Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Current Opinion in Psychology

AI in mental health

Simon D'Alfonso

With the advent of digital approaches to mental health, modern artificial intelligence (AI), and machine learning in particular, is being used in the development of prediction, detection and treatment solutions for mental health care. In terms of treatment, AI is being incorporated into digital interventions, particularly web and smartphone apps, to enhance user experience and optimise personalised mental health care. In terms of prediction and detection, modern streams of abundant data mean that data-driven AI methods can be employed to develop prediction/detection models for mental health conditions. In particular, an individual's 'digital exhaust', the data gathered from their numerous personal digital device and social media interactions, can be mined for behavioural or mental health insights. Language, long considered a window into the human mind, can now be quantitatively harnessed as data with powerful computer-based natural language processing to also provide a method of inferring mental health. Furthermore, natural language processing can also be used to develop conversational agents used for therapeutic intervention.

Address
The University of Melbourne School of Computing and Information Systems, Australia

Corresponding author: D'Alfonso, Simon (dalfonso@unimelb.edu.au)

Current Opinion in Psychology 2020, **36**:112–117

This review comes from a themed issue on **Cyberpsychology**

Edited by **Jon Elhai** and **Dmitri Rozgonjuk**

https://doi.org/10.1016/j.copsyc.2020.04.005

Introduction

The field of mental health care, like a range of other fields, has been impacted by the revolution in digital technology and artificial intelligence (AI), with the field of digital mental health now firmly established as well as continued and emerging work on AI-driven solutions for mental health [1]. AI is a broad term, encompassing a range of techniques and approaches to developing computational systems that perform cognitive processes characteristic of humans, such as learning, the ability to reason and problem solve, pattern recognition, generalisation, and predictive inference. This review will focus on three of the main ways that AI is being applied in mental health:

- Personal sensing or digital phenotyping
- Natural language processing of clinical texts and social media content
- Chatbots

Smartphones and digital phenotyping

Digital phenotyping [2] or personal sensing [3] involves using sensor and usage data from personal digital devices, particularly smartphones, to infer contextual and behavioural information about an individual that can then be used as input for machine learning methods to predict psychological/psychometric outcomes and mental health conditions. Apart from smartphones there has also been work on using wearables such as smartwatches and actigraphy devices [4,5], and there has even been some incipient discussion on the use of the Internet of Things for mental health [6]. For a more general discussion of digital phenotyping beyond this AI-relative overview, the reader can consult the article on digital phenotyping in this special issue [7].

Being two of the most prevalent mental health conditions, much of the digital phenotyping research has been on depression and anxiety. One connection that has been researched is that between movement or physical activity (as tracked by geolocation and accelerometer sensors) and mental ill-health, particularly depression, as measured by scales such as the PHQ-9. [8•,9•,10].

In terms of smartphones and screen input, recent work suggests that keystroke dynamics of clicking, tapping, scrolling and swiping can provide mental health clues, without the need to linguistically analyse inputted character content. A recent publication on touchscreen typing pattern analysis for depression detection [11••] proposes a machine learning-based method for determining depressive individuals as measured by the PHQ-9, based on smartphone typing patterns as input.

Smartphone and other digital device sensing has been used to study conditions other than depression and anxiety, including early work on schizophrenia or psychotic disorders [12,13•]. In [13•] the study authors found that reductions in the number and duration of outgoing calls, as well as number of text messages were associated with relapses of schizophrenia.

Smartphones have also proved to be a very effective way to administer ecological momentary assessments for mental health monitoring (EMA), providing an easy and efficient way to send individuals repeated questionnaires to obtain *in-situ* assessments of contexts, behaviours, psychological states and moods [14–16]. Moving beyond just assessments but retaining ecological momentariness, the novel idea of Ecological Momentary Interventions (EMI) is to provide momentary psychological interventions or behavioural prompts delivered via personal mobile devices during an individual's daily life, often informed by their responses to EMAs [17,18,19•].

The true potential for smartphone EMIs however, and something which is yet to be fully realised, is to incorporate AI systems that can deliver contextually relevant and personalized therapy recommendations informed by smartphone sensing information and digital phenotyping insights [20]. Such a system would also learn about the individual and evolve over time, continuously improving its responses based on user patterns and system interactions.

Language, voice and mental health

The idea that the language we use, and our vocalisations, can indicate our psychological states, coupled with advances in the AI fields of natural language processing (NLP) and audio analysis, has led to an emergence of research on associations between language/voice characteristics and mental health. Transcriptions of clinical interviews/sessions are traditional sources of textual content for mental health language analysis. However, Internet technologies such as social media, online forums and instant messaging offer rich new sources of non-clinical text for such analysis [21].

Properties characteristic of language disturbance, such as impoverished vocabulary, semantic incoherence and reduced syntactic complexity are indicators of severe mental illness, particularly schizophrenia/psychosis. Such properties can be quantified using NLP techniques and the resulting figures used as inputs to machine-learning models for mental health classification/prediction. An early research instance used semantic coherence and syntactic complexity to predict later psychosis development with 100% accuracy (n = 34) [22]. Recent research by the same team [23••] applied the same natural language processing approach to a larger dataset, developing a similar machine-learning classifier that had an 83% accuracy in predicting psychosis onset and a 72% accuracy in discriminating the speech of recent-onset psychosis patients from that of healthy individuals. Whilst in a sense such research simply adds to the long-observed association between language disturbance and severe mental illness, the development of such automated tools offers a precise, scalable and rapidly executable means for detection/prediction.

The newsfeeds and forums of services such as Facebook, Twitter and Reddit provide a rich source of material for input into natural language processing systems. By analysing linguistic features in social media content, it is possible to generate machine learning models that can be used to infer an individual's mental health earlier than traditional approaches. Recent work continues research that started to appear after the advent of social media on analysing posted content for signs of mental health issues, with most of this research focusing on depression [24–28,29••]. In [29••], the Facebook posts of a relatively large patient cohort were analysed to predict, with an accuracy approximately matching screening surveys, depression as recorded in their electronic medical records. The researchers found that language predictors of depression include emotional (sadness), interpersonal (loneliness, hostility), and cognitive (preoccupation with the self, rumination) processes. There has also been recent notable work on the detection of psychosis [30] and suicidal ideation/risk [31,32]. Facebook themselves have implemented a machine learning system to identify post content that indicates people who might be at risk of self-harm [33]. Beyond the natural language processing of text, work on the audio analysis of paralinguistic or acoustic (e.g. volume, pitch and intonation) aspects of speech using AI has shown that such properties of speech can also be computationally analysed to infer mental health information [34].

Chatbots and virtual agents

A chatbot is a computer program that mimics conversation with users via a chat interface, either text or voice based. The underlying system can be based on a variety of foundations, ranging from a set of simple rule-based responses and keyword matching to sophisticated NLP techniques [35,36]. The history of chatbots is intimately tied with psychology. Apart from the interesting philosophical and psychological questions they raise, the first well-established chatbot, ELIZA was actually programmed (in 1966) to simulate a Rogerian psychotherapist [37]. Recent reviews indicate that a few dozen chatbots have been developed, for a range of conditions, including depression, autism and anxiety. User satisfaction with chatbots is good and preliminary evidence for efficacy is reasonably favourable [38,39].

The simplest of these chatbots can be used as conversational search assistants or recommendation system interfaces, leading users to relevant mental health information or therapy content after a basic and brief dialogical interaction [40]. Whilst the arrival of an AI agent capable of replicating a human therapist is not on the near horizon, if at all, more advanced AI agents incorporating sophisticated NLP are able to simulate a modest conversation employing therapeutic techniques. Whilst not intended to replace the human therapist, such therapeutic chatbots can provide their own form of interaction with users.

They can be available at any time to communicate, can be used by individuals who experience stigma or discomfort with seeing a therapist, and can be accessed by those with limited access to traditional mental health services.

Three of the most prominent therapeutic mental health chatbots that have emerged over the last few years are Woebot [41], Wysa [42] and Tess [43]. Woebot delivers cognitive behavioural therapy in the form of brief, daily conversations and mood tracking to help clients with depression and anxiety. A randomized controlled trial study to determine the feasibility, acceptability, and preliminary efficacy of Woebot found that after two weeks of use, the Woebot group experienced a significant reduction in depression, as measured by the PHQ-9, compared to the information-only control group, who were provided with an NIMH e-book on depression and did not experience an overall reduction. However, both groups significantly reduced anxiety as measured by the Generalized Anxiety Disorder scale (GAD-7) [44]. Wysa employs several methods such as cognitive behavioural therapy, behavioural reinforcement and mindfulness to help clients with depression. A preliminary study of Wysa [45•] showed that the group of users with high engagement had a significantly higher average improvement of the PHQ-9 measure compared with the group of users with low engagement. In user-provided feedback responses, a modest 68% found the app experience helpful and encouraging. Tess appears to have the most published research out of all these chatbot options [46]. Similar in nature to the Woebot study, a recent study to assess the feasibility and efficacy of Tess [47••] showed improvements in depression and anxiety in a cohort of college students. The Tess user group had statistically significant differences over the control group (which was also provided with the same NIMH e-book on depression) for measures of PHQ-Q, GAD-7 as well as the Positive and Negative Affect Scale (PANAS).

Chatbots for mental health show promise, however further work is required to obtain stronger findings and validate them in larger samples and across longer durations. Apart from the technical sophistication of language processing techniques, work on mental health chatbots will also need to consider the aspect of affective and empathic AI [48,49]. Beyond textual chatbots, virtual therapy agents with an avatar representation such as ELLIE, which can also process nonverbal signals, extend the ambit and abilities of AI therapy agents [50].

Finally, there are ethical dimensions to consider in deploying AI agents for mental health [51•,52•]. It is also important to realise the limitations of chatbots, and that they can serve as complements or supplements, rather than replacements for professionally trained human therapists. Responses to emergencies such as disclosures of immediate harm or suicidal ideation are limited and sometimes dangerously inappropriate, as one test of Woebot and Wysa demonstrated [53]. Such scenarios need to be handled outside of the chatbot. Linguistic detection of problems beyond the bot's purview should at least be followed by the immediate presentation of contact information for live help, a further possibility being to notify a relevant other.

Supplementary topics

Before ending, it is worth quickly mentioning two current topics involving AI.

Ethics and AI mental health research

Mental health care is already a field that by its nature raises particular ethical and legal considerations as well as the need for regulation. The development of AI and its increasing application to an array of sectors, including mental health, has brought with it the need to ethically scrutinise and regulate this application. Thus, we are now at a new intersection point, where the combination of AI and mental health raises its own novel considerations [54•].

If not kept in check, AI could exacerbate traditional ethical problems in mental health care. Furthermore, AI brings in its own ethical issues such as fairness, inclusiveness, transparency, accountability, privacy, security, reliability and safety [55]. A recent editorial article [56], which touches upon some of these themes, raises the issue of patient and public involvement in AI mental health research. It makes the interesting case that patients, service users and carers should participate as 'domain experts' in the design, research and development of AI mental health solutions.

The digital therapeutic alliance

The therapeutic alliance, the relationship that develops between a therapist and a patient, is a significant factor in the outcome of psychological therapy. As mental healthcare starts to increasingly adopt digital technologies and AI, offering therapeutic interventions that may not involve human therapists, the notion of the therapeutic alliance in digital mental healthcare requires exploration.

The term digital therapeutic alliance (DTA) is a broad one that can apply to a range of types of digital mental health care. In its simplest sense it could apply to the alliance between client and therapist in the case of therapy sessions conducted via e-mail, online chat or videoconferencing. The more interesting cases however are those involving a human client and a computerised therapeutic intervention, whether that be a smartphone/ web app or sophisticated conversational agent. The human interaction aspect of conversational agents was touched upon earlier. However, given their central presence in current digital mental health, it is the notion of a DTA in terms of app interventions that has generated attention in recent years. The two main considerations are how can such a DTA be measured and how can it be fostered in apps? Some attempts have been made to devise or test measures, but thus far even the most developed of these attempts have largely just taken an existing measure such as the Working Alliance Inventory or Agnew Relationship Measure and made some adjustments to suit the digital app context [57•,58•]. Whatever measures of the DTA do emerge from research into constructing purpose-built measures for digital mental health interventions, it is most likely the case that the incorporation of artificial intelligence will be crucial to fostering it.

Credit author statement

Simon D'Alfonso is the sole author of this article.

Conflict of interest statement

Nothing declared.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Graham S, Depp C, Lee EE, Nebeker C, Tu X, Kim H-C, Jeste DV: **Artificial intelligence for mental health and mental illnesses: an overview**. *Curr Psychiatry Rep* 2019, **21**.

2. Onnela J-P, Rauch SL: **Harnessing smartphone-based digital phenotyping to enhance behavioral and mental health**. *Neuropsychopharmacology* 2016, **41**:1691-1696.

3. Mohr DC, Zhang M, Schueller SM: **Personal sensing: understanding mental health using ubiquitous sensors and machine learning**. *Annu Rev Clin Psychol* 2017, **13**:23-47.

4. Difrancesco S, Lamers F, Riese H, Merikangas KR, Beekman ATF, van Hemert AM, Schoevers RA, Penninx BWJH: **Sleep, circadian rhythm, and physical activity patterns in depressive and anxiety disorders: A 2-week ambulatory assessment study**. *Depress Anxiety* 2019, **36**:975-986.

5. Scott J, Grierson A, Gehue L, Kallestad H, MacMillan I, Hickie I: **Can consumer grade activity devices replace research grade actiwatches in youth mental health settings?** *Sleep Biol. Rhythms* 2019, **17**:223-232.

6. de la Torre Díez I, Alonso SG, Hamrioui S, Cruz EM, Nozaleda LM, Franco MA: **IoT-based services and applications for mental health in the literature**. *J Med Syst* 2019, **43**.

7. Montag C, Sindermann C, Baumeister H: **Digital phenotyping in psychological and medical sciences: a reflection about necessary prerequisites to reduce harm and increase benefits**. *Curr Opin Psychol* 2020.

8. • Masud MT, Mamun MA, Thapa K, Lee D, Griffiths MD, Yang S: **Unobtrusive monitoring of behavior and movement patterns to detect clinical depression severity level via smartphone**. *Journal of Biomedical Informatics Biomed Inform* 2020, **103**.
Investigates associations between features extracted from physical activity and movement patterns (accelerometer and GPS sensors) and PHQ-9 psychometric tests. Develops a linear regression model to estimate PHQ-9 score and a support vector machine algorithm to classify the depression severity level among individuals.

9. • Ware S, Yue C, Morillo R, Lu J, Shang C, Bi J, Kamath J, Russell A, Bamis A, Wang B: **Predicting depressive symptoms using smartphone data**. *Smart Health* 2020, **15**.
Investigates associations and machine learning models between various features extracted from geolocation data and PHQ-9/QIDS psychometric tests.

10. Saeb S, Lattie EG, Schueller SM, Kording KP, Mohr DC: **The relationship between mobile phone location sensor data and depressive symptom severity**. *PeerJ* 2016, **4**.

11. •• Mastoras R-E, Iakovakis D, Hadjidimitriou S, Charisis V, Kassie S, Alsaadi T, Khandoker A, Hadjileontiadis LJ: **Touchscreen typing pattern analysis for remote detection of the depressive tendency**. *"Scientific ReportsSci Rep* 2019, **9**.
Proposes a machine learning-based method for determining depressive individuals as measured by the PHQ-9, based on smartphone typing patterns as input.

12. Barnett I, Torous J, Staples P, Sandoval L, Keshavan M, Onnela J-P: **Relapse prediction in schizophrenia through digital phenotyping: a pilot study**. *Neuropsychopharmacology* 2018, **43**:1660-1666.

13. • Buck B, Scherer E, Brian R, Wang R, Wang W, Campbell A, Choudhury T, Hauser M, Kane JM, Ben-Zeev D: **Relationships between smartphone social behavior and relapse in schizophrenia: A preliminary report**. *Schizophrenia Research Res* 2019, **208**:167-172.
Uses exploratory mixed effects models to examine relationships of social behaviour to schizophrenia relapse, finding that reductions in the number and duration of outgoing calls, as well as number of text messages were associated with relapses.

14. Bell IH, Lim MH, Rossell SL, Thomas N: **Ecological momentary assessment and intervention in the treatment of psychotic disorders: a systematic review**. *Psychiatr Serv* 2017, **68**:1172-1181.

15. Robinaugh DJ, Brown ML, Losiewicz OM, Jones PJ, Marques L, Baker AW: **Towards a precision psychiatry approach to anxiety disorders with ecological momentary assessment: the example of panic disorder**. *General Psychiatry* 2020, **33**.

16. Triantafillou S, Saeb S, Lattie EG, Mohr DC, Kording KP: **Relationship between sleep quality and mood: ecological momentary assessment study**. *JMIR Ment Health* 2019, **6**.

17. Myin-Germeys I, Klippel A, Steinhart H, Reininghaus U: **Ecological momentary interventions in psychiatry**. *Curr Opin Psychiatry* 2016, **29**:258-263.

18. Schueller S, Aguilera A, Mohr D: **Ecological momentary interventions for depression and anxiety**. *Depress Anxiety* 2017, **34**.

19. • Hanssen E, Balvert S, Oorschot M, Borkelmans K, van Os J, Delespaul P, Fett A: **An ecological momentary intervention incorporating personalised feedback to improve symptoms and social functioning in schizophrenia spectrum disorders**. *Psychiatry Res* 2020, **earch284**.
Examined the feasibility and effectiveness of an ecological momentary intervention for schizophrenia, finding that personalised feedback informed by momentary ecological assessments may impact momentary psychosis in daily life.

20. D'Alfonso S, Carpenter N, Alvarez-Jimenez M: **Making the MOST out of smartphone opportunities for mental health**. In *Proceedings of the 30th Australian Conference on Computer-Human Interaction*; *Melbourne: 2018*.

21. Calvo R, Milne D, Hussain S, Christensen H: **Natural language processing in mental health applications using non-clinical texts**. *Nat Lang Eng* 2017:1-37.

22. Bedi G, Carrillo F, Cecchi GA, Slezak DF, Sigman M, Mota NB, Ribeiro S, Javitt DC, Copelli M, Corcoran CM: **Automated analysis of free speech predicts psychosis onset in high-risk youths**. *NPJ Schizophr* 2015, **1**.

23. •• Corcoran CM, Carrillo F, Slezak DF, Bedi G, Klim C, Javitt DC, Bearden CE, Cecchi GA: **Prediction of psychosis across protocols and risk cohorts using automated language analysis**. *World Psychiatry* 2018, **17**:67-75.
Develops natural language processing machine learning models that predict psychosis and discriminate the speech of psychosis patients based on measures of semantic coherence and syntactic complexity.

24. Briand A, Almeida H, Meurs M: **Analysis of social media posts for early detection of mental health conditions**. *Advances in Artificial Intelligence. Canadian AI 2018. Lecture Notes in Computer Science*; *Toronto: 2018*.

25. Biradar A, Totad S: **Detecting depression in social media posts using machine learning**. *Recent Trends in Image Processing and Pattern Recognition. RTIP2R 2018. Communications in Computer and Information Science*; *Solapur: 2019*.

26. Ziwei BY, Chua HN: **An application for classifying depression in tweets**. In *Proceedings of the 2nd International Conference on Computing and Big Data*; *Taichung: 2019*.

27. Syarif I, Ningtias N, Badriyah T: **Study on mental disorder detection via social media mining**. *4th International Conference on Computing, Communications and Security (ICCCS)*; *Rome: 2019*.

28. Tadesse MM, Lin H, Xu B, Yang L: **Detection of depression-related posts in reddit social media forum**. *IEEE Access* 2019, **7**:44883-44893.

29. Eichstaedt JC, Smith RJ, Merchant RM, Ungar LH, Crutchley P,
•• Preoţiuc-Pietro D, Asch DA, Schwartz HA: **Facebook language predicts depression in medical records**. *Proceedings of the National Academy of Sciences115no. 44 Natl Acad Sci U S A* 2018, **115**:11203-11208.
Study that analyses the Facebook posts of a relatively large patient cohort to predict, with an accuracy approximately matching screening surveys, depression as recorded in their electronic medical records.

30. Birnbaum ML, Ernala SK, Rizvi AF, Arenare E, Van Meter AR, De Choudhury M, Kane JM: **Detecting relapse in youth with psychotic disorders utilizing patient-generated and patient-contributed digital data from Facebook**. *NPJ Schizophr* 2019, **5**.

31. Huang Y, Liu X, Zhu T: **Suicidal ideation detection via social media analytics**. *Human Centered Computing. HCC 2019. Lecture Notes in Computer Science*; *Čačak: 2019*.

32. Coppersmith G, Leary R, Crutchley P, Fine A: **Natural language processing of social media as screening for suicide risk**. *Biomed Inform Insights* 2018, **10**.

33. Card C: *How Facebook AI Helps Suicide Prevention*. 2018. 10 September [Online]. Available: https://about.fb.com/news/2018/09/inside-feed-suicide-prevention-and-ai/. [Accessed 29 March 2020].

34. Low D, Bentley K, Ghosh S: **Automated assessment of psychiatric disorders using speech: a systematic review**. *Laryngosc Investig Otolaryngol* 2020, **5**:96-116.

35. Laranjo L, Dunn AG, Tong HL, Kocaballi AB, Chen J, Bashir R, Surian D, Gallego B, Magrabi F, Lau AYS, Coiera E: **Conversational agents in healthcare: a systematic review**. *J Am Med Inform Assoc* 2018, **25**:1248-1258.

36. Ramesh K, Ravishankaran S, Joshi A, Chandrasekaran K: **A Survey of Design Techniques for Conversational Agents," in Information, Communication and Computing Technology, ICICCT 2017**. *Communications in Computer and Information Science*. Singapore: Springer; 2017.

37. Weizenbaum J: **ELIZA a computer program for the study of natural language communication between man and machine**. *Commun ACM* 1966, **9**:36-45.

38. Abd-alrazaq AA, Alajlani M, Alalwan AA, Bewick BM, Gardner P, Househ M: **An overview of the features of chatbots in mental health: a scoping review**. *Int J Med Inform* 2019, **132**.

39. Vaidyam AN, Wisniewski H, Halamka JD, Kashavan MS, Torous JB: **Chatbots and conversational agents in mental health: a review of the psychiatric landscape**. *Can J Psychiatry* 2019, **64**:456-464.

40. D'Alfonso S, Santesteban-Echarri O, Rice S, Wadley G, Lederman R, Miles C, Gleeson J, Alvarez-Jimenez M: **Artificial intelligence-assisted online social therapy for youth mental health**. *Front Psychol* 2017, **8**.

41. Woebot Labs, Woebot [Online]. Available: https://woebot.io/. [Accessed 29 March 2020].

42. Wysa Ltd, "Wysa," [Online]. Available: https://www.wysa.io/ . [Accessed 29 March 2020].

43. X2 Foundation, "X2," [Online]. Available: https://www.x2ai.com/ . [Accessed 29 March 2020].

44. Fitzpatrick K, Darcy A, Vierhile Molly: **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 Ment Health* 2017, **4**.

45. Inkster B, Sarda S, Subramanian Vinod: **An Empathy-Driven,
• Conversational Artificial Intelligence Agent (Wysa) for Digital Mental Well-Being: Real-World Data Evaluation Mixed-Methods StudyJMIR mHealth and uHealth6no. 11empathy-driven, conversational artificial intelligence agent (Wysa) for digital mental well-being: real-world data evaluation mixed-methods study**. *JMIR Mhealth Uhealth* 2018, **6**.
Study evaluating the effectiveness and engagement levels of the Wysa chatbot, showing that the group of users with high engagement had a significantly higher average improvement of the PHQ-9 measure compared with the group of users with low engagement.

46. X2 Foundation, "Outcomes," [Online]. Available: https://www.x2ai.com/outcomes. [Accessed 29 March 2020].

47. Fulmer R, Joerin A, Gentile B, Lakerink L, Rauws M: **Using
•• Psychological Artificial Intelligence (Tess) to Relieve Symptoms of Depression and Anxiety: Randomized Controlled TrialJMIR Mental Health5no. 4relieve symptoms of depression and anxiety: randomized controlled trial**. *JMIR Ment Health* 2018, **5**.
Study to assess the feasibility and efficacy of the Tess chatbot, showing improvements in depression and anxiety, as measured by the PHQ-Q, GAD-7 and PANAS scales, in a cohort of college students.

48. Morris R, Kouddous K, Kshirsagar R, SM S: **Towards an artificially empathic conversational agent for mental health applications: system design and user perceptions**. *J Med Internet Res* 2018, **20**.

49. George S: **From sex and therapy bots to virtual assistants and tutors: how emotional should artificially intelligent agents be?** *Proceedings of the 1st International Conference on Conversational User Interfaces*; *Dublin: 2019*.

50. Rizzo A, Shilling R, Forbell E, Scherer S, Gratch J, Morency LP: **Autonomous virtual human agents for healthcare information support and clinical interviewing**. *Artificial Intelligence in Behavioral and Mental Health Care*. Academic Press; 2016:53-79.

51. Fiske A, Henningsen P, Buyx A: **Your Robot Therapist Will See
• You Now: Ethical Implications of Embodied Artificial Intelligence in Psychiatry, Psychology, and PsychotherapyJournal of Medical Internet Research21no. 5robot therapist will see you now: ethical implications of embodied artificial intelligence in psychiatry, psychology, and psychotherapy**. *J Med Internet Res* 2019, **21**.
This paper assesses the ethical and social implications of using bots and AI agents in mental health care. It develops a set of preliminary recommendations on how to address ethical and social challenges in current and future applications of such systems.

52. Kretzschmar K, Tyroll H, Pavarini G, Manzini A, Singh I, N.Y.P.A.
• Group: **Can Your Phone Be Your Therapist? Young People's Ethical Perspectives on the Use of Fully Automated Conversational Agents (Chatbots) in Mental Health SupportBiomedical Informaticsyour phone be your therapist? Young people's ethical perspectives on the use of fully automated conversational agents (Chatbots) in mental health support**. *Insights* 2019, **11**:1-9.
This article looks at the strengths and limitations of using chatbots in mental health support. It outlines minimum ethical standards for these platforms, including issues surrounding privacy and confidentiality, efficacy and safety, and reviews the Woebot, Joy and Wysa bots/platforms according to the proposed framework.

53. BBC, "Child advice chatbots fail to spot sexual abuse," BBC, [Online]. Available: https://www.bbc.com/news/technology-46507900. [Accessed 30 March 2020].

54. Gooding P: **Mapping the rise of digital mentalE health
• technologies: emerging issues for law and society**. *Mapping*

the Therapeutic Alliance in Digital Mental Health technologies: Emerging issues for law and society 2019, **vol. 67**.
This article maps established and emerging technologies in the mental health context with an eye to locating major socio-legal and ethical issues. One of the first/main articles that has a comprehensive focus on these issues within the field of digital mental health.

55. Microsoft, "Responsible AI principles from Microsoft," 25 February 2020. [Online]. Available: https://www.microsoft.com/en-us/ai/responsible-ai.

56. Carr S: **'AI gone mental': engagement and ethics in data-driven technology for mental health**. *J Ment Health* 2020, **11**.

57. • Berry K, Salter A, Morris R, James S, Bucci S: **Assessing Therapeutic Alliance in the Context of mHealth Interventions for Mental Health Problems: Development of the Mobile Agnew Relationship Measure (mARM) QuestionnaireJournal Medical Internet Research20no. 4therapeutic alliance in the context of mHealth interventions for mental health problems: development of the mobile agnew relationship measure (mARM) questionnaire**. *J Med Internet Res* 2018, **20**.
An early attempt to devise a measure of the digital therapeutic alliance with smartphone apps based on the Agnew Relationship Measure of the therapeutic alliance.

58. • Henson P, Peck P, Torous J: **Considering the rise of digital mental health Interventions**. *Harvard Review of Psychiatry27no. 4 Rev Psychiatry* 2019, **27**:268-273.
Discusses the nascent topic of the digital therapeutic alliance and offers a quick measure for it based on the Working Alliance Inventory of the therapeutic alliance.