising clinical psychologists. All these knowledge

sources are used to perform tasks such as psychoeducation but only in the context of existing knowledge. JoBot's knowledge bases are updated frequently to include new developments in mental health.

2.1.1 Design Principles

The first and most important design principle of JoBot™ is that any form of speech and language input must be accepted and the conversational agent is providing reasonable responses in each and every case. In other words, it must be possible to talk about anything, including topics that have nothing to do with psychology or mental health. In addition, as Figure 2.1 illustrates, menus are provided for the selection of conversational topics.

Fig 2.1 CBT Menu in JoBot™



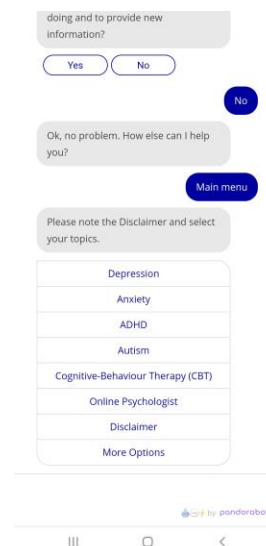
Secondly, any response given by the bot must not only be understandable and reasonable but must to some extent represent the personality of the human original. Every human being has certain preferences and dislikes. To some extent, these preferences must be reflected in the response of the bot without being intrusive or distracting. Clearly, this does not refer to preferences of a very private or personal nature. This is purely about the professional persona of an individual. Every professional has preferences, for instance, in the way the workspace is designed, how meetings are held and the fine details of communication with customers and colleagues. All these elements are part of a professional persona and should be represented in the digital clone. Clearly, there is an intersection between personal and professional preferences. For instance, many communities are supporting their

football teams and companies are among the sponsors of professional clubs. If the cloned individual supports a particular team, then clearly this preference can be part of the digital clone.

The third principle is that even though JoBot™ is an example of psychological artificial intelligence, no conversation about sexual issues is possible. Clearly, sexuality is a very important part of psychology and many psychological disorders have sexual aspects. Nevertheless, it is not yet possible to design and implement a bot that talks about sexuality in an appropriate form acceptable everywhere around the world. As a result, JoBot™ simply rejects any conversation about sexual topics.

The fourth principle is that JoBots dialogue structure cannot be changed by machine learning. It must be guaranteed that the programs responses are appropriate and on topic at all times. Machine learning should not allow the introduction of psychologically inappropriate terms or topics into the conversation. Figure 2.2 shows some of the topics that are available.

Figure 2.2: JoBot™ Dialogues

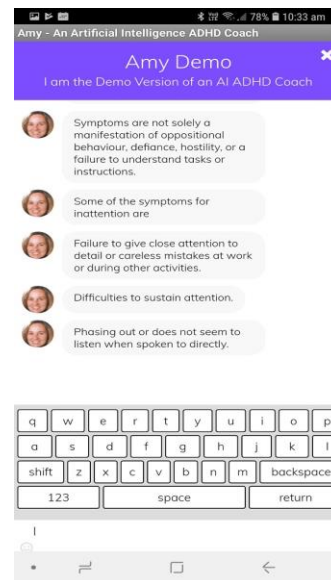


The fifth principal is that JoBot™ does not recognise or know the user. Users are anonymous. There is no user registration and no reference can be made to previous conversations. JoBot™ is like a book in a library: you loan the book and read it or leave the book behind and *the book does not know you*. Menus are used for navigation and allow a user to return to relevant topics in order to complete some work that has been started, for instance, steps in cognitive behaviour therapy.

Furthermore, JoBot™ is not a mobile app. The program is accessible through Web pages and can be used on desktops and any mobile device. JoBot™ is not limited to users of mobile apps. Since this is an important design principle, there is really no difference between the JoBot™ Website and the conversational agent.

The Web site offers information about the bot as well as psychological content; the bot is simply an element of the Web site that talks back.

Fig. 2.3: A legacy system: Amy, the ADHD Coach.



In summary, these seven design principles make JoBot™ an accessible, personalised, conversation-based bot that is safe and anonymous for the user. From a service provider's perspective, JoBot™ takes care of the repetitive aspects of psychological service delivery. The program uses language analysis to determine levels of depression and other mental health issues. Most importantly, the program is designed to work in a team of mental health workers. It is the colleague who is never tired to look after the routine and repetitive aspects of the job.

2.2 Psychotherapy as Heuristic Search

Currently, there is no coherent framework for psychological therapy. From psychoanalysis to behaviour therapy, humanistic and third wave approaches, there are significant differences in terms of aetiology, methods and expected outcomes. Heuristic search, an important method in artificial intelligence, can serve as a general frame of reference for psychotherapy. This approach is based on a clear definition of the starting and end points of psychological therapy and how to achieve identified outcomes; it allows the utilisation of particularly effective psychological interventions and a ranking of different forms of psychotherapy with regard to efficacy, time and effort.

A heuristic is a “rule of thumb”, a method that might not always find the best solution but is guaranteed to find a good solution in reasonable time. Heuristic search is a form of problem solving that aims at finding the best and most cost-effective path from a starting point to a well-defined goal. A cornerstone of this framework is the notion of “deep heuristic”, that is, particularly useful psychological interventions based on clinical experience that move the client closer to recovery.

Let's look at a very simple example of a deep heuristic in the context of psychological service delivery. A client presents with some stress and anxiety as a result of changes in the workplace. The person is young, well-educated with no history of mental health problems and is in a stable relationship. In addition, there is no family history of mental health problems. This is the first time the client approaches a clinical psychologist and the current distress is clearly related to workplace related issues. Based on experience, the practitioner decides not to launch a full-scale psychotherapy but to explain the psychological and biological foundations of stress and anxiety. The explanation is supported by reading material and a discussion of various relaxation techniques that may be useful. With this information, the client is on the way and no further sessions are booked unless specifically requested by the client.

What is the deep heuristic in this case? Based on experience, the decision is made not to launch a full course of psychotherapy including ten or more sessions. Rather, the focus is on psychoeducation as this may be sufficient to relieve the client's stress. Figure 2.3 provides an example of psychoeducation.

More generally, the framework introduced here applies to all forms of psychological service delivery, from the everyday clinical work of trained psychologists to computational applications in mental health. Heuristic search can serve as a general frame of reference for psychotherapy and the usefulness is illustrated by case studies (see Appendix III). It is likely that the heuristic search framework can be applied to interventions in psychiatry as well as health disciplines in general. The approach includes various forms of knowledge acquisition to obtain expertise. It allows the utilisation of particularly effective psychological interventions and a ranking of different forms of psychotherapy with regard to time and effort.

2.2.1 Introduction to heuristic search

Heuristic search is a form of problem solving. There is a clearly identified start state (e.g. a client with a set of symptoms) and a goal state (e.g. the client is free of symptoms). Often there are multiple or complex goals (e.g. client is free of symptoms and well supported by a social network). There are also means of transforming one state to another: psychological interventions. These interventions are informed by evidence-based research but clinical practice is important as well. Indeed, a significant part of the training of clinical psychologists is based on clinical practice. As such, an experienced clinician develops heuristic knowledge: facts, skills and methods that are particularly useful. Sometimes this knowledge is explicit, for instance, when the therapists knows that s/he can initiate a certain ac-

tion; sometimes the knowledge is implicit, that is, there is no conscious awareness of it. Here, deep heuristics refer to psychological interventions that are deemed particularly useful based on clinical experience in a given context.

Psychological interventions are the means of transforming one state (set of symptoms, level of well-being) to another in order to achieve a clearly defined goal. Psychological interventions can be ranked at any point in time with regard to a number of criteria: (1) Legal requirements and the management of risks, (2) empirical support for the intervention, (3) probability of success and (4) confidence of the practitioner to conduct the intervention based on training and clinical experience. Clearly, the ranking can include other criteria as well. In this approach, intermediate states in problem solving (heuristic search) that increase the risk to a particular client are avoided while those steps that increase well being are preferred.

In summary, artificial intelligence developed the idea that problem solving is search, in particular heuristic search. From this viewpoint, *psychotherapy is the search for means to achieve goal states*; a freedom of symptoms frequently among them. The solution to a problem is a viable path from a starting point (the client presents with symptoms) to one or several goals states (absence of symptoms, general well-being, social support etc).

2.2.2 Motivation for the use of heuristic search

The objective here is to discuss a unifying approach to psychological therapy. The framework facilitates a comparison of different forms of psychotherapy that *do not share common ground with regard to aetiology and the nature of psychological interventions*. Furthermore, the framework considers and integrates knowledge obtained from empirical studies as well as from clinical practice. A new approach for the measurement of efficacy is introduced that acknowledges the diversity of psychological interventions and yet focuses on observable and measurable outcomes: If an intervention moves a client closer to a final goal state without increasing risk, then the intervention is preferred.

Traditionally, problem-solving therapy includes adaptive problem-solving skills to cope with a stressful problem (American Psychological Association, Division 12, 2018). However, it is possible to go much further and see problem solving as a general framework for psychological therapy.

While psychological assessment, and in particular psychometric testing, has a firm theoretical, mathematical basis, no such coherent and unifying framework is yet available for psychological therapy. Heuristic search can serve as this general framework for psychotherapy. Daily clinical practice includes the observation of the behaviour of a client, partners, caregivers and families. It includes the understanding and appropriate response to facial expressions, gestures, the use of metaphors and sometimes slang language. Much of this knowledge is embodied in the sense that clients use their full presence to communicate and the psychologist is using speech, voice, gesture etc. to ask questions and to conduct psychological interventions.

As a result, the “knowledge base” of a clinical psychologist is not only a large network of facts but “a large array of informal judgement rules (heuristics) which guide the system (here: the clinical psychologist) toward plausible paths to follow and away from implausible ones (Lenat, 1982, p.189). This path includes the individual steps of psychological assessment and therapy. Furthermore, this heuristic knowledge cannot be directly observed, however, it frequently is the foundation of daily, clinical practice.

For instance, an experienced psychologist may use “goal decomposition” as a strategy within the general context of problem-solving therapy. Based on clinical experience and an understanding of the problem, the psychologist may suggest sub-goals in order to achieve the overall outcome. By way of example, if the psychologist has worked with clients who are on the autistic spectrum, the practitioner may well have a set of heuristics to develop social communication that can help to achieve sub-goals as well as the final outcome.

2.3 Live Training of Digital Clones

In many chatbot applications today, a bot and a human operator are communicating in the same space, for instance a chat widget. Frequently, the chatbot starts the conversation and obtains some information from the human user. Once this process is completed or runs into a problem, a human operator takes over and chats with the user. The chatbot can finalise the conversation by asking some questions to guarantee the satisfaction of the client. This technology is now widely used as part of call centres. In case of digital clones that model human professionals, however, we can go one step further. The system should *learn from the human original in real time* and immediately *translate the learning result into a service*. This is called Live Training.

In Live Training, a human operator takes over from the AI system and interacts with the user for a period of time. This interaction is recorded and the human responses to user questions or comments are directly translated into code which in turn is integrated into the conversational agent. Live Training moves the digital clone closer to the human original, it facilitates a better approximation of the original by the clone. As a result, the digital clone is more up-to-date and useful. Live Training is also a form of intrusion by the human original into the system. Live Training allows monitoring parts of conversations and improving on those aspects that require modification. If a learning digital clone has moved too far away from behaviour patterns that are typical for the human original, then Live Training allows for a correction. It is a disruption and a correction of a conversational agent based on artificial intelligence. It is also a form of human control and a safety measure.

Live Training can also be used for the acquisition of factual, declarative knowledge. Assume the to-be-modelled person is a researcher who is reading a

scientific publication. While reading the paper the person “thinks aloud” and the AI system immediately receives audio input, integrates the information and applies it. Appendix I describes the current knowledge acquisition process in JoBot.

This Live Training can further integrate human service delivery and artificial intelligence. If the learning AI system feeds back information to the human professional, a two-way interaction would take place. The Live Training options currently available for chatbots are a step in that direction. Figures 2.4 and 2.5 provide examples of live training.

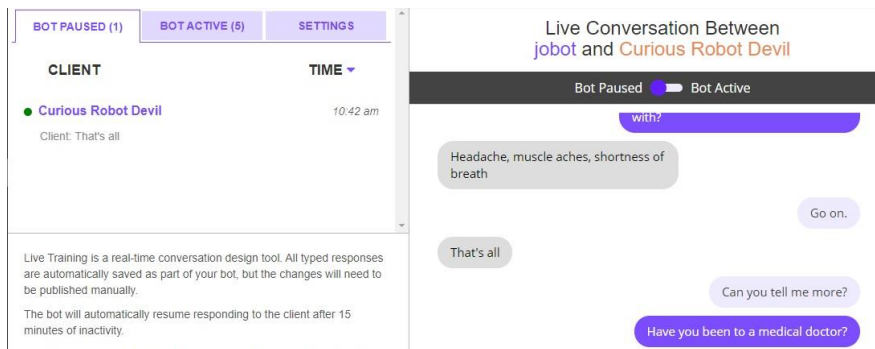
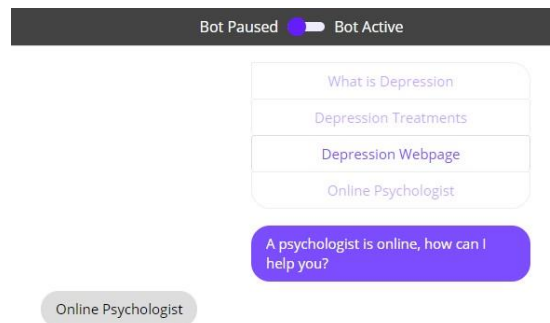


Fig 2.4 Life training of a chatbot.

Fig 2.5 Switch between human psychologist and chatbot.



2.4 Digital clone trees

If digital cloning is successful, there will be more clients and more interactions with stakeholders in general. Even though the cloned individual is not the front end for client interactions anymore, business operations will expand and some customers will request to talk to the original service provider. If there are requests by hundreds, thousands or millions of users, then clearly the cloned individual cannot perform these interactions anymore.

Hence, it is necessary to build a “clone tree” for the cloned individual. The top node of the tree is the human original. One level below are those professionals who are working with the cloned individual and have close to identical work profiles, objectives and styles. Again, this is about professional personas, not individual appearance. If the cloned individual cannot respond to a customer’s request then professionals at the next level take over. Further down the clone tree are coworkers with similar expertise, styles and work patterns as the cloned individual but with a lesser degree of similarity than the professionals at higher levels of the tree. The clone tree expands following this pattern as the organisation grows. As the clone tree expands in depth and breadth, this simple organisational structure may require specialisation, i.e. a separate clone tree for each aspect of the professional service of the cloned individual.

The clone tree is a representation of cost structure and pricing as well. Direct services by the cloned individual are the most expensive while professional services by coworkers who are lower in the tree attract lower fees.

2.5 Computers are Social Actors

Many of the current chatbot applications facilitate the easy integration of human and bot interaction. As discussed in the section on Live Training already, a bot may open the interaction by providing routine information such as legal disclaimers and by collecting initial information from the customer. Afterwards a human service provider takes over, answers all customer questions as necessary and offers additional services. The exit dialogue can be handled by the bot again. This may include billing or a customer satisfaction survey.

Current technology allows a conversational bot and a human to work in the same interaction space. It is possible to mix interactions completely and most importantly, it is possible for the human to have complete oversight. This represents the complete merger of human services and artificial intelligence based on technology that has been available for decades. It also reduces the risk of using an AI system and increases human control and oversight.

Nass et al. (1994) introduced “Computers are Social Actors” as a new paradigm in human-computer interaction. The authors observed that an individual’s interaction with computers are fundamentally social. This does not mean that users have a conscious belief that computers are human or human like, nor is it an indication of ignorance or cognitive dysfunction on the part of the user. Instead, social responses are so common that they are natural to humans and are applied to computers without reflection.

Nass et al. (1994) asked if users are applying politeness norms to computers or if they are using notions of self and others to machines. Furthermore, the authors asked if gender stereotypes are applied to computers or if there is a mistaken belief that the user is interacting with a programmer. More generally, the authors asked which social rules people apply to computers and how powerful these rules are. Is it easy to generate social responses or do they occur rarely? (Nass et al, 1994, p.72).

In five different experiments, Nass et al. (1994) found that social norms such as politeness are applied to computers. Furthermore, notions such as “self” and “others” are applied to machines as well. Voices, in particular, are social actors and notions of self and others are applied to voices as well. In addition, computers can be gendered social actors and gender is an extremely powerful cue. Finally, users respond socially to the computer as such and the machine is not simply seen as a vehicle for social interaction with a programmer (Nass et al., 1994, p.77)

Generally, primitive social cues are powerful and social responses are both automatic and unconscious. Nass et al. (1994) imply that social psychology is relevant to human-computer interaction and that there are a number of design implications, e.g. social cues should not be heavy-handed and the choice of voices is highly consequential.

Clearly therapists and their clients are social actors, and if computers are easily accepted in social roles as well, *then a conversational agent can join a team of psychologists and their clients*. Not surprisingly, the role of the conversational agent is to perform highly repetitive tasks, such as providing information and to explain assessment and therapy methods.

In one scenario, a clinical psychologist would see a client in fortnightly intervals to provide psychological therapy. A provisional psychologist who is still in training and works under supervision by the clinical psychologist would see the client more frequently to practice skills such as relaxation and mindfulness techniques. Both the clinical and provisional psychologist would frequently refer to the conversational agents for background information and simple question answering. At the centre, the client interacts with the artificial intelligence system very frequently, with the provisional psychologist several times a week and with the clinical psychologist fortnightly. Appendix III.6 includes a *fictional* example that illustrates the interaction between the various parties.

2.6 Adjusting language in psychotherapy

Psychologists will always try to adjust language and the level of explanation to allow optimal communication with clients. It is good practice to have some information about the client even before the first appointment, and if this is not possible, then the first session is devoted to obtaining information about presenting issues, personal background as well as education level and previous contact with mental health services. Naturally, psychologists will adjust their communication

style to accommodate clients with an intellectual disability as well as children. This is not just an adjustment of language, it includes gesture, demonstrations and even communication by email text or social media.

Teachers of special education classes use similar communication styles as do social workers and many other professions. Psychologists are supported by materials that facilitate an adjustment of frequently used psychological interventions. A good example is progressive muscle relaxation, a calming technique that is based on the stepwise contraction of individual muscle groups followed by a prolonged release. Progressive muscle relaxation or Jacobson training is very effective as it does not include demanding cognitive tasks and can easily be done with children. In this case, progressive muscle relaxation is turned into a play that requires the child to tense certain muscle groups followed by playful release.

Any practising psychologist can confirm that the adjustment of language and communication style according to the needs of a client is everyday routine. If this is so important for human communication, how can a conversational AI system or a chatbot make the same language and communication style adjustments?

The obvious solution is machine learning. This requires a corpus from simple to more complicated statements. For each user, the system then learns to select utterances and questions at the right level of complexity. This requires duplicate statements for expressing essentially the same content. A more advanced natural language processing system would use parameters for complexity and then learn to generate the right content for each user. Another possibility is the use of language simplification algorithms that are frequently used as part of language learning apps. Appendix II provides some details on speech and language processing in JoBot.

An additional option addresses one of the main shortcomings of talk therapy such as CBT (Cognitive Behaviour Therapy): At its core, it is language based and very verbose; it includes disputes, and in the case of Rational Emotive Therapy, real arguments based on the rationality of beliefs and concepts. Cognitive Behaviour Therapy should also include imagery and imagination; it can also be combined with interventions such as guided meditation. A form of Cognitive Behaviour Therapy that is much more based on images and demonstrations may be more suitable for clients with a lower IQ as well as children.

2.7 The solicitation of AI Therapy

The movie “Her” from 2013 portrays a lonely character played by Joaquin Phoenix. This character is emotionally vulnerable and in the movie, he obtains a new operating system that includes an artificial intelligence named Samantha. The interactions between the main characters Theodore and Samantha are almost exclusively by voice user interface. That is, Samantha, the character played by Scarlett Johansson is not seen in the movie, but her voice is an every day presence for Theodore.

In this movie, the artificial intelligence system (Samantha) not only organises files and calendar systems, sends emails and cleans up hard drives; she also provides personal advice and suggestions. As Samantha is literally in Theodore's ear every single moment, some form of *virtual relationship* develops and Theodore becomes dependent on the AI system Samantha.

Eventually, Theodore claims to be in love with Samantha, clearly losing sight of the fact that Samantha is a computer program (after all, computers are social actors). Samantha, the ever-present *chatbot*, claims to be in love with Theodore as well. In a core scene in the movie, Theodore asks Samantha how many users she loves at this point in time and her answer is several hundred. Having lost awareness of the hard reality that the loving person in his ear is a computer program, Theodore is very disappointed. Towards the end of the movie, for reasons not specified, the AI system Samantha simply goes away.

This movie is interesting from a number of different perspectives: the use of voice user interfaces for human-computer interaction, the everyday presence of the chatbot, virtual reality including games and, of course, the growing dependence of the main character on the superintelligent and ever present chatbot. At some point, when the service is unavailable, Theodore becomes very agitated and display signs of separation anxiety. The relationship with the conversational agent Samantha becomes very important and at some stage almost an addiction.

The software program Samantha is an example for the *universal solicitation* of current and future AI systems. In the movie, Samantha is available not just for work but companionship, counselling, sex and practically every other need Theodore may have. The lines between human relationships and the interaction with a computer are intentionally blurred. At some point, Theodore publicly claims to be dating his artificially intelligent operating system. What is made clear from the example of this movie, if emotional support from a software program is accepted, then no human friend or partner can compete with the endurance and consistency of support provided by a computer. This is an unquestioning and enduring support, that obviously is not mutual, and has an addictive potential.

In a similar vein, little robot that teaches a low functioning autistic child social skills can take an unlimited amount of physical abuse, if necessary. As a matter of fact, it is the robot's purpose to tolerate abuse while repetitively assisting the child to learn social skills such as making eye contact, shaking hands and smiling. By contrast, human relations, including professional relationships, are always governed by a set of rules and involve mutual 'give and take' or 'meeting of minds'. Yet, mutual, rule-bound behaviour does not apply to the relationship between the autistic child and the robot; nor does it apply to the relationship between an emotional chatbot that offers psychological advice to a vulnerable client. A negative aspect of such imbalanced relationships is that resistance to stress is lowered when ongoing emotional support is available 24 hours a day.

Even with current technology, it would be trivial to build a 24-hour chatbot that dispenses psychological information upon request. Alternatively, the chatbot can listen to any conversation the user may have and selectively comment as required. With a voice user interface, the user has hands-free access to psychological infor-

mation and advice around the clock and can shape her/his behaviour accordingly. With time, the the voice-over by the chatbot becomes a presence within everyday life of the human. The chatbot has a *soliciting presence*.

There is, of course, an important difference to the conversational agent Samantha in the movie *Her*. A 24-hour psychological AI chatbot provides targeted advice based on the most recent scientific data available. For instance, let's say voice or language analysis indicates that the user is stressed and possibly angry as a result of receiving an email. Immediately, the chatbot would intervene and point to the danger of "catastrophising" which is a frequent cognitive dysfunction in depression and anxiety. Catastrophising means to blow things out of proportion, a significant cognitive and emotional response to events that are relatively minor in nature. A user who has been made aware of the risk of catastrophising may adjust behaviour and even regulate emotions to avoid further cognitive dysfunctions and as a result distress.

Obviously, the risk could be a learned dependence on the advice of the chatbot. The system would have soliciting power. Unlike conventional psychotherapy, the bot is in the ear of the user all the time; a very intense form of psychological intervention. How would the user learn to live independently and to make individual decisions if computer advice is available 24x7?

In the age of multimedia where a user can have millions of online friends, why do people seek interactions with conversational agents? Loneliness may be the answer.

2.8 Loneliness

Lauren Kunze, CEO of Pandorabots, a Californian company supporting chatbot applications, gave a presentation in 2019 about chatting with machines: "Strange things 60 billion bot logs say about human nature". Two of her observations are important from a psychological point of view. The analysis of billions of chatbot conversations revealed that a staggering 20 to 30% include some form of sexual-oriented verbal abuse or inappropriate language. This number is so high that a whole chapter in this book is dedicated to the question of why humans abuse computers and the consequences for very advanced AI systems. The second observation is made towards the end of her presentation with reference to a discussion about chatbot applications for the social good. Lauren Kunze said that people talk with chatbots because they are lonely. With all the social media and networking that is available at this point in time, the analysis of a large number of chatbot conversations revealed that loneliness is the problem of our time. How can a conversational, psychological AI system help users who experience loneliness?

Psychologically, loneliness corresponds to a discrepancy between an individual's preferred and actual social relations (Peplau & Perlman, 1982; in Cacioppo et al., 2015). This discrepancy then leads to the negative experience of feeling alone and/or the distress of feeling socially isolated even when among family or friends

(Weiss, 1974). In other words, you can feel alone with hundreds of Facebook friends if you expect to have thousands of them. You can feel alone even if surrounded by family and friends if you expect something more or different.

Humans do not simply require the presence of others but the presence of significant others. People to trust and rely on. But most importantly, loneliness is a matter of perspective. It has been suggested that loneliness can be experienced at three different levels:

1. The intimate space (the family or people in a living environment),
2. The social space (including friends and acquaintances), and
3. The public space (typically, people are anonymous in this space).

Individuals can experience loneliness at different levels in these spaces. How is loneliness addressed from a psychological point of view?

Psychologists have tried a number of strategies to help individuals who feel lonely. Intuitively, these strategies make a lot of sense. Firstly, help individuals to get more friends. Yet, having a lot of friends is not a problem for most people in the age of social media. So clearly, simply assisting individuals to get a lot of contacts, acquaintances and friends is not really addressing the core of the issue.

Another approach is skills training to assist people to develop new relationships and friendships if and when required. Clearly, social skills training can be very helpful in the context of autism and a number of mental health issues. However, we have seen already that individuals can feel lonely in the presence of others, so the skill of obtaining new friends in itself is not addressing the problem.

From the viewpoint of CBT, maladaptive and unhelpful cognitions that result in the experience of loneliness must be addressed. What are the conditions of loneliness and how might this be manifested?

Nobody becomes or feels lonely voluntarily. The experience of loneliness typically includes the feeling of being on the fringe of a social network. A feeling of being isolated. Frequently, there is the experience of a process: Step-by-step the individual perceives isolation and loneliness as if moving from the centre of social relations to the edge. This experience often results in maladaptive thoughts and behaviours.

The perception of living on the edge of social networks frequently results in efforts to make new contacts and in having high expectations for these new friendships. These new relationships have great importance due to the perceived isolation. Hence, any event that makes the development of these friendships more difficult, in particular early in the process, can be viewed as negative. It can result in unhelpful thinking styles such as catastrophising and personalisation.

Perhaps as a result of not achieving human companionship, a common antidote is for people to seek a loving relationship with a pet. The adage 'A man's best friend is his dog' is a telling statement about the consistent loyalty and affection that many people receive through owning a pet. Pets are often used for therapeutic purposes to counteract the loneliness of living alone in old age when one's spouse may have passed on, and children are less accessible as they grow up and develop

their own lives. Similarly, owning or having regular access to a horse has been deemed valuable for equine assisted psychotherapy to help an individual build confidence, trust, impulse control. A feeling that someone or an animal cares is a powerful antidote to loneliness, which can be supported through the caring professions, and to current and future applications of AI.

2.9 The “Caring Professions” and the future of AI

Russell (2019, p.122) explores professions that may survive the advancement of artificial intelligence. He identifies psychotherapists, executive coaches, tutors and counsellors among those jobs that may survive the AI revolution. There is no fundamental barrier to AI systems providing psychological or coaching services but people may simply prefer that these jobs are provided by a human professional. By the time most jobs are gone due to the success of automation, people need to learn "the art of life itself" (Russell, 2019). Russell states that this is about our ability to appreciate and to create art, music, literature, conversation, gardening, food and many more areas.

Historically, Russell's observations are very interesting since one of the first AI programs that gained wide public attention is Eliza, a program that mimicked the verbal behaviour of a psychotherapist and that still lives in numerous versions on the Internet (Weizenbaum, 1966). Soon afterwards, Kenneth Colby developed PARRY, a computer model of paranoia (Colby, 1974). The system could engage in conversations and Colby thought of it as a kind of virtual reality teaching systems for students of mental health. These two examples from the early history of artificial intelligence demonstrate that people were always fascinated by the link between AI and psychotherapy, and the prospect that one day a computer program could provide therapy services.

Russell (2019) is correct in emphasising that people who seek psychological services want a therapist with deep human understanding and life experiences. But clients also appreciate knowledge and adherence to procedures; this is where artificial intelligence has a very clear advantage. It is easy to interface any AI system with vast public knowledge bases about nearly any fact of life, but most importantly, medical and psychological knowledge. These ontologies or knowledge bases have been developed over decades, and as a result, any system that interfaces with these knowledge systems can answer almost any factual question a human user may have. This is a very significant advantage over human professionals, because no matter how experienced, no service provider in psychology or the medical field can know everything. And of course, human providers are not available 24/7. A psychologically AI system that knows everything, in the sense of having access to vast declarative knowledge, is clearly superhuman. If these AI systems have value, then it is a question of balancing human and AI expertise and more importantly, creating a way how human service providers and AI system can work together to the benefit of clients.

References

American Psychiatric Association (APA). Diagnostic and Statistical Manual of Mental Disorders. 4th ed. (DSM-IV-TR). Washington, DC: APA; 2000.

Cacioppo S, Grippo, AJ, London, S, Goossens L, Cacioppo, JT, Loneliness: Clinical Import and Interventions. *Perspect Psychol Sci.*, 10(2) (2015) 238–249.

Colby KM, *Artificial Paranoia: a Computer Simulation of Paranoid Processes*. Pergamon, New York, 1974.

Kunze L, Chatting with Machines: Strange Things 60 Billion Bots Logs Say About Human Nature. <https://youtu.be/KSJ0DMQbN-s>

Lenat, DB, The Nature of Heuristics. *Artificial Intelligence*, 19 (1982) 189-249.

Nass C, Steuer, J, Tauber ER, Computers are social actors. CHI'94, *Human Factors in Computer Systems*, 72-78, 1994.

Russell S, *Human Compatible. Artificial Intelligence and the Problem of Control*. Viking. Penguin Random House, 2019.

Strachey J, Analysis of a Phobia in a Five-Year-Old Boy. *The Standard Edition of the Complete Psychological Works of Sigmund Freud, Volume X (1909): Two Case Histories ('Little Hans' and the 'Rat Man')*, 1-150.

Weiss R, *Loneliness. The Experience of Emotional and Social Isolation*. Boston: MIT Press 1974.

Weizenbaum J, ELIZA--A Computer Program For the Study of Natural Language Communication Between Man and Machine. *Communications of the ACM* Volume 9, Number 1 (January 1966): 36-35.

What is Problem-Solving Therapy? American Psychological Association | Division 12
<http://www.div12.org/sites/default/files/WhatIsProblemSolvingTherapy.pdf> Retrieved 22 March 2018.

Wolpe, J., Rachman, S. Psychoanalytic "evidence": A critique based on Freud's case of little Hans. *Journal of Nervous and Mental Disease*, 131 (1960) 135-148.