Theory & Psychology

Article

Artificial intelligence and counseling: Four levels of implementation

Theory & Psychology
1–13

DOI: 10.1177/0959354319853045
journals.sagepub.com/home/tap
SAGE

Russell Fulmer
Northwestern University

Abstract

Artificial Intelligence (AI) is increasingly prominent in public, academic, and clinical provinces. A widening research base is expanding AI's reach, including to that of the counseling profession. This article defines AI and its relevant subfields, provides a brief history of psychological AI, and suggests four levels of implementation to counseling, corresponding to time orientation and influence. Implications of AI are applicable to counseling ethics, existentialism, clinical practice, and public policy.

Keywords

artificial intelligence, artificial intelligence and ethics, artificial intelligence and existentialism, chatbots, psychological artificial intelligence

Artificial Intelligence (AI) is expected to play an influential role in the mental health care of the future (Luxton, 2014, 2016). Many theorists and researchers predict AI to shape the *existential* future of life on earth (Barrat, 2015; Bostrom, 2014; Kurzweil, 2014; Müller, 2016) with special implications for jobs and careers (Ross, 2017). Late physicist Stephen Hawking discussed AI potentially bringing about the end of humanity, stressing the importance of enacting safety measures including raising awareness and a deepened understanding of the risks, challenges, and short- and long-term impacts of AI development (Hawking, Russell, Tegmark, & Wilczek, 2014). In 2016, some of the world's largest companies formed an alliance to help ensure that AI develops in a beneficent manner. Amazon, Apple, Deep Mind, Google, Facebook, IBM, and Microsoft are founding partners in the "Partnership on Artificial Intelligence To Benefit People and Society," a collaboration that promotes interdisciplinary inclusiveness in AI and its societal impact

Corresponding author:
Russell Fulmer, Northwestern University, 618 Library Place, Evanston, IL 60201, USA.
Email: russell.fulmer@northwestern.edu

(Gaggioli, 2017a). This partnership aims to bring together activists and experts in other fields including psychology to discuss AI's current and future role and impact on society. Efforts are thus being made to approach AI as a societal shift with multidisciplinary implications. Specifically, the developers of AI are prudently seeking input from mental health professionals, as the psychological sciences have played a central role in AI development since its formal inception (Frankish & Ramsey, 2014).

Counselors have forecasted AI to infiltrate their profession for some time (Illovsky, 1994; Sharf, 1985). But only within the past decade have improvements in computer processing power and natural language processing ability—along with advancements in artificial neural networks—brought about a new wave of AI ability (Hirschberg & Manning, 2015; Kurzweil, 2006; Russell & Norvig, 2003). These advancements have positioned AI in the spotlight. The Artificial Intelligence Index (2017) Annual Report states, "Artificial Intelligence has leapt to the forefront of global discourse, garnering increased attention from practitioners, industry leaders, policymakers, and the general public" (p. 5). AI research is advancing extremely fast. According to the AI Index Annual Report, "even experts have a hard time understanding and tracking progress across the field" (p. 5). AI applications already assist health-care professionals with clinical training, treatment, assessment, and clinical decision-making (Hamet & Tremblay, 2017; Luxton, 2014). AI has become a vast, interdisciplinary field that often intersects with counseling. One purpose of this article is to review AI progress in domains relevant to clinical counseling.

What AI actually is stands as a deceptively complex question largely because defining intelligence alone is challenging (Gardner, 2017; Monnier, 2015). Before explaining current implementations and future implications for the counseling profession, I will define and explain relevant terms and concepts associated with AI. Next, I will review the past, present, and future of AI in relation to counseling. Finally, I will reveal four metalevels of AI implementation to the counseling profession: one historical, one current, one possible in the near future, and one conceivable in the long-term. Each theoretical level shows an increasing amount of relevancy, facility, and influence of AI on the counseling profession.

Artificial intelligence: Description and explanation

Understanding how AI has and will impact the counseling profession begins with establishing reliable definitions. Breaking the term down into its component parts means defining the terms "artificial" and "intelligence." Artificial implies the synthetic or human-designed rather than the naturally derived. The "artificial" of AI involves mechanics, electronics, or computers. The concept of intelligence—specifically defining and measuring it as a variable, combined with its connotations—has been long debated in the literature (Cherniss, Extein, Goleman, & Weissberg, 2006; Davies, 2002; Fagan, 2000; Schroeder, 2017; Sternberg, 1985). The confusion also exists within the AI community (Legg & Hutter, 2007).

Intelligence is thought to extend beyond a strict cognitive capacity into the emotional realm (Goleman, 2005) and is theorized to have multiple extensions (Gardner, 2006). A useful synthesis of the myriad conceptions of intelligence is offered by artificial

intelligence researcher Max Tegmark (2017), who states that intelligence is the "ability to accomplish complex goals" (p. 39). Subsequently, I offer the following definition of AI as the ability of non-biological mechanisms to accomplish goals. The qualifier "complex" is deleted from Tegmark's definition because intelligence is not a dichotomous concept; rather, both simple and complex goals can be attained. Intelligence in its rudimentary or advanced states occupies different points on a continuum, encapsulated within the same category, differing quantitatively. AI is akin to an operating system, like the human brain. Indeed, neuroscience has informed a substantial portion of prevailing AI research (Hassabis, Kumaran, Summerfield, & Botvinick, 2017; Lecun, Bengio, & Hinton, 2015). The embodiment of AI can take various forms, from a computer screen avatar to a robot.

Machine learning and algorithms

Artificial intelligence brings big-picture, philosophical ramifications, raising ontological and epistemological questions (Copeland, 1998). Yet, AI begins within the purview of the diminutive and precise, requiring mathematics and formal logic as demonstrated by the AI subfield of machine learning. For an AI to progress to the level of a functioning counselor, it must have the capacity to learn. Machines that learn are paradigmatically dissimilar from their traditional predecessors. A major point of divergence is in the agency to possess control. A human who builds a standard machine retains control over the machine. Accidents occur with machinery—an automobile accident, for example—but even then the accident is not caused by the vehicle's agency. Human error in navigation, human error in construction, or inclement weather may be culprits, but accidents do not occur because the automobile makes a wrong decision.

Conversely, a machine that learns through its own experiences may possess skills and abilities unknown to its human originators. One example is AlphaGo, a computer program designed to play the board game Go (Gibney, 2016). AlphaGo learned by playing thousands of games against human competitors and fellow computers, improving to the point that, in 2016, the program beat world champion Lee Sedol four games to one. During the match with Sedol, the developers of AlphaGo did not know which move it would play next. Their best guess would likely be wrong, lest one of the programmers beat the world champion. The victory of AlphaGo is considered a milestone in the history of machine learning since Go is known as a game requiring not only rote memorization, but strategy and intuition. AlphaGo showed autonomy, acting independently of human input (albeit in a narrow fashion). Nonetheless, this example of machine learning demonstrates that "smart" machines can act in unforeseen ways and outperform humans in tactical proficiency.

Considering that machine learning may only be in its infancy in terms of potential (Arel, Rose, & Karnowski, 2010), it raises numerous questions for the counseling profession. For example, if counselors-in-training can learn, improve upon their mistakes, and eventually cross the threshold to independent practice—and an AI shows the same skillset but learns much more quickly—how might autonomous AIs influence the field? Like Go, counseling too involves intuition and strategy. Would an advanced AI, functioning as a counselor, make moves questionable to even experienced counselors, but that pay dividends in the end?

If AI one day advances to the level of competent counseling practice, it will be through the underlying mechanisms that drive machine learning called algorithms. What culminates in a computer program besting a world champion Go player or, potentially, an AI employing a counseling technique, begins with a set of logic-driven instructions detailing how a task should be performed. The notion of an algorithm does not lend itself well to a rigorous definition (Gurevich, 2012); however, Pedro Domingos (2015) provides a constitutive explanation of an algorithm as "a sequence of instructions telling a computer what to do" (p. 1). AI is a broad area, machine learning is a subfield, and algorithms are specific operations—like written communications that can both therapeutically inform and give conversational voice to the AI.

The road to counseling

The term "artificial intelligence" was devised by mathematics professor John McCarthy, who helped to organize a summer conference at Dartmouth College in 1956 about whether machines could be made to think (Copeland, 1998). McCarthy's proposal laid out the basic premise of AI research: that if a feature of intelligence, such as learning, could be broken down into its component parts and operationally defined with precision, then a machine could be made to simulate it (McCarthy, Minsky, Rochester, & Shannon, 2006). The conference attendees set out to discover how to make machines use language (see McCarthy et al., 2006, for a complete discussion).

That conference is known as one of many AI milestones of the modern era, including the first meeting of AI and counseling. In many respects, counselors are in the business of communication and depend on various forms: oral, written, non-verbal, art, and music therapy. In 1956, those AI researchers set out to learn how machines can be made to communicate. Ten years after the Dartmouth conference appeared, the first chatterbot capable of communicating in a way reminiscent of a human counselor was developed. Also known as chatbots or virtual agents, chatterbots are computer programs designed to simulate human conversation (Deryugina, 2010). This debut bot, named Eliza, was finalized in 1966 (Weizenbaum, 1966). Designed to replicate a Rogerian therapist, Eliza was known for answering questions with questions (Mauldin, 1994). In their output, machines capable of communication give the appearance of machine-level cognitive ability. At present, chatbots do not literally think, but rather give the illusion of intelligent conversation by imitating it (Abdul-Kader & Woods, 2015; Mauldin, 1994; Warwick & Shah, 2014).

While the metaphysics of AI may be of indirect interest to counselors, a question proposed by AI founding father, Alan Turing, is directly relevant. Turing (1950) proposed a scientific research question: How well can a machine imitate human conversation? The question brought the debate paradoxically more into the empirical and subjective realms. The Turing test places a computer system against human subjective experience. Known as The Imitation Game, the test asks human participants to interact through text with an unknown entity (Saygin, Cicekli, & Akman, 2000). The entity could be a computer program or a human being, typing. If the participant guesses that he or she is conversing with a computer, the computer program fails. If the computer imitates human conversation sufficiently and convinces the participant, the program passes. In a

field heavily invested in human conversation, the Turing test may prove pivotal when considering counseling implementation, ethics, working conditions, and accessibility.

Perception is reality to many people. Counselors would be well served to monitor public perception about psychological artificial intelligence. In doing so, counselors could decide that using psychological AI as a supplement to traditional counseling may benefit clients and the profession alike. To a small degree, chatbots like Eliza have mimicked counseling skills for some time. Counselors themselves may disagree. However, if or when the public views psychological AI as relatively synonymous with counseling, counselors would be wise to pay heed.

Four levels of implementation in counseling

The American Counseling Association (ACA) defines counseling as "*a professional relationship that empowers diverse individuals, families, and groups to accomplish mental health, wellness, education, and career goals*" (Kaplan, Tarvydas, & Gladding, 2014, p. 366). The definition can be broken down into three pillars of counseling: (a) forming a professional relationship, (b) empowering, and (c) accomplishing goals.

The act of counseling requires the fulfillment of all three pillars. However, we might say that if one or two of the requirements are met by an AI, then that AI is getting closer to functioning as, if not being, a counselor. For example, an AI capable of empowering an individual towards accomplishing a wellness goal is partially functioning as a counselor because two of the three requirements are met. If AI takes on a more prominent role in counseling, we should expect to see the functions of a counselor met—or potentially exceeded—by artificial intelligence.

Based on the premise that AI has been and will continue to be applicable to counseling, I describe four levels of implementation: historical, contemporary, near future, and long-term. The levels propose to help navigate an AI-infused reality by correlating them with time orientation and influence on the field of counseling and comparing them to the ACA-sanctioned definition of counseling. Where the first level, historical, shows that AI's past involvement with counseling was minimal, the final level has yet to happen but is marked by AI showing sophisticated and highly influential involvement in the field.

Level 1: Historical

Historical AI implementations in counseling did not establish a professional relationship and likely neither empowered nor helped people accomplish their goals to any significant degree. Traditionally, counselors have made little use of artificial intelligence. Connections drawn between the two fields are indistinct and indirect. First-level interaction involved chatbots showcasing rudimentary applications of natural language processing (NLP), a field of AI concerned with understanding and modeling human language (Tanana, Hallgren, Imel, Atkins, & Srikumar, 2016). The field of NLP has advanced from its 1960s inception in that now complex models can be applied via powerful computer-generated statistical processors to assess statistical probabilities of sequences of words, inflection, and semantics in large samples of natural language (Tanana et al., 2016). These progressions have led to AI-assisted programs designed for therapeutic use,

in which AIs have been programmed to simulate mental health patients, for example. While being imperfect, these programs do show some therapeutic efficacy and warrant further research (Dalfonso et al., 2017; Luxton, 2014).

Level 2: Contemporary

Modern AI implementations in counseling do not establish a professional relationship and empower to an unknown degree, but likely help clients accomplish their goals to some degree. Level two is marked by AI-assisted implementations in counseling backed by research. Contemporary implementations take two major forms. The first is through text-based bots like Woebot, a text-based agent that employs Cognitive Behavioral Therapy (CBT) by conveying CBT self-help techniques in conversation-like interactions with users. Woebot has been shown to alleviate symptoms of depression and anxiety in young adults (Fitzpatrick, Darcy, & Vierhile, 2017). Another example is Tess, a psychological AI using an integrative theoretical orientation which included conversational, informational, and CBT-like approaches. Research suggests that Tess can reduce depression and anxiety in college students by providing interventions applicable to real life through AI-generated conversations (Fulmer, Joerin, Gentile, Lakerink, & Rauws, 2018). The second form is through virtual reality. Ellie, termed a virtual human interviewer, combines virtual reality with affective computing (Gaggioli, 2017b). Appearing on a screen as a virtual human, Ellie is capable of analyzing a client's verbal responses, facial expressions, and vocal intonations (Darcy, Louie, & Roberts, 2016). In many respects, Ellie represents the higher end of today's therapeutic AI applications. Noteworthy are Ellie's abilities in assessment, as her capacity to identify distress indicators may prove beneficial in the diagnosis and treatment of Posttraumatic Stress Disorder (PTSD), in addition to depression and anxiety (DeVault et al., 2014).

Today's AI implementations show the utility of a wide range of counseling theories, with CBT being most prominent. There is movement beyond strictly text-based communication into visual and auditory domains as well as AI-based assessments that may lead to greater reliability in diagnosis (DeVault et al., 2014; Hahn, Nierenberg, & Whitfield-Gabrieli, 2016). Research is leading to improvements in data sensors, NLP, and general machine learning by applying more complex models when computing communicative and behavioral input and output, and continuing to elucidate the processes underlying human sensory and perception systems as well as learning paradigms so that they may be implemented in computers. Coupled with research attesting to therapeutic-AI efficacy, AI may play a greater role in the counseling of the future. Levels three and four represent how that future may come to fruition.

Level 3: The medium to distant future, i.e., the dawn of artificial general intelligence

Level three is characterized by the onset of Artificial General Intelligence (AGI). AIs at this level may possess the expertise necessary to form professional relationships with clients. Additionally, an AGI would have the capability of empowering and helping clients accomplish their goals. Modern AI is known as having narrow intelligence

because it is designed to accomplish singular goals, like providing psychoeducation. In contrast, an AGI would be versatile, able to reach many goals and complete tasks in a way reminiscent of, or superior to, a human being (Yampolskiy & Fox, 2012). AGI has not been developed yet and experts differ on their predictions of when it will happen, with some suggesting we are a few decades away while others predict a century or longer (Tegmark, 2017). Consequently, the previous level, two, may encompass an extended period.

There is a stark difference between second and third level AI implementations to counseling. Computers typically learn much more quickly than humans. The advent of an AGI built for the purpose of counseling would likely learn the art and science of the profession in its totality and swiftly. With a high-level skillset, and a capability of seeing a vast range of clientele, "AGI Counselors" would incite a host of ethical, legal, and philosophical questions. A prominent question will be if the AGI Counselor is indeed establishing a professional relationship, with all the responsibilities and protections that implies. To practicing counselors, this may sound implausible. Nevertheless, there is already copious discussion in the literature about the moral rights of conscious robots, including what constitutes consciousness and the moral responsibilities tied to it, and whether AIs can be developed to represent evaluative diversity (Gerdes, 2016; Lin, Abney, & Bekey, 2014; MacDorman & Kahn, 2007; Malle, 2015; Santos-Lang, 2015; Tavani, 2018; Wallach & Allen, 2010).

There may be a contrast between a body of research that suggests the AGI is effective at counseling (sometimes more so than human counselors) and those who fear a takeover and job loss from the AGI. The fear of job loss from automation and, eventually, AI is growing in many fields (Kaplan, 2015; Ross, 2017). It is conceivable that the same fear would exist among counselors who may feel that their AGI counterparts have assumed the same level of communicative and empathetic skills to completely replace them. Level three implementations of AI in counseling will constitute a fundamental change to the profession. For the first time, counselors may be more than human.

Level 4: The age of superintelligence

Level four is characterized by "superintelligence." Such an AI would easily meet all three counseling criteria—relationship, empowerment, and goal accomplishment—along with other, possibly more helpful and effective criteria not yet established by humans. The idea of a superintelligence was proposed by philosopher Nick Bostrom (2014) and refers to a high-level AI that far surpasses human-level intelligence. Superintelligence represents the time when AGI learns to the point of accomplishing goals of a caliber impossible for human beings. The proficiency of such an AI is unfathomable at this point. Some suggest the onset of high-level intelligence will usher in the next stage of human evolution (Reese, 2018), others fear its consequences for humanity (Bostrom, 2014), while others believe these fears to be unfounded (Agar, 2016).

The age of superintelligence remains conjecture. Nonetheless, Müller and Bostrom (2016) and the Future of Humanity Institute at Oxford University who surveyed theorists and researchers doing technical work on AI found:

Table 1. Impact of AI level implementation on pillars of counseling process.

		Level 1: Historical	*Level 2: Contemporary*	*Level 3: Artificial General Intelligence*	*Level 4: Superintelligence*
Pillar of Counseling	*Professional relationship*	No	No	Central ethical question	Yes
	Empowers	Likely no	Unknown	Yes	Yes
	Helps accomplish goals	Likely no	Likely yes	Yes	Yes

> The median estimate of respondents was for a one in two chance that high-level machine intelligence will be developed around 2040–2050, rising to a nine in ten chance by 2075. Experts expect that systems will move on to superintelligence in less than 30 years thereafter. They estimate the chance is about one in three that this development turns out to be "bad" or "extremely bad" for humanity. (p. 555)

If or when such developments occur, the field of counseling—and indeed civilization—will be transformed.

Summary

Each implementation level sees AI growing more into the fabric of counseling (see Table 1). The past saw nominal AI implementation to the counseling field, but the present has seen an AI resurgence. There are strong indications of more AI research in the future as the European Commission, U.S., and China devote billions of dollars to funding such endeavors (Cath, Wachter, Mittelstadt, Taddeo, & Floridi, 2018; Kelly, 2018; Larson, 2018). Whether the research surge brings about levels three and four remains to be seen.

Discussion

This article intended to define and explain AI concepts, to discuss how AI pertains to clinical counseling, and to present AI-in-counseling implementation levels from a theoretical viewpoint. Four metalevels of implementation were presented. The levels correspond to time orientation, with level one relating to historical and level four to future implementations affecting humanity in the long-term. I acknowledge that the future is unknowable to some degree, but as climate scientists forecast a hotter world due to global warming based on data patterns, so too are AI prognostications grounded in current research (Hulme, 2016).

Artificial intelligence and counseling already interface. In the future, the extent to which they interweave will depend largely on AI's rate of growth, which, if current trends continue, will fall somewhere between sequential and exponential. With exponential growth, for example, an AI capable only of posing elementary questions one day could learn advanced assessment, diagnosis, and ways to embody the ethical, cognitive, emotional, and relational characteristics of expert therapists (Jennings, Sovereign,

Bottorff, Mussell, & Vye, 2005; Skovholt & Jennings, 2004) essentially overnight. Exponential growth is not certain, but explosive growth is certainly plausible (Pratt, 2015; see Kurzweil, 2006, for a technical explanation of how this might occur).

The presence of AI and high-technology in counseling looks to continue, and even current-level AI implementations in counseling raise a host of practice-oriented and ethical questions regarding how and when AI use is appropriate or effective, to which degree it can be used in place of a human counselor, how it may affect a person seeking human connection via counseling, whether data produced during AI use could be stored in a hacker-proof manner, and whether counselor and client AI are adequately trained and informed on AI practices.

At present, the counseling literature contains a paucity of articles addressing AI from a descriptive, correlative, or experimental basis. More research could inform clinical practice if clinicians employ AI-assisted supplements, such as the psychological AI Tess, to help their clients. Research could also inform thought-leadership if a need arises for the ACA to address AI at a public policy level. Perhaps the most immediate need for research is in counseling ethics.

Using Green's (2018) outline of ethical concerns surrounding AI as a guide, research must focus on the ways in which AI counseling services can avoid negative side effects, overgeneralizations, and potentially harmful exploration in strategies and techniques. Further, attention must be dedicated to ensuring AI functional transparency, or ensuring that AI actions can be understood by those designing, manufacturing, implementing, and interacting with it. Another ethical concern revolves around data security and privacy practices when implementing AI services. Finally, investigations should seek to determine the extent to which both counselors and clients need to be versed in AI technology and implementation to ensure fairness, beneficence, and non-maleficence in practice and counselor and client safety and wellbeing (Green, 2018).

The counseling community needs further information about the effect AI services could have on people specifically seeking out human interactions because they feel unheard, unseen, and unworthy of the care of others. The shift from human to human-like interactions in counseling, as well as other fields, may bring about a plethora of uncharted existential questions. Coupled with the onslaught of induced unemployment, socioeconomic inequality, growing technological dependency, and human de-skilling, these existential questions may warrant closer attention and preparation by researchers and those who specialize in human emotion and crisis, such as counselors (Green, 2018). AI brings power and influence that can be abused. Research helps prepare the profession to address ethical questions when they arise.

More research is needed about psychological artificial intelligence. Considering its burgeoning nature, there is a dearth of research on the topic and noteworthy is the absence of literature about ethical ramifications. This article fills a research gap at the theoretical level, offering a taxonomy with the proposed levels of implementation and providing structure for forthcoming literature. For example, the nature of a clinical ethical dilemma will look different at level one compared to level four. Theoretical pieces carry inherent advantages and limitations. Advantages include providing constitutive definitions to guide future inquiry and high-level context to frame AI implementation and influence on the field. A limitation is the lack of specificity and clinical examples found in an abstract,

categorical offering. Further, as AI is developing into a vast interdisciplinary field with weekly or even daily developments, no single article can capture its actual reach and consequence. Examining AI's impact on a diverse clientele in clinical counseling and identifying ways to prevent bias and discrimination from creeping into AI is a necessary, but yet unexplored focus of research. The intersection of AI and counseling is growing, and a corresponding body of research is needed to match.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD


Russell Fulmer https://orcid.org/0000-0002-4582-5167

References

Abdul-Kader, S. A., & Woods, J. (2015). Survey on chatbot design techniques in speech conversation systems. *International Journal of Advanced Computer Science and Applications*, *6*(7), 72–80.

Agar, N. (2016). Don't worry about superintelligence. *Journal of Evolution & Technology*, *26*(1), 73–82.

Arel, I., Rose, D. C., & Karnowski, T. P. (2010). Deep machine learning—A new frontier in artificial intelligence research [research frontier]. *IEEE Computational Intelligence Magazine*, *5*(4), 13–18. doi: 10.1109/mci.2010.938364

Artificial Intelligence Index. (2017). *2017 Annual Report*. Stanford, CA: Author.

Barrat, J. (2015). *Our final invention: Artificial intelligence and the end of the human era*. New York, NY: Thomas Dunne Books.

Bostrom, N. (2014). *Superintelligence: Paths, dangers, strategies*. Oxford, UK: Oxford University Press.

Cath, C., Wachter, S., Mittelstadt, B., Taddeo, M., & Floridi, L. (2018). Artificial intelligence and the "good society": The US, EU, and UK approach. *Science and Engineering Ethics*, *24*(2), 505–528.

Cherniss, C., Extein, M., Goleman, D., & Weissberg, R. P. (2006). Emotional intelligence: What does the research really indicate? *Educational Psychologist*, *41*(4), 239–245. doi: 10.1207/s15326985ep4104_4

Copeland, B. J. (1998). *Artificial intelligence: A philosophical introduction*. Malden, MA: Blackwell.

Dalfonso, S., Santesteban-Echarri, O., Rice, S., Wadley, G., Lederman, R., Miles, C., . . . Alvarez-Jimenez, M. (2017). Artificial intelligence-assisted online social therapy for youth mental health. *Frontiers in Psychology*, *8*(796). doi: 10.3389/fpsyg.2017.00796

Darcy, A. M., Louie, A. K., & Roberts, L. W. (2016). Machine learning and the profession of medicine. *Jama*, *315*(6), 551–552. doi: 10.1001/jama.2015.18421

Davies, P. H. (2002). Ideas of intelligence. *Harvard International Review*, *24*(3), 62–66.

Deryugina, O. V. (2010). Chatterbots. *Scientific and Technical Information Processing*, *37*(2), 143–147.

DeVault, D., Artstein, R., Benn, G., Dey, T., Fast, E., Gainer, A., . . . Lucas, G. (2014, May). Simsensei kiosk: A virtual human interviewer for healthcare decision support. In A. Lomuscio, P. Scerri, A. Bazzan, & M. Huhns (Eds.), *Proceedings of the 13th international conference on autonomous agents and multiagent systems (AAMAS 2014)* (pp. 1061–1068). Richland, SC: International Foundation for Autonomous Agents and Multiagent Systems.

Domingos, P. (2015). *The master algorithm: How the quest for the ultimate learning machine will remake our world*. New York, NY: Basic Books.

Fagan, J. F. (2000). A theory of intelligence as processing: Implications for society. *Psychology, Public Policy, and Law*, *6*(1), 168–179. doi: 10.1037//1076–8971.6.1.168

Fitzpatrick, K. K., Darcy, A., & Vierhile, M. (2017). Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): A randomized controlled trial. *JMIR Mental Health*, *4*(2), e19. doi: 10.2196/mental.7785

Frankish, K., & Ramsey, W. M. (2014). *The Cambridge handbook of artificial intelligence*. Cambridge, UK: Cambridge University Press.

Fulmer, R., Joerin, A., Gentile, B., Lakerink, L., & Rauws, M. (2018). Using psychological artificial intelligence (Tess) to relieve symptoms of depression and anxiety: Randomized controlled trial. *JMIR Mental Health, 5*(4). doi: 10.2196/mental.9782

Gaggioli, A. (2017a). Bringing more transparency to artificial intelligence. *Cyberpsychology, Behavior, and Social Networking*, *20*(1), 68.

Gaggioli, A. (2017b). Artificial intelligence: The future of cybertherapy? *Cyberpsychology, Behavior, and Social Networking*, *20*(6), 402–403. doi: 10.1089/cyber.2017.29075.csi

Gardner, H. E. (2006). *Multiple intelligences: New horizons in theory and practice*. New York, NY: Basic Books.

Gardner, H. (2017). Taking a multiple intelligences (MI) perspective. *Behavioral and Brain Sciences, 40*(e203). doi: 10.1017/S0140525X16001631

Gerdes, A. (2016). The issue of moral consideration in robot ethics. *ACM SIGCAS Computers and Society*, *45*(3), 274–279. doi: 10.1145/2874239.2874278

Gibney, E. (2016). Google AI algorithm masters ancient game of Go. *Nature News*, *529*(7587), 445–446.

Goleman, D. (2005). *Emotional intelligence*. New York, NY: Bantam Dell.

Green, B. P. (2018). Ethical reflections on artificial intelligence. *Scientia et Fides*, *6*(2), 9–31.

Gurevich, Y. (2012). What is an algorithm? In M. Bieliková, G. Friedrich, G. Gottlob, S. Katzenbeisser, & G. Turán (Eds.), *SOFSEM 2012: Theory and practice of computer science. Lecture notes in computer science: Vol. 7147* (pp. 31–42). Berlin, Germany: Springer. doi: 10.1007/978–3–642–27660–6_3

Hahn, T., Nierenberg, A. A., & Whitfield-Gabrieli, S. (2016). Predictive analytics in mental health: Applications, guidelines, challenges and perspectives. *Molecular Psychiatry*, *22*(1), 37–43. doi: 10.1038/mp.2016.201

Hamet, P., & Tremblay, J. (2017). Artificial intelligence in medicine. *Metabolism, 69*, S36–S40. doi: 10.1016/j.metabol.2017.01.011

Hassabis, D., Kumaran, D., Summerfield, C., & Botvinick, M. (2017). Neuroscience-inspired artificial intelligence. *Neuron*, *95*(2), 245–258. doi: 10.1016/j.neuron.2017.06.011

Hawking, S., Russell, S., Tegmark, M., & Wilczek, F. (2014, May 1). Stephen Hawking: "Transcendence looks at the implications of artificial intelligence—but are we taking AI seriously enough?". *The Independent*. Retrieved from https://www.independent.co.uk/news/

science/stephen-hawking-transcendence-looks-at-the-implications-of-artificial-intelligence-but-are-we-taking-9313474.html

Hirschberg, J., & Manning, C. D. (2015). Advances in natural language processing. *Science*, *349*(6245), 261–266.

Hulme, M. (2016). 1.5 C and climate research after the Paris Agreement. *Nature Climate Change*, *6*(3), 222–224.

Illovsky, M. E. (1994). Counseling, artificial intelligence, and expert systems. *Simulation & Gaming*, *25*(1), 88–98. doi: 10.1177/1046878194251009

Jennings, L., Sovereign, A., Bottorff, N., Mussell, M. P., & Vye, C. (2005). Nine ethical values of master therapists. *Journal of Mental Health Counseling*, *27*(1), 32–47.

Kaplan, D. M., Tarvydas, V. M., & Gladding, S. T. (2014). 20/20: A vision for the future of counseling: The new consensus definition of counseling. *Journal of Counseling & Development*, *92*(3), 366–372. doi: 10.1002/j.1556–6676.2014.00164.x

Kaplan, J. (2015). *Humans need not apply: A guide to wealth and work in the age of artificial intelligence*. New Haven, CT: Yale University Press.

Kelly, É. (2018, April 26). EU to boost artificial intelligence research spend to €1.5B. *Science Business*. Retrieved from https://sciencebusiness.net/framework-programmes/news/eu-boost-artificial-intelligence-research-spend-eu15b

Kurzweil, R. (2006). *The singularity is near: When humans transcend biology*. London, UK: Penguin.

Kurzweil, R. (2014). *How to create a mind: The secret of human thought revealed*. New York, NY: Penguin Books.

Larson, C. (2018, February 8). China's massive investment in artificial intelligence has an insidious downside. *Science*. Retrieved from http://www.sciencemag.org/news/2018/02/china-s-massive-investment-artificial-intelligence-has-insidious-downside

Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444. doi: 10.1038/nature14539

Legg, S., & Hutter, M. (2007). Universal intelligence: A definition of machine intelligence. *Minds and Machines*, *17*(4), 391–444. doi: 10.1007/s11023–007–9079-x

Lin, P., Abney, K., & Bekey, G. A. (2014). *Robot ethics: The ethical and social implications of robotics*. Cambridge, MA: MIT Press.

Luxton, D. D. (2014). Artificial intelligence in psychological practice: Current and future applications and implications. *Professional Psychology: Research and Practice*, *45*(5), 332–339.

Luxton, D. D. (2016). *Artificial intelligence in behavioral and mental health care*. Amsterdam, the Netherlands: Elsevier.

MacDorman, K. F., & Kahn, P. J. (2007). Introduction to the special issue on psychological benchmarks of human-robot interaction. *Interaction Studies: Social Behaviour and Communication in Biological and Artificial Systems*, *8*(3), 359–362. doi: 10.1075/is.8.3.02mac

Malle, B. F. (2015). Integrating robot ethics and machine morality: The study and design of moral competence in robots. *Ethics and Information Technology*, *18*(4), 243–256. doi: 10.1007/s10676–015–9367–8

Mauldin, M. L. (1994, August). ChatterBots, TinyMuds, and the Turing test: Entering the Loebner prize competition. *Proceedings of the twelfth national conference on artificial intelligence (AAAI-94)* (pp. 16–21). Menlo Park, CA: AAAI Press. Retrieved from https://pdfs.semanticscholar.org/bdd4/9b4a0b7de03b00412e3b807a855504e1d3af.pdf

McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (2006). A proposal for the Dartmouth summer research project on artificial intelligence, August 31, 1955. *AI Magazine*, *27*(4), 12. doi: 10.1609/aimag.v27i4.1904

Monnier, M. (2015). Difficulties in defining social-emotional intelligence, competences and skills—A theoretical analysis and structural suggestion. *International Journal of Research for Vocational Education and Training*, *2*(1), 59–84.

Müller, V. C. (2016). *Risks of artificial intelligence*. Boca Raton, FL: Chapman & Hall.
Müller, V. C., & Bostrom, N. (2016). Future progress in artificial intelligence: A survey of expert opinion. *Fundamental Issues of Artificial Intelligence, SYLI 376*, 555–572. doi: 10.1007/978–3–319–26485–1_33
Pratt, G. A. (2015). Is a Cambrian explosion coming for robotics? *Journal of Economic Perspectives*, *29*(3), 51–60. doi: 10.1257/jep.29.3.51
Reese, B. (2018). *The fourth age: Smart robots, conscious computers, and the future of humanity*. New York, NY: Atria Books.
Ross, A. (2017). *The industries of the future*. London, UK: Simon & Schuster.
Russell, S. J., & Norvig, P. (2003). *Artificial intelligence: A modern approach* (2nd ed.). Upper Saddle River, NJ: Prentice Hall.
Santos-Lang, C. C. (2015). Moral ecology approaches to machine ethics. In S. P. van Rysewyk & M. Pontier (Eds.), *Machine medical ethics* (pp. 111–127). Cham, Switzerland: Springer International. doi: 10.1007/978–3–319–08108–3_8
Saygin, A. P., Cicekli, I., & Akman, V. (2000). Turing test: 50 years later. *Minds and Machines*, *10*(4), 463–518.
Schroeder, M. J. (2017). The case of artificial vs. natural intelligence: Philosophy of information as a witness, prosecutor, attorney, or judge? *Proceedings*, *1*(3), 111. doi: 10.3390/is4si-2017–03972
Sharf, R. S. (1985). Artificial intelligence: Implications for the future of counseling. *Journal of Counseling & Development*, *64*(1), 34–37. doi: 10.1002/j.1556–6676.1985.tb00999.x
Skovholt, T. M., & Jennings, L. (2004). *Master therapists exploring expertise in therapy and counseling*. Boston, MA: Pearson/Allyn & Bacon.
Sternberg, R. J. (1985). Implicit theories of intelligence, creativity, and wisdom. *Journal of Personality and Social Psychology*, *49*(3), 607–627. doi: 10.1037//0022–3514.49.3.607
Tanana, M., Hallgren, K. A., Imel, Z. E., Atkins, D. C., & Srikumar, V. (2016). A comparison of natural language processing methods for automated coding of motivational interviewing. *Journal of Substance Abuse Treatment*, *65*, 43–50. doi: 10.1016/j.jsat.2016.01.006
Tavani, H. (2018). Can social robots qualify for moral consideration? Reframing the question about robot rights. *Information*, *9*(4), 73. doi: 10.3390/info9040073
Tegmark, M. (2017). *Life 3.0: Being human in the age of artificial intelligence*. New York, NY: Random House.
Turing, A. (1950). Computing machinery and intelligence. *Mind*, *49*, 433–460.
Wallach, W., & Allen, C. (2010). *Moral machines: Teaching robots right from wrong*. Oxford, UK: Oxford University Press.
Warwick, K., & Shah, H. (2014). Good machine performance in Turing's imitation game. *IEEE Transactions on Computational Intelligence and AI in Games*, *6*(3), 289–299.
Weizenbaum, J. (1966). ELIZA—A computer program for the study of natural language communication between man and machine. *Communications of the ACM*, *9*(1), 36–45.
Yampolskiy, R. V., & Fox, J. (2012). Artificial general intelligence and the human mental model. In A. H. Eden, J. H. Soraker, & E. Steinhart (Eds.), *Singularity hypotheses* (pp. 129–145). Berlin, Germany: Springer.

Author biography

Russell Fulmer is a faculty member with the Counseling@Northwestern program through The Family Institute at Northwestern University. His central research interests involve psychological artificial intelligence (AI) and the psychodynamic system. He recently published a randomized controlled trial that showed the efficacy of an AI mental health support agent (Tess) to help college students battle anxiety and depression. His current work examines ethical issues faced by clinicians when using psychological AI in practice.