Academic Paper

A design framework to create Artificial Intelligence Coaches

Nicky Terblanche ☑ (University of Stellenbosch Business School, Cape Town, South Africa)

Abstract

There is on going debate about the potential impact of artificial intelligence (AI) on humanity. The application of AI in the helping professions is an active research area, but not in organisational coaching. Guidelines for designing organisational AI Coaches adhering to international coaching standards, practices and ethics are needed. This conceptual paper presents the Designing AI Coach (DAIC) framework that uses expert system principles to link human coaching efficacy (strong coach-coachee relationships, ethical conduct, focussed coaching outcomes underpinned by proven theoretical models) to established AI design approaches, creating a baseline for empirical research.

Keywords

artificial intelligence coaching, e-coaching, chatbot coach, chatbot design, executive coaching, organisational coaching,

Article history

Accepted for publication: 17 July 2020 Published online: 03 August 2020



© the Author(s) Published by Oxford Brookes University

Introduction

Coaching, as a helping profession, has made significant inroads in organisations as a mechanism to support people's learning, growth, wellness, self-awareness, career management and behavioural change (Passmore, 2015; Segers & Inceoglu, 2012). At the same time, the rise of AI is hailed by some as the most significant event in recent human history with the potential to disrupt virtually all aspects of human life (Acemoglu & Restrepo, 2018; Brynjolfsson & McAfee, 2012; Mongillo, Shteingart, & Loewenstein, 2014). However, claims of the abilities and potential of AI are often overstated and it seems unlikely that we will have AI that matches human intelligence in the near future (Panetta, 2018). This does not mean that AI is not already having a meaningful impact in many contexts, including helping professions, such as healthcare and psychology (Pereira & Diaz, 2019). It seems poised for further refinement, growth and possible disruption and it is therefore inevitable that all coaching stakeholders will need to pre-emptively consider how to leverage, create and adopt AI responsibly within the coaching industry (Lai, 2017). The use of AI in organisational coaching is under-researched and specifically, it is not known how to effectively design an organisational AI Coach.

For the purposes of this paper, 'coaching' is defined as 'a human development process that involves structured, focused interaction and the use of appropriate strategies, tools and techniques to promote desirable and sustainable change for the benefit of the client and potentially for other stakeholders' (Bachkirova, Cox, & Clutterbuck, 2014, p. 1). Furthermore, this paper limits its scope to organisational coaching that includes genres such as executive coaching, workplace coaching, managerial coaching, leadership coaching and business coaching (Grover & Furnham, 2016). The organisational coaching industry is growing rapidly and has become a global phenomenon used by numerous organisations worldwide to develop their employees (Theeboom, Beersma, & Vianen, 2013). As a growing industry and emerging profession, coaching evolves continuously and it seems inevitable that AI will play an increasingly significant role in organisational coaching in the future.

With sparse research available on the creation and application of AI in organisational coaching, this conceptual paper asks what needs to be considered, in principle, when designing an AI Coach. In answer, the Designing AI Coach (DAIC) framework is presented. This framework uses principles from expert systems to guide the design of AI Coaches based on widely agreed predictor of coaching success: strong coach-coachee relationships (De Haan et al., 2016; Graßmann, Schölmerich & Schermuly, 2019; McKenna & Davis, 2009), ethical conduct (Diochon & Nizet, 2015; Gray, 2011; Passmore, 2009) and focussed coaching outcomes all underpinned by proven theoretical models (Spence & Oades, 2011).

This paper proceeds as follows: It starts by contextualising AI Coaching within the current organisational coaching literature and shows that current definitions are inadequate. Next a brief overview of AI is provided where it is argued that chatbots, a type of AI and expert system, have potential for immediate application in organisational coaching. Since chatbot AI Coaching in organisations has not been well researched, this paper explores how chatbots have been designed and applied in related fields. The perceived benefits and challenges of coaching chatbots are described. Finally, the novel DAIC framework is presented before concluding with suggestions for further research.

Situating AI Coaching within the organisational coaching literature

Although the purpose of this paper is not to provide a systematic literature review of AI in organisational coaching, a literature search was conducted to understand how AI is currently positioned within organisational coaching. Using Google Scholar a search for 'artificial intelligence coach' and 'artificial intelligence coaching' did not produce any peer reviewed journal articles on AI and organisational coaching within the first 10 results pages. Neither did replacing 'coach/coaching' with 'organisational coach/coaching' yield any desired results. A number of papers on AI Coaching in healthcare did however appear, and these papers used the term 'e-coaching' to describe the use of AI in that context, with Kamphorst (2017) providing a particularly useful overview.

Using 'e-coaching' as a search term revealed a number of relevant results in relation to coaching. Clutterbuck (2010, p. 7) described e-coaching as a developmental relationship, mediated through e-mail and potentially including other media. E-coaching is described by Hernez-Broome and Boyce (2010, p. 285) as 'a technology-mediated relationship between coach and client'. Geissler et al. (2014, p. 166) defined it as 'coaching mediated through modern media', while Ribbers and Waringa (2015, p. 91) described e-coaching as 'a non-hierarchical developmental partnership between two parties separated by a geographical distance, in which the learning and reflection process was conducted through both analogue and virtual means'.

In healthcare and psychology, the search for 'e-coaching' revealed a broader definition with applications like the promotion of physical activity (Klein, Manzoor, Middelweerd, Mollee & Te Velde, 2015); regulating nutritional intake (Boh et al., 2016); treatment of depression (Van der Ven et al., 2012); and insomnia (Beun et al., 2016). In these domains, 'e-coaching' refers to the process

of not just facilitating the coaching, but includes autonomous entities doing the actual coaching (Kamphorst, 2017, p. 627). This extended definition contrasts with the organisational coaching literature where currently 'e-coaching' seems to refer only to varying communication modalities between a human coach and client. To demarcate what this paper argues to be a new area of practice and research in organisational coaching, a term to capture the use of autonomous coaching agents that could completely replace or at least augment human coaches is proposed: Artificial Intelligence (AI) Coaching. It is proposed that AI Coaching be defined independent of e-coaching to clearly distinguish it as a type of coaching *entity* and not merely another coaching modality. To fully grasp the concept of AI Coaching, a basic understanding of the nature of AI itself is required.

Artificial intelligence (AI)

'Artificial Intelligence' may be defined as 'the broad collection of technologies, such as computer vision, language processing, robotics, robotic process automation and virtual agents that are able to mimic cognitive human functions' (Bughin & Hazan, 2017, p. 4). Al is also described as a computer program combined with real-life data, which can be trained to perform a task and can become smarter about its users through experience with its users (Arney, 2017, p. 6). Another view states that AI is a science dedicated to the study of systems that, from the perspective of an observer, act intelligently (Bernardini, Sônego, & Pozzebon, 2018). Al research started in the early 1950s. It is an interdisciplinary field that applies learning and perception to specific tasks including potentially coaching.

A distinction can be made between artificial general intelligence (Strong AI) and artificial narrow intelligence (Weak AI). Strong AI is embodied by a machine that exhibits consciousness, sentience, the ability to learn beyond what was initially intended by its designers and can apply its intelligence in more than one specific area. Weak AI focusses on specific, narrow tasks, such as virtual assistants and self-driving cars (Siau & Yang, 2017). Expert systems are considered a form of Weak AI and is described as complex software programs based on specialised knowledge, able to provide acceptable solutions to individual problems in a narrow topic area (Chen, Hsu, Liu & Yang, 2012; Telang, Kalia, Vukovic, Pandita & Singh, 2018).

To match a human coach, a Strong AI entity would be needed since it promises to do everything a human can do, and more. This field of research is however in its infancy with some projections indicating that we may not see credible examples of Strong AI in the foreseeable future (Panetta, 2018). The implication of this is that it is highly unlikely for an AI entity to convincingly perform all the functions that a human coach currently performs any time soon.

While Strong AI may not yet be a possibility for coaching, Weak AI in the form of expert systems provides options worth exploring. For the purpose of this discussion, the focus therefore is on a particular embodiment of Weak AI that is currently showing potential for application in coaching, namely conversational agents or chatbots.

Conversational agents (chatbots)

A 'conversational agent or chatbot system' is defined as a computer programme that interacts with users via natural language either through text, voice, or both (Chung & Park, 2019). Chatbots typically receive questions in natural human language, associate these questions with a knowledge base, and then offer a response (Fryer & Carpenter, 2006). Various terms are used to describe chatbots, including conversational agents, talkbots, dialogue systems, chatterbots, machine conversation systems and virtual agents (Saarem, 2016; Saenz, Burgess, Gustitis, Mena, & Sasangohar, 2017). The origins of chatbot type systems can be traced back to the 1950s, when Allan Turing proposed a five-minute test (also known as the Imitation Game) based on a text

message conversation, where a human had to predict whether the entity they were communicating with via text was a computer program or not (Turing, 1950).

Two famous chatbots of the past are Eliza, developed in 1966 and PARRY, developed in the 1970s. Eliza imitated a Rogerian psychologist, using simple pattern matching techniques to restructure users' sentences into questions. Not withstanding the simplistic approach, its performance was considered remarkable, partly due to people's inexperience with this type of technology (Bradeško & Mladnic, 2012). PARRY imitated a paranoid person and when its transcripts were compared to real paranoia patients, psychiatrists were able to distinguish between the two sets only 48% of the time (Bradeško & Mladnic, 2012). More recently, chatbots have found new applications in the services industry where they are used to assist with customer queries, advice and fulfilment of orders (Araujo, 2018). Chatbots have proliferated with more than 100 000 chatbots created in one year on Facebook Messenger alone (Johnson, 2017).

As a form of Weak AI and expert system, chatbots are usually designed by a set of scripted rules (retrieval-based), AI (generative-based), or a combination of both to interact with humans (De Angeli & Brahnam, 2008). Driven by algorithms of varying complexity and optionally employing AI technologies such as machine learning, deep learning and Natural Language Processing, Generation and Understanding (NLP, NLG, NLU), chatbots respond to users by deciding on the appropriate response given a user input (Neff & Nagy, 2016; Saenz et al., 2017). From the expert system perspective, chatbots attempt to mimic human experts in a particular narrow field of expertise (Telang et al., 2018).

The perceived benefits and challenges of coaching chatbots in related fields

In organisational coaching, no empirical studies on the design and effectiveness of organisational coaching chatbots seem to be available. A broader assessment including fields related to organisational coaching such as health, well-being and therapy provide some evidence of the application of chatbots (Laranjo da Silva et al., 2018). Research has been conducted on the efficacy of chatbots that assist people with aspects such as eating habits, depression, neurological disorders and promotion of physical activity (Bickmore, Schulman, & Sidner, 2013; Bickmore, Silliman et al., 2013; Pereira & Diaz, 2019; Watson, Bickmore, Cange, Kulshreshtha & Kvedar, 2012).

Research from the healthcare domain claims that AI Coaching provides a wide range of strategies and techniques intended to help individuals achieve their goals for self-improvement (Kamphorst, 2017; Kaptein, Markopoulos, De Ruyter, & Aarts, 2015) and can potentially play an essential role in supporting behavioural change (Kocielnik, Avrahami, Marlow, Lu, & Hsieh, 2018). Other advantages of chatbot coaches include the possibility of interacting anonymously, especially in the context of sensitive information (Pereira & Diaz, 2019). People who interact with chatbots may therefore feel less shame and be more willing to disclose information, display more positive feelings towards chatbots and feel more anonymous, as opposed to interacting with real humans (Lucas, Gratch, King, & Morency, 2014). This is especially important in organisational settings where different stakeholders are involved in the coaching process and coachees are often caught between the firm's expectations and their own needs (Polsfuss & Ardichvili, 2008).

Although research on actual efficacy and benefits of coaching chatbots is rare, there seems to be agreement that a chatbot can do the following (Bii, 2013; Driscoll & Carliner, 2005; Klopfenstein, Delpriori, Malatini & Bogliolo, 2017):

- Keep a record of most, if not all, communications;
- Be trained with any text in any language;

- Facilitate a conversation, ask appropriate questions and allow the client to figure things out for themselves;
- Help clients to develop their self-coaching approach that is inexpensive and accessible;
- Be ethical, respect the client's choices and remain neutral and unbiased;
- Create and monitor a new dynamic environment for achieving coaching outcomes and make learning lasting and applicable to concrete goals;
- Collect trends and understand how clients talk about their challenges and desires;
- Support and complement coaching services.

However, there are also still numerous unsolved challenges regarding chatbots described in literature (Britz, 2016), including:

- Incorporating context: to produce sensible responses chatbots need to include both linguistic context and physical context.
- Coherent personality: when generating responses, a chatbot should ideally produce consistent answers to semantically identical inputs.
- Evaluation of models: the best way to evaluate a conversational agent is to measure whether it is fulfilling its task; this evaluation is easier when there is a specific goal.
- Intention and diversity: chatbots tend to produce generic responses as opposed to humans who produce responses that carry purpose and are specific to the input. Chatbots lack this kind of diversity due to current limiting AI capabilities.

The main challenge of creating realistic chatbots is seen to be the difficulty to maintain the ongoing context of a conversation (Bradeško & Mladenic, 2012). Current approaches use pattern-matching techniques to map input to output, but this approach rarely leads to purposeful, satisfying conversations. Understanding the benefits offered by chatbots enable us to better understand how they can realistically contribute to the organisational coaching domain.

A proposed design framework for organisational chatbot AI Coaches

Given the current use and perceived benefits of chatbot AI Coaches in related fields, the question about how to design these entities for organisational coaching arises. In an attempt to answer this question, an expert system design approach was followed. Expert system design dictates that the system should be modelled on how human experts execute a task (Lucas & Van Der Gaag, 1991). For AI Coaching this implies stipulating what constitutes effective human coaching, and mapping this to acknowledged chatbot design principles. This approach was followed to derive the DAIC framework. The two facets of the DAIC framework, effective human coaching and chatbot design principles are discussed next, where after the framework itself is presented.

Facet one - Effective human coaching

Following the expert system approach, a knowledge-base based on how human coaches operate effectively guided the development of the DAIC framework and consists of four principles: (i) widely agreed human-coach efficacy elements; (ii) the use of recognised theoretical models; and (iii) ethical conduct. The fourth principle (narrow coaching focus) stems from the inherent limitations of chatbots as Weak AI and expert systems namely the ability to only focus on one particular task. Each principle is elaborated on next.

To determine what constitutes the first design principle (widely agreed human-coach efficacy elements), the actively researched and growing body of knowledge on 'how' coaching works (De Haan, Bertie, Day, & Sills, 2010; Theeboom et al., 2014) was consulted. It appears that there are different opinions on the matter. Grant (2012, 2014) found that goal-orientation is the most

important determinant of coaching success while Bozer and Jones (2018) identified seven factors that contribute to successful workplace coaching: self-efficacy, coachee motivation, goal orientation, trust, interpersonal attraction, feedback intervention, and supervisory support. While there are varying opinions, a number of scholars agree that one aspect contributes more than any other to coaching success: the coach-coachee relationship (De Haan et al., 2016; Graßmann, Schölmerich & Schermuly, 2019; McKenna & Davis, 2009).

Aspects that help build a strong coach-coachee relationship include trust, empathy and transparency (Grant, 2014; Gyllensten & Palmer, 2007) with trust in particular being linked to higher levels of coachee commitment to the coaching process (Bluckert, 2005). Human coaches can build trust by being predictable, transparent (about their ability) and reliable (Boyce, Jeffrey Jackson & Neal, 2010). Perceived trustworthiness of another person is another important contributor to strong relationships (Kenexa, 2012). As a construct, trustworthiness consists of three elements. *Ability* is the trust instilled by the skills and competencies of a person (Mayer, Davis & Schoorman, 1995). *Benevolence* refers to the perception of being acted towards in a well-meaning manner (Schoorman, Mayer & Davis, 2007). *Integrity* is a measure of adherence to agreed-upon principles between two parties (Mayer et al., 1995). In summary, it appears that the following aspects of a coach are important contributors to strong coaching relationships and resultant successful coaching interventions: trust, empathy, transparency, predictability, reliability, ability, benevolence and integrity.

The second principle of the DAIC framework, is the need for evidence-based practice (Grant, 2014). One of the criticism often levelled at coaching is that practitioners use models and frameworks that are borrowed from other professions without having been empirically verified for the coaching context (Theeboom et al., 2013). Therefore, in addition to a strong coach-coachee relationship, an AI Coach must also be based on theoretically sound coaching approaches (Spence & Oades, 2011).

The third principle that underpins the DAIC framework is ethically sound practice. Ethics in coaching is an important and active research area (Diochon & Nizet, 2015; Gray, 2011; Passmore, 2009). The introduction of intelligent autonomous AI Coaching systems raises unique ethical concerns (Kamphorst, 2017). These concerns are underscored by users' increasing demand for assurance that the algorithms and AI used in their AI Coaches are structurally unbiased and ethical. Intended users of technologies like chatbots must be confident that the technology will meet their needs, will align with existing practices, and that the benefits will outweigh the detriments (Kamphorst, 2017).

Four pertinent types of ethical and legal issues were identified by Kamphorst (2017) and are applicable to AI Coaches in the context of organisational coaching: (i) privacy and data protection; (ii) autonomy; (iii) liability and division of responsibilities; and (iv) biases.

Firstly, both the need and the ability of AI Coaching systems to continuously collect, monitor, store, and process information raise issues regarding privacy and data protection. Questions such as 'who owns and can access the data obtained from a coaching session?' must be made clear to users.

Secondly, since AI Coaching combines the paradigm of empowering people and enhancing selfregulation while simultaneously entering their spheres, the personal autonomy of the users may be affected in relatively new ways, both positive and negative. It also raises the question of how to deal with potential manipulation by an AI Coaching system. Autonomous AI Coaching systems may offer users suggestions for action, thereby affecting their decision-making process (Luthans, Avey & Patera, 2008). Being influenced by its decision making seems to conflict with the classical understanding of self-directedness as professed in coaching. Thirdly, many stakeholders of various levels of diversity, specialisation, and complex interdependencies are involved in creating an AI Coaching system. Therefore, the division of liabilities and responsibilities among the relevant stakeholders involved (producers, providers, consumers, and organisations) cannot be ignored. Creating responsible AI Coaches also require alertness to the possibility that some clients need to work with a different specialist and not a coach (Braddick, 2017). The acceptance of AI Coaching applications will be constrained if the design and use of the system adhere to a different set of ethical principles than their intended users.

Lastly, machine learning typically used in AI relies on large amounts of data. Data originates from many sources and data is not necessarily neutral. Care must be taken to ensure that potential biases inherent in data are not transferred to the AI Coach via the learning process, or if not avoidable, these biases must be made explicit (Schmid Mast, Gatica-Perez, Frauendorfer, Nguyen & Choudhury, 2015).

Al Coaching in the organisational context presents additional ethical challenges. In traditional human-to-human coaching, contracting for coaching in organisations typically involves three parties: the coach, the coachee and the sponsoring organisation paying for the coaching. Although the organisation pays for the coaching, there is usually a confidential agreement between coach and coachee to the exclusion of the organisation (Passmore & Fillery-Travis, 2011). If an Al Coach is used and paid for by the organisation, the ethical question about who owns the details of the conversation must be made clear. It would potentially be unethical for the organisation to have access to the Al Coach-coachee conversation.

The final principle that underpins the DAIC framework relates to coaching focus. In traditional human-to-human coaching, several coaching facets could be pursued simultaneously, including for example goal-attainment, well-being, creation of self-awareness and behavioural change. Weak AI, however, at best acts in a narrow, specialised field (Siau & Yang, 2017). A prerequisite imposed by Weak AI and specially expert systems is that the focus of the system must be limited to a narrow area of expertise (Chen et al., 2012). This implies a specific coaching focus. The focus of chatbot AI Coaches should therefore initially be limited to one aspect typically associated with coaching.

Facet two - Chatbot design best practices

There are five chatbot design principles included in the DAIC framework: (i) level of human likeness; (ii) managing capability expectations; (iii) changing behaviour; (iv) reliability; and (v) disclosure.

The first principle raises the question of how human-like a chatbot AI Coach should be. Based on the desired human coach attributes described earlier, it seems logical that a chatbot AI Coach would need identity and presence as well as the ability to engage emotionally (Xu, Liu, Guo, Sinha, & Akkiraju, 2017). This is not an easy problem to solve as demonstrated by the 'uncanny valley' phenomenon, where objects that visually appear very human-like trigger negative impressions or feelings of eeriness (Ciechanowski, Przegalinska, Magnuski & Gloor, 2019; Sasaki, Ihaya & Yamada, 2017). Creating a chatbot that closely mimics a human is therefore counter intuitively not necessarily the best approach. Ciechanowski et al. (2019), for example, showed that people experience less of the 'uncanny effect' and less negative effect when interacting with a simple textbased chatbot as opposed to a more human-like avatar chatbot. That said, the uncanny valley is a continuum, implying that some human-like aspects may be beneficial. Araujo (2018), for example, showed that when a chatbot employs human-like cues, such as having a name, conversing in the first person and using informal human language including 'hello' and 'good-bye', users experienced a higher level of social presence than if these factors are absent. These cues automatically imply a sense of human self-awareness or self-concept by the chatbot, making it more anthropomorphic and relatable than a chatbot without a self-concept (Araujo, 2018).

For the second principle, it is important to set and manage expectations of the AI's capabilities by being clear on its limitations (Lovejoy, 2019). If clients interacting with a chatbot AI Coach expect the same level of intelligence as from a human coach, they are bound to be disappointed, which will in turn jeopardise the trust relationship. The chatbot AI coach must therefore clearly communicate its purpose and capabilities (Jain et al., 2018).

The third principle relates to the fact that the chatbot AI coaches could and likely will change their behaviour as they learn from continued usage. Users must therefore be made aware that their interactions are used to improve the AI Coach and that because of this, the interactions may change. It must also be clear that the AI Coach could ask for feedback from the user on how it performs (Lovejoy, 2019).

The fourth principle, reliability addresses the fact that because an AI Coach is continuously learning, it is bound to make mistakes. When it fails, the AI Coach needs to do so gracefully and remind the user that it is in an on going cycle of learning to improve (Lovejoy, 2019; Thies et al., 2017).

The fifth principle, disclosure states that even though the aim of an AI Coach is to eventually replace a human coach, at this stage of technological maturity, it is probably best to clearly communicate to the user that the AI Coach is in fact not a human and does not have human capabilities. This knowledge may assist users in managing their expectations and not, for example, disclose information that they would to a human coach (Lee & Choi, 2017).

The Designing AI Coach (DAIC) framework

Having discussed the two facets that inform the DAIC framework (human-coach effectiveness and chatbot design principles), this paper now presents the organisational DAIC Framework in Figure 1. The framework postulates that an effective chatbot AI Coach must focus on a specific coaching outcome, such as goal-attainment, well-being, self-awareness or behavioural change. Furthermore, the internal operating model of the chatbot must be based on validated theoretical models that support the specific coaching outcome. In addition, the important predictors of coaching success (a strong coach-coachee relationship) must be embedded in the chatbot interaction model. Finally, the chatbot's behaviour must be guided by an acceptable organisational coaching ethical code, all the while being cognisant of the requirements, restrictions and conventions of a typical organisational context.

To implement the DAIC framework, chatbot design best practices must be used. Table 1 provides a mapping between aspects of strong coach-coachee relationships and chatbot design considerations.

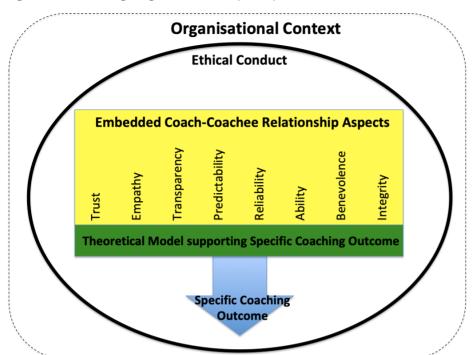




Table 1: Chatbot design practices to support strong coach-coachee relationships

| Coach attribute | Chatbot design consideration |
|-----------------|---|
| Trust | Avoid the 'uncanny valley' effect (Ciechanowski et al., 2019) |
| | Communicate data privacy agreement (Bakker, Kazantzis, Rickwood & Rickard, 2016) |
| | Create consistent chatbot personality (Shum et al. 2018) |
| | Reduce security and privacy concerns (Thies et al., 2017) |
| Empathy | Use a human name and human-like conversational cues (Araujo, 2018) |
| | Remember user's likes, dislikes and preferences across sessions (Thies et al., 2017) |
| Transparency | Reveal non-humanness (Lovejoy, 2019) |
| | Practice self-disclosure (Lee & Choi 2017) |
| | Showcase purpose and ethical standards (Neururer et al. 2018) |
| Predictability | State possible behaviour change due to continuous learning (Lovejoy, 2019) |
| | • Find a balance between a predictable personality and sufficient human-like variation (Sjödén et al. 2011) |
| | Use conversational context in interactions (Chaves & Gerosa 2018) |
| Reliability | • Fail gracefully (Lovejoy, 2019) |
| | Monitor chatbot performance and reliability (Lovejoy, 2019) |
| | Provide confirmation messages (Thies et al., 2017) |
| Ability | • Use established theoretical models (e.g. goal attainment) (Geissler et al., 2014; Poepsel, 2011) |
| | Use personalisation and avoid generic responses (Tallyn et al. 2018) |
| Benevolence | Communicate positive intent (Lovejoy, 2019) |
| | Demonstrate a positive attitude and mood (Thies et al., 2017) |
| Integrity | Clearly communicate limitations (Lovejoy, 2019) |
| | Clarify purpose in the introductory message (Jain et al. 2018) |

Future directions for research

The use of AI in the helping professions is a relatively new research area, and even more so in organisational coaching. Numerous opportunities exist for scientific investigation with a focus on the application of Weak AI in the form of chatbots. Two broad areas of research need immediate focus: (i) efficacy studies looking at how well a chatbot coach is able to fulfil certain coaching tasks; and (ii) technology adoption studies considering the factors that encourage or dissuade users from using chatbot coaches.

In terms of coaching efficacy studies, typical research focus areas in human coach studies could be performed with AI Coaches including how effective an AI Coach is to help a client with, for example, goal attainment, self-awareness, emotional intelligence, well-being and behavioural change. How do these results compare to clients using human coaches? As with human coach research, it is important to design robust studies that employ large-scale, random control trials, longitudinal research strategies. A positive aspect of using AI Coaches once they have been built, is that it is much cheaper to perform large-scale studies since there is no need to recruit and compensate human coaches. Logistically there are also fewer barriers, since chatbot AI Coaches are typically available on smartphone devices, tablets and personal computers.

In terms of technology adoption research, questions about which factors influence the adoption of a chatbot AI Coach by users need answering. What influences trust in AI Coaching? What role does the personality type of the client play in trusting and engaging with a chatbot coach? What level of humanness of a chatbot is optimal, for example, should a chatbot have a gender? When is it appropriate for the AI Coach to ask for user specified input (a much more difficult AI problem to solve) versus presenting users with predefined options? What should be included in the initial conversation to set realistic expectations and build trust, for example, what is the optimal balance between a chatbot trying to be as human as possible or admit that it has limitations? Which factors play a role in technology adoption? The well-known Technology Adoption Model (TAM) (Davis, Bagozzi, & Warshaw, 1989) and its numerous variants could be used to explore answers to these questions.

Perhaps researchers could use existing coaching competency frameworks, such as those of the International Coach Federation (ICF), as a guide to evaluate AI Coaches. One approach could be to ask credential adjudicators of various coaching bodies to officially evaluate the AI Coach. A cursory glance at the ICF competency model (ICF, 2017) suggests that currently, existing coaching chatbots could very well already pass some of the entry-level coach certification criteria.

Finally, it must be acknowledged that modelling AI Coaches on human coaches, the approach taken by this paper may not be the most optimal. It could be that AI Coaches need to have distinctly different skills and characteristics to human coaches. However, since no empirical evidence currently exists to prove or refute this assumption, this paper argues that in order to gathering empirical evidence, it is acceptable to start with the human-based expert system approach as a baseline. In time and with more empirical evidence, the true nature of AI Coaches will hopefully emerge.

Conclusion

Al is not currently sufficiently advanced to replace a human coach and given the trajectory of development in Strong AI, it is unlikely that we will see an AI Coach match a human coach any time soon. Human coaches will continue to outperform AI Coaches in terms of understanding the contexts surrounding the client, connecting with the client as a person, and providing socioemotional support. However, AI technology will inevitably improve as machine learning and the processing and understanding of natural language continues to improve exponentially, leading to AI Coaches that may excel at specific coaching tasks.

In order to guide and monitor the rise of AI Coaches in organisational coaching, the various stakeholders, such as practicing coaches, coaching bodies (such as ICF, COMENSA and EMCC), coach training providers and purchasers of coaching services (such as Human Resource professionals), are encouraged to educate themselves on the nature and potential of AI Coaching. They could actively participate in securing an AI Coaching future that ethically and effectively contributes to the coaching industry. It is hoped that the DAIC framework presented in this paper will provide some direction for this important emerging area of coaching practice and research.

References

- Acemoglu, D. and Restrepo, P. (2018) 'The race between man and machine: Implications of technology for growth, factor shares, and employment', *American Economic Review*, 108(6), pp.1488-1542. DOI: 10.3386/w22252.
- Araujo, T. (2018) 'Living up to the chatbot hype: The influence of anthropomorphic design cues and communicative agency framing on conversational agent and company perceptions', *Computers in Human Behavior*, 85(August), pp.183-189. DOI: 10.1016/j.chb.2018.03.051.
- Arney, K. (2017) 'Algorithm's gonna get you', *The Times Educational Supplement*. Available at: https://www.tes.com/magazine/article/algorithms-gonna-get-you.
- Bachkirova, T., Cox, E. and Clutterbuck, D. (2014) 'Introduction', in Cox, E., Bachkirova, T. and Clutterbuck, D. (eds.) *The complete handbook of coaching* (2nd edn.). London: Sage, pp.1-20.
- Bakker, D., Kazantzis, N., Rickwood, D. and Rickard, N. (2016) 'Mental Health Smartphone Apps: Review and Evidence-Based Recommendations for Future Developments', *JMIR Mental Health*, 3(1). DOI: 10.2196/mental.4984.
- Chatbots: An analysis of the state of art of literature (2018). Proceedings of the First Workshop on Advanced Virtual Environments and Education (WAVE2), 4–5 October 2018, Florianópolis, Brazil. DOI: 10.5753/wave.2018.1.
- Beun, R.J., Brinkman, W-P., Fitrianie, S. and et al, (2016) *Improving adherence in automated e-coaching: A case from insomnia therapy.* International Conference on Persuasive Technology, 5-7 April 2016, Salzburg, Austria, pp.276-287. DOI: 10.1007/978-3-319-31510-2 24.
- Bickmore, T.W., Schulman, D. and Sidner, C. (2013) 'Automated interventions for multiple health behaviors using conversational agents', *Patient Education & Counselling*, 92(2), pp.142-148. DOI: 10.1016/j.pec.2013.05.011 .
- Bickmore, T.W., Silliman, R.A., Nelson, K. and et al, (2013) 'A randomized controlled trial of an automated exercise coach for older adults', *Journal of the American Geriatrics Society*, 61(10), pp.1676-1683. DOI: 10.1111/jgs.12449.
- Bii, P.K. (2013) 'Chatbot technology: A possible means of unlocking student potential to learn how to learn', *Educational Research*, 4(2), pp.218-221. Available at: https://www.interesjournals.org/articles/chatbot-technology-a-possible-means-of-unlocking-student-potential-to-learn.pdf.
- Bluckert, P. (2005) 'Critical factors in executive coaching-the coaching relationship', *Industrial and Commercial Training*, 37(7), pp.336-340. DOI: 10.1108/00197850510626785.
- Boh, B., Lemmens, L., Jansen, A. and et al, (2016) 'An ecological momentary intervention for weight loss and healthy eating via smartphone and internet: Study protocol for a randomised controlled trial', *Trials*, 17(1). DOI: 10.1186/s13063-016-1280-x.
- Boyce, L.A., Jackson, R.J. and Neal, L.J. (2010) 'Building successful leadership coaching relationships: Examining impact of matching criteria in a leadership coaching program', *Journal of Management Development*, 29(10), pp.914-931.
- Braddick, C. (2017) Coaching at work: An artificial reality. Available at: https://www.coaching-at-work.com/2017/08/31/an-artificial-reality/.
- Bradeško, L. and Mladenić, D. (2012) A survey of chatbot systems through a Loebner Prize competition. Proceedings of the Slovenian Language Technologies Society, Eighth Conference of Language Technologies, 8-9 October 2012, Ljubljana, Slovenia, pp.34-37. Available at: http://nl.ijs.si/isjt12/proceedings/isjt2012_06.pdf.
- Britz, D. (2016) Deep learning for chatbots, Part 1: Introduction. Available at: http://www.wildml.com/2016/04/deep-learningfor-chatbots-part-1-introduction/.
- Brynjolfsson, E. and McAfee, A. (2012) Race against the machine: How the digital revolution is accelerating innovation, driving productivity, and irreversibly transforming employment and the economy. MIT Center for Digital Business.
- Bughin, J. and Hazan, E. (2017) 'The new spring of artificial intelligence: A few early economies', *VoxEU and CEPR*. Available at: https://voxeu.org/article/new-spring-artificial-intelligence-few-early-economics.
- Chaves, A.P. and Gerosa, M.A. (2018) Single or Multiple Conversational Agents?: An Interactional Coherence Comparison. Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, April 2018, Montreal, Canada.
- Chen, Y., Hsu, C., Liu, L. and Yang, S. (2012) 'Constructing a nutrition diagnosis expert system', *Expert Systems with Applications*, 39(2). DOI: 10.1016/j.eswa.2011.07.069.
- Chung, K. and Park, R.C. (2019) 'Chatbot-based healthcare service with a knowledge base for cloud computing', *Cluster Computing*, 22(1), pp.S1925-S1937. DOI: 10.1007/s10586-018-2334-5.
- Ciechanowski, L., Przegalinska, A., Magnuski, M. and Gloor, P.A. (2019) 'In the shades of the uncanny valley: An experimental study of human–chatbot interaction', *Future Generation Computer Systems*, 92(March), pp.539-548. DOI: 10.1016/j.future.2018.01.055.
- Clutterbuck, D. (2010) 'Coaching reflection: The liberated coach', *Coaching: An International Journal of Theory, Research and Practice*, 3(1), pp.73-81. DOI: 10.1080/17521880903102308.
- Davis, F.D., Bagozzi, R.P. and Warshaw, P.R. (1989) 'User acceptance of computer technology: a comparison of two theoretical models', *Management science*, 35(8), pp.982-1003. DOI: 10.1287/mnsc.35.8.982.

- De Angeli, A. and Brahnam, S. (2008) 'I hate you! Disinhibition with virtual partners', *Interacting with Computers*, 20(3), pp.302-310. DOI: 10.1016/j.intcom.2008.02.004.
- de Haan, E., Bertie, C., Day, A. and Sills, C. (2010) 'Clients' Critical Moments of Coaching: Toward a "Client Model" of Executive Coaching', Academy of Management Learning & Education, 9, pp.607-621. DOI: 10.5465/amle.2010.56659879.
- de Haan, E., Grant, A.M., Burger, Y. and Eriksson, P.O. (2016) 'A large-scale study of executive and workplace coaching: The relative contributions of relationship, personality match, and self-efficacy', *Consulting Psychology Journal: Practice and Research*, 68(3), pp.189-207. DOI: 10.1037/cpb0000058.
- Diochon, P.F. and Nizet, J. (2015) 'Ethical codes and executive coaches: One size does not fit all', *The Journal of Applied Behavioral Science*, 51(2), pp.277-301. DOI: 10.1177/0021886315576190.
- Driscoll, M. and Carliner, S. (2005) Advanced web-based training strategies: Unlocking instructionally sound online learning. San Francisco: John Wiley & Sons .
- Fryer, L. and Carpenter, R. (2006) 'Bots as language learning tools', Language Learning & Technology, 10(3), pp.8-14.
- Geissler, H., Hasenbein, M., Kanatouri, S. and Wegener, R. (2014) 'E-Coaching: Conceptual and empirical findings of a virtual coaching programme', *International Journal of Evidence Based Coaching and Mentoring*, 12(2), pp.165-186. Available at: https://radar.brookes.ac.uk/radar/items/585eb4f9-19ce-49e1-b600-509fde1e18c0/1/.
- Gessnitzer, S. and Kauffeld, S. (2015) 'The working alliance in coaching: Why behavior is the key to success', *The Journal of Applied Behavioral Science*, 51(2), pp.177-197. DOI: 10.1177/0021886315576407.
- Graßmann, C., Schölmerich, F. and Schermuly, C.C. (2020) 'The relationship between working alliance and client outcomes in coaching: A meta-analysis', *Human Relations*, 73, pp.35-58. DOI: 10.1177/0018726718819725.
- Grant, A.M. (2014) 'Autonomy support, relationship satisfaction and goal focus in the coach-coachee relationship: Which best predicts coaching success?', *Coaching: An International Journal of Theory, Research and Practice*, 7(1), pp.18-38. DOI: 10.1080/17521882.2013.850106.
- Gray, D.E. (2011) 'Journeys towards the professionalisation of coaching: Dilemmas, dialogues and decisions along the global pathway', *Coaching: An International Journal of Theory, Research and Practice*, 4(1), pp.4-19. DOI: 10.1080/17521882.2010.550896.
- Grover, S. and Furnham, A. (2016) 'Coaching as a developmental intervention in organisations: A systematic review of its effectiveness and the mechanisms underlying it', *PLoS ONE*, 11(7). DOI: 10.1371/journal.pone.0159137.
- Gyllensten, K. and Palmer, S. (2007) 'The coaching relationship: An interpretative phenomenological analysis', *International Coaching Psychology Review*, 2(2), pp.168-177.
- Hernez-Broome, G. and Boyce, L.A. (eds.) (2010) Advancing executive coaching: Setting the course for successful leadership coaching. San Francisco: John Wiley & Sons.
- International Coach Federation. (ICF) (2017) *ICF Core Competencies*. Available at: https://coachfederation.org/app/uploads/2017/12/CoreCompetencies.pdf.
- Jain, M., Kumar, P., Kota, R. and Patel, S.N. (2018) *Evaluating and Informing the Design of Chatbots.* DIS '18: Designing Interactive Systems Conference 2018, June 2018, Hong Kong.
- Johnson, K. (2017) Facebook Messenger hits 100,000 bots. Available at: https://venturebeat.com/2017/04/18/facebookmessenger-hits-100000-bots/.
- Kamphorst, B.A. (2017) 'E-coaching systems: What they are, and what they aren't', *Personal and Ubiquitous Computing*, 21(4), pp.625-632. DOI: 10.1007/s00779-017-1020-6.
- Kaptein, M., Markopoulos, P., De Ruyter, B. and Aarts, E. (2015) 'Personalizing persuasive technologies: Explicit and implicit personalization using persuasion profiles', *International Journal of Human-Computer Studies*, 77, pp.38-51. DOI: 10.1016/j.ijhcs.2015.01.004.
- Kenexa (2012) High Performance Institute Work Trends report. Available at: http://www.kenexa.com/ThoughtLeadership/WorkTrendsReports/TrustMatters.
- Klein, M.C.A., Manzoor, A., Middelweerd, A. and et al, (2015) 'Encouraging physical activity via a personalized mobile system', *IEEE Internet Computing*, 19(4), pp.20-27. DOI: 10.1109/MIC.2015.51.
- Klopfenstein, L.C., Delpriori, S., Malatini, S. and Bogliolo, A. (2017) The rise of bots: A survey of conversational interfaces, patterns, and paradigms. DIS '17: Designing Interactive Systems Conference 2017, June 2017, Edinburgh, pp.555-565. DOI: 10.1145/3064663.3064672.
- Kocielnik, R., Avrahami, D., Marlow, J. and et al, (2018) Designing for workplace reflection: A chat and voice-based conversational agent. DIS '18: Designing Interactive Systems Conference 2018, June 2018, Hong Kong, pp.881-894. DOI: 10.1145/3196709.3196784.
- Lai, P.C. (2017) 'The literature review of technology adoption models and theories for the novelty technology', *Journal of Information Systems and Technology Management*, 14(1), pp.21-38. DOI: 10.4301/s1807-17752017000100002.

- Laranjo da Silva, L., Dunn, A.G., Tong, H.L. and et al, (2018) 'Conversational agents in healthcare: A systematic review', Journal of the American Medical Informatics Association, 25(9), pp.1248-1258. DOI: 10.1093/jamia/ocy072.
- Lee, S. and Choi, J. (2017) 'Enhancing user experience with conversational agent for movie recommendation: Effects of self-disclosure and reciprocity', *International Journal of Human Computer Studies*, 103, pp.95-105.

Lovejoy, J. (2019) The UX of AI. Available at: https://design.google/library/ux-ai/.

- Lucas, G.M., Gratch, J., King, A. and Morency, L.P. (2014) 'It's only a computer: Virtual humans increase willingness to disclose', *Computers in Human Behavior*, 37, pp.94-100. DOI: 10.1016/j.chb.2014.04.043.
- Luthans, F., Avey, J.B. and Patera, J.L. (2008) 'Experimental analysis of a web-based training intervention to develop positive psychological capital', *Academy of Management Learning & Education*, 7(2), pp.209-221. DOI: 10.5465/amle.2008.32712618.
- Mayer, R.C., Davis, J.H. and Schoorman, F.D. (1995) 'An integrative model of organizational trust', Academy of Management Review, 20(3), pp.709-734. DOI: 10.2307/258792.
- McKenna, D.D. and Davis, S.L. (2009) 'Hidden in plain sight: The active ingredients of executive coaching', *Industrial and Organizational Psychology*, 2(3), pp.244-260. DOI: 10.1111/j.1754-9434.2009.01143.x.
- Mongillo, G., Shteingart, H. and Loewenstein, Y. (2014) 'Race against the machine', *Proceedings of the IEEE*, 102(4), pp.542-543. DOI: 10.1109/JPROC.2014.2308599.
- Neff, G. and Nagy, P. (2016) 'Talking to bots: Symbiotic agency and the case of Tay', *International Journal of Communication*, 10, pp.4915-4931.
- Neururer, M., Schlögl, S., Brinkschulte, L. and Groth, A. (2018) 'Perceptions on Authenticity in Chat Bots', *Multimodal Technologies and Interaction*, 2(3). DOI: 10.3390/mti2030060.
- Panetta, K. (2018) Widespread artificial intelligence, biohacking, new platforms and immersive experiences dominate this year's Gartner Hype Cycle. Available at: https://www.gartner.com/smarterwithgartner/5-trends-emerge-in-gartner-hype-cycle-for-emerging-technologies-2018/.
- Passmore, J. (2009) 'Coaching ethics: Making ethical decisions novices and experts', *The Coaching Psychologist*, 5(1), pp.6-10.

Passmore, J. (2015) Excellence in Coaching: The Industry Guide. London: Kogan Page.

- Passmore, J. and Fillery-Travis, A. (2011) 'A critical review of executive coaching research: A decade of progress and what's to come', An International Journal of Theory, Research and Practice, 4(2), pp.70-88. DOI: 10.1080/17521882.2011.596484.
- Pereira, J. and Diaz, O. (2019) 'Using Health Chatbots for Behavior Change: A Mapping Study', *Journal of Medical Systems*, 43(5). DOI: 10.1007/s10916-019-1237-1.
- Poepsel, M.A. (2011) The impact of an online evidence-based coaching program on goal striving, subjective well-being, and level of hope. Capella University. Available at: https://pqdtopen.proquest.com/doc/872553863.html.
- Polsfuss, C. and Ardichvili, A. (2008) 'Three principles psychology: Applications in leadership development and coaching', Advances in developing human resources, 10(5), pp.671-685. DOI: 10.1177/1523422308322205.
- Provoost, S., Lau, H.M., Ruwaard, J. and Riper, H. (2017) 'Embodied conversational agents in clinical psychology: A scoping review', *Journal of Medical Internet Research*, 19(5). DOI: 10.2196/jmir.6553.
- Ribbers, A. and Waringa, A. (2015) *E-coaching: Theory and practice for a new online approach to coaching.* New York: Routledge.
- Saarem, A.C. (2016) Why would I talk to you? Investigating user perceptions of conversational agents. Norwegian University of Science and Technology.
- Saenz, J., Burgess, W., Gustitis, E. and et al, (2017) The usability analysis of chatbot technologies for internal personnel communications. Industrial and Systems Engineering Conference 2017, 20-23 May 2017, Pittsburgh, Pennsylvania, USA, pp.1375-1380. Available at: http://toc.proceedings.com/36171webtoc.pdf.
- Sasaki, K., Ihaya, K. and Yamada, Y. (2017) 'Avoidance of novelty contributes to the uncanny valley', *Frontiers in Psychology*, 8. DOI: 10.3389/fpsyg.2017.01792.
- Schmid Mast, M., Gatica-Perez, D., Frauendorfer, D. and et al, (2015) 'Social Sensing for Psychology', *Current Directions in Psychological Science*, 24(2), pp.154-160. DOI: 10.1177/0963721414560811.
- Schoorman, F.D., Mayer, R.C. and Davis, J.H. (2007) 'An integrated model of organizational trust: Past, present, and future', *The Academy of Management Review*, 32(2), pp.334-354. DOI: 10.5465/amr.2007.24348410.
- Segers, J. and Inceoglu, I. (2012) 'Exploring supportive and developmental career management through business strategies and coaching', *Human Resource Management*, 51(1), pp.99-120. DOI: 10.1002/hrm.20432.
- Shum, H., He, X. and Li, D. (2018) 'From Eliza to Xiaolce: challenges and opportunities with social chatbots', *Frontiers of Information Technology & Electronic Engineering*, 19(1), pp.10-26. DOI: 10.1631/FITEE.1700826.

- Siau, K.L. and Yang, Y. (2017) *Impact of Artificial Intelligence, Robotics, and Machine Learning on Sales and Marketing.* Midwest United States Association for Information Systems 12th Annual Conference, 18-19 May 2017, Springfield, Illinois. Available at: http://aisel.aisnet.org/mwais2017/48.
- Sjödén, B., Silvervarg, A., Haake, M. and Gulz, A. (2011) 'Extending an Educational Math Game with a Pedagogical Conversational Agent: Facing Design Challenges'', in De Wannemacker, S., Clarebout, G. and De Causmaecker, P. (eds.) Interdisciplinary Approaches to Adaptive Learning: A Look at the Neighbours. Springer, pp.116-130.
- Spence, G.B. and Oades, L.G. (2011) 'Coaching with self-determination theory in mind: Using theory to advance evidencebased coaching practice', *International Journal of Evidence-Based Coaching and Mentoring*, 9(2), pp.37-55. Available at: https://radar.brookes.ac.uk/radar/items/59c36762-42b9-432e-b070-04c193b48f71/1/.
- Tallyn, E., Fried, H., Gianni, R. and et al, (2018) *The Ethnobot: Gathering Ethnographies in the Age of IoT*. CHI '18: CHI Conference on Human Factors in Computing Systems, April 2018, Montreal, Canada.
- Telang, P.R., Kalia, A.K., Vukovic, M. and et al, (2018) 'A Conceptual Framework for Engineering Chatbots', *IEEE Internet Computing*, 22(6), pp.54-59. DOI: 10.1109/MIC.2018.2877827.
- Theeboom, T., Beersma, B. and Van Vianen, A. (2013) 'Does coaching work? A meta-analysis on the effects of coaching on individual level outcomes in an organizational context', *The Journal of Positive Psychology*, 9(1). DOI: 10.1080/17439760.2013.837499.
- Thies, I.M., Menon, N., Magapu, S. and et al, (2017) *How do you want your chatbot? An exploratory Wizard-of-Oz study with young, urban Indians.* IFIP Conference on Human-Computer Interaction, 25–29 September 2017, Mumbai, India, pp.441-459.
- Turing, B.A.M. (1950) 'Computing machinery and intelligence', *Mind*, 49, pp.433-460. Available at: https://www.csee.umbc.edu/courses/471/papers/turing.pdf.
- Van der Ven, P., Henriques, M.R., Hoogendoorn, M. and et al, (2012) A mobile system for treatment of depression. Computing Paradigms for Mental Health. 2nd International Workshop on Computing Paradigms for Mental Health -MindCare 2012, 1-4 February 2012, Vilamoura, Portugal. Available at: https://www.scitepress.org/Papers/2012/38917/pdf/index.html.
- Watson, A., Bickmore, T.W., Cange, A. and et al, (2012) 'An internet-based virtual coach to promote physical activity adherence in overweight adults: randomized controlled trial', *Journal of Medical Internet Research*, 14(1). DOI: 10.2196/jmir.1629.
- Xu, A., Liu, Z., Guo, Y. and et al, (2017) A new chatbot for customer service on social media. 2017 CHI Conference on Human Factors in Computing Systems, 6-11 May 2017, Denver, Colorado, USA. DOI: 10.1145/3025453.3025496.

About the authors

Dr Nicky Terblanche is a senior lecturer/researcher in the MPhil in Management Coaching and MBA Information Systems at the University of Stellenbosch Business School.