Conversational AI Therapist for Daily Function Screening in Home Environments

Jingping Nie
Columbia University
New York, New York, USA
jn2551@columbia.edu

Hanya Shao
The Soho Center for Mental Health Counseling, Kensington Wellness
New York, New York, USA
verashao@thesohocenter.com

Minghui Zhao
Columbia University
New York, New York, USA
mz2866@columbia.edu

Stephen Xia
Columbia University
New York, New York, USA
stephen.xia@columbia.edu

Matthias Preindl
Columbia University
New York, New York, USA
matthias.preindl@columbia.edu

Xiaofan Jiang
Columbia University
New York, New York, USA
jiang@ee.columbia.edu

ABSTRACT

The growth of smart devices is making typical homes more intelligent. In this work, in collaboration with therapists, we introduce a home-based *AI therapist* that takes advantage of the smart home environment to screen the day-to-day functioning and infer mental wellness of an occupant. Unlike existing "chatbot" works that identify the mental status of users through conversation, our *AI therapist* additionally leverages smart devices and sensors throughout the home to infer mental well-being and assesses a user's daily functioning. We propose a series of 37 dimensions of daily functioning, that our system observes through conversing with the user and detecting daily activity events using sensors and smart sensors throughout the home. Our system utilizes these 37 dimensions in conjunction with novel natural language processing architectures to detect abnormalities in mental status (e.g., angry or depressed), well-being, and daily functioning and generate responses to console users when abnormalities are detected. Through a series of user studies, we demonstrate that our system can converse with a user naturally, accurately detect abnormalities in well-being, and provide appropriate responses consoling users.

CCS CONCEPTS

• **Human-centered computing → Sound-based input / output**; • **Computer systems organization → Sensor networks**.

KEYWORDS

smart homes, daily functioning screening, mental health, artificial intelligence

ACM Reference Format:

Jingping Nie, Hanya Shao, Minghui Zhao, Stephen Xia, Matthias Preindl, and Xiaofan Jiang. 2022. Conversational AI Therapist for Daily Function Screening in Home Environments. In *1st ACM International Workshop on Intelligent Acoustic Systems and Applications (IASA'22), July 1, 2022, Portland, OR, USA.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3539490.3539603

IASA'22, July 1, 2022, Portland, OR, USA

ACM ISBN 978-1-4503-9403-1/22/07…$15.00
https://doi.org/10.1145/3539490.3539603

1 INTRODUCTION

Monitoring the mental well-being of people outside of visits to healthcare providers is important to improving our quality of life, especially for people who live alone or are exhibiting early signs of mental disorders. In 2021, the United States Census Bureau reported 37 million one-person households [1]. Additionally, the current COVID-19 pandemic has made in-person therapist and doctor visits more challenging. There are a variety of smart wearables and devices for monitoring physical and mental health [2–4]. Additionally, the growth of IoT devices and sensors has enabled around 70% of homes in the United States to have at least one smart home device [5]. A system that leverages smart home sensors and devices already present in the environment could enable continuous and comfortable mental health monitoring.

Researchers in psychology highlight the correlation between mental health adjustments and daily functioning, suggesting that people with mental health adjustment issues, in general, have more challenges and limitations in their day-to-day functioning, such as work, social relationships, and self-care [6, 7]. Various psychological measurements and standards, widely used by researchers and practitioners, assess daily functioning to evaluate mental health status, including the Daily Living Activities−20 (DLA-20), the Diagnostic and Statistical Manual of Mental Disorders (DSM)-IV, and the MOS 36-Item Short-Form Health Survey (SF-36) [8–10]. In collaboration with therapists, we define 37 dimensions of day-to-day functioning used in our system to screen the daily functioning of a user, which could also be used to infer mental health status. Researchers find routine monitoring helpful in mental health management and interventions [11]. Extensive studies have documented the effectiveness of measuring biosignals and markers, such as measuring heart rate variability, in monitoring specific mental health adjustments and conditions [12], but few have looked into environment based or non-invasive assessments of daily functioning to evaluate the general mental health status of users.

We propose a system for smart home environments that acts as an "AI therapist" to provide 24/7 daily functioning and mental status screening for users. Our system leverages smart devices

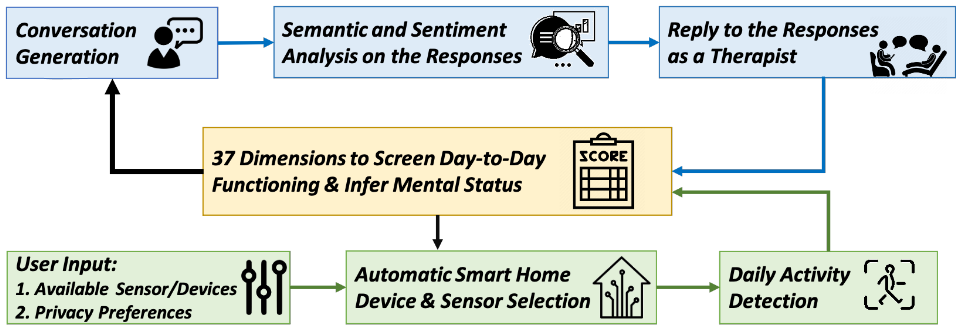

Figure 1: *AI therapist*'s system architecture. There are three modules in the proposed system: 1) 37 dimensions and scoring criteria to evaluate the day-to-day functioning and mental status of the user (yellow); 2) "Chat room" that leverages novel NLP techniques to converse with the user (blue); and 3) smart home devices and sensors that are used to monitor the condition of the user (green).

and ecosystems, that are becoming increasingly common in home environments, to monitor mental status and well-being and does not require users to wear anything additional on their bodies. This system takes advantage of the prevalence of low-cost smart home devices to 1) assess a user's daily functioning and monitor his/her mental well-being and 2) provide preliminary support at home. To monitor the condition of the user based on the 37 dimensions, our system primarily leverages privacy-aware sensors on smart home devices with lower fidelity, such as weight sensors and inertial measurement units (IMU), and using data-driven natural language processing (NLP) methods to "chat" like a therapist with the user through devices such as Alexa. We introduce a reinforcement learning (RL) based system that is able to learn from a user's responses and provide personalized interventions, such as modifying smart home settings, driving home robots, or reaching out to related personnel in an emergency. We integrate our system into the Amazon Alexa smart home ecosystem. Our system is also compatible with common smart home sensors and devices. We envision that our system can be compatible with a wide range of smart home ecosystems (e.g., Google Home, and Ring).

2 RELATED WORK

2.1 Psychological Measurements and Standards

Traditional screening of mental health involves an initial interview conducted by a trained interviewer, who is able to collect information, identify symptoms and signs, and suggest further psychological measurements [13]. There are several work that attempt to automate the initial mental health screening by replacing the interviewer with a series of close-ended web-based questions that assess major aspects of biopsychosocial well-being [14–16]. Patients using these web-based exams can obtain a quick evaluation on their mental health based on a few closed-ended questions; the assessments are limited to basic activities and events, but overlook the details that could be obtained from open-ended questions or face-to-face interactions, such as a person's affect or tone. Other assessments, such as the Montreal Cognitive Assessment (MOCA), Beck anxiety inventory, and Beck depression inventory, provide more specific screening options with emphasis on different mental health diagnoses [17–19]. Though questions from these assessments are generally easy to understand for lay persons, the results require a professional to interpret. Therefore, it's difficult to integrate such assessments into systems for the general population.

2.2 Acoustic Systems and Platforms for Daily Functioning Assessment

Audio is one of the most descriptive sensing modalities that can inform us about our health and physical world. Audio-based modalities are commonly used or being explored as an effective treatment for a variety of illnesses, such as cancer [20], or being used to detect a variety of illnesses, such as respiratory diseases [21] in adults or detecting periods of non-breathing during sleep, which is especially important for preventing sudden infant death syndrome in infants [22, 23].

There are a variety of systems and platforms that leverage acoustics and audio to monitor daily activities. Several works utilize microphone arrays, filtering techniques, and deep learning and machine learning classifiers to detect various sounds in a home, such as voice commands, speech, falls, and doors opening/closing [24–26]. Other works leverage sensor fusion and multiple sensing modalities, including leveraging both video and audio, to detect various activities of daily life [27]. In all of these applications, the key contributions involve developing new algorithms and systems for detecting specific events that could give insight into a person's well-being; for example, if a scream is detected, a person might have experienced a sudden or frightening event. None of these works assess a person's mental well-being. Additionally, [28] introduces a deep learning architecture for detecting relapses in depression using audio and visual cues. These works take a passive approach to monitoring daily life and mental well-being. In this work, we take an active approach. We introduce an AI therapist that can be integrated into smart home ecosystems that directly converses with users to understand a person's well-being and daily functioning and gives advice and consolation to users who may be exhibiting signs of poor well-being (e.g., stressed, sad, angry, etc.).

There are a few works that introduce chatbots that listen to users, assess their mental well-being, and provide responses to assist in improving mental well-being [29]. In this work, rather than take a

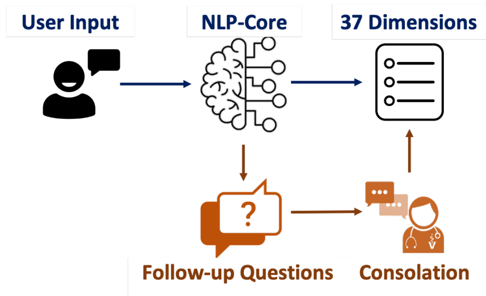

Figure 2: The workflow of the "chat room" module in our proposed *AI therapist* after the user replies to a question.

completely black-box deep neural network approach, we base our decisions based on 37 different dimensions of daily functioning that are commonly used to evaluate patients by mental health specialists. Additionally, unlike existing chatbots, we do not only use voice and speech; we also leverage other smart sensors and devices that may be found within the home environment to determine the best consolation and responses to users.

3 SYSTEM ARCHITECTURE

Figure 1 shows the architecture of our proposed "AI therapist" system, which is made up of three components denoted by three different colors. The 37 dimensions and evaluation criteria for day-to-day functioning screening is the critical component for assessing the mental state of users (yellow). These 37 dimensions were determined in collaboration with therapists, based off of their clinical experience. Table 1 lists the 37 dimensions of day-to-day functioning we propose.

The second component of our system is aimed at detecting and monitoring daily activities that correspond to the 37 dimensions (green). Our system automatically selects smart devices and sensors based on user preferences and the user's historical performance on the 37 dimensions. Our system prioritizes using more privacy-aware sensors unless specified by users, such as particulate matter sensors, alcohol gas sensors, and temperature sensors, to detect the common daily activities. The details for the study setup will be elaborated in Section 4.1.

The third component in this system is the "chat room" module of our proposed *AI therapist* (blue). This module converses with the user in a natural way after evaluating users on the 37 dimensions. Unlike traditional paper/web-based surveys and questionnaires used to assess daily activity functioning and mental health, our system takes advantage of smart home ecosystems (e.g., Alexa) and smartphones to "talk and chat" with the user in a more natural setting, mimicking an actual human therapist.

We use data-driven natural language processing (NLP) methods, incorporating speech-to-text and text-to-speech techniques, to perform mental health and daily functioning assessments in a more human-like way to increase the usability of our system and improve the user's overall experience, which makes them more likely to converse with our "AI therapist" on a regular basis. We integrate an RL-based recommender system to generate our human-like conversations considering: 1) the importance of the 37 dimensions scored by the therapists; 2) the user's previous responses and results; and 3) the daily activities detected by the smart home sensors. The

No.	Dimension
1	Maintaining stable weight;
2	Managing mood;
3	Taking medication as prescribed;
4	Participating primary and mental health care;
5	Organizing personal possessions & doing housework;
6	Talking to other people;
7	Expressing feelings to other people;
8	Managing personal safety;
9	Managing risk;
10	Following regular schedule for bedtime & sleeping enough;
11	Maintaining regular schedule for eating;
12	Managing work/school;
13	Having work-life balance;
14	Showing up for appointments and obligations;
15	Managing finance and items of value;
16	Getting adequate nutrition;
17	Problem solving and decision making capability;
18	Family support;
19	Family relationship;
20	Alcohol usage;
21	Tobacco usage;
22	Other substances usage;
23	Enjoying personal choices for leisure activities;
24	Creativity;
25	Participation in community;
26	Support from social network;
27	Relationship with friends and colleagues;
28	Managing boundaries in close relationship;
29	Managing sexual safety;
30	Productivity at work or school;
31	Motivation at work or school;
32	Coping skills to de-stress;
33	Exhibiting control over self-harming behaviour;
34	Law-abiding;
35	Managing legal issue;
36	Maintaining personal hygiene;
37	Doing exercises and sports;

Table 1: Dimensions to Screen Daily Functioning

recommender system asks the user questions based on the types of activities detected (or not detected) from the smart home sensors to obtain answers to as many of the 37 dimensions as possible. After the user responds to each question, the user responses will be analyzed semantically and for tone. With the analyzed results, the "AI therapist" will generate conversational replies to the user responses, to either inquire for further information or provide consolation if it determines that the user's mental well-being is affected (e.g., angry or sad) based on how it scores the 37 dimensions. Next, we describe in detail how our "chat room" accomplishes this.

3.1 "Chat room"

Figure 2 shows the data processing pipeline of the "chat room" module. After receiving responses from the user, we use the Google

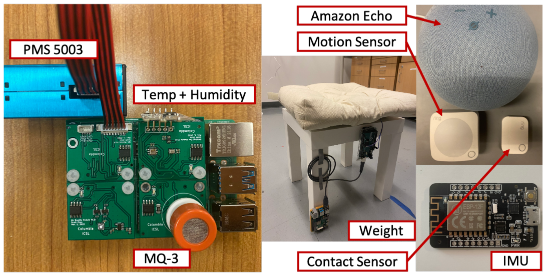

Figure 3: Sensors and devices placed in the in-lab experiment.

Cloud Speech API to convert speech to text [30]. We preprocessed this text to correct for punctuation and breaks in the text before handing it to the NLP-core. The NLP-core generates conversational-like dialogue based on user responses. The NLP-core is based on OpenAI GPT-3 model [31], which is a powerful cloud-based deep neural network language model that can be accessed from the cloud. We tune GPT-3 using our own dataset that maps user responses to one or more of the 37 dimensions and scores them. To perform this tuning, we created a dataset that includes: 1) 2000 user responses that are mapped to one or more of the 37 dimensions and scored. The mapping and scores are provided by the therapist. 2) 300 general responses (e.g., "Yes", "No", "Maybe", "I don't know", "Stop", and "I don't understand your question") that our "AI therapist" uses to converse with the user and transition from one question to the next question. To train our GPT-3 model to score each dimension, therapists provide scores ranging from 0 to 2. A score of 0 means that the user provided a response that does not indicate any abnormal mental status, while a score of 2 means that more attention should be paid by health care providers to the user in this dimension.

When the NLP-core detects any response that needs further attention from therapists (score of 2), the proposed *AI therapist* will generate follow-up questions accordingly to learn more about the user's situation. Using the newly obtained information from the follow-up questions, our system will *console* the user based on his/her situation, just like how a typical person would console someone going through hard times. We used around 150 common phrases for *consolation*, provided by our therapist, to fine-tune a second GPT-3 model to generate phrases for consoling users.

The *AI therapist* will generate a report after each conversation to imitate the diagnosis report that each health care provider needs to fill out after mental health screening. To preserve user privacy, the report only contains the following: 1) the original user responses; 2) the followup responses from the *AI therapist*; and 3) consolation phrases generated by the *AI therapist* when it detects issues in the user's well-being (responses that score 2).

4 STUDY DESIGN

In this section, we discuss the process of recruiting participants, experiment setup, and data collection process. Seven participants voluntarily participated in this study, including four males and three females aged between 21 and 30. All subjects live in single-occupant homes (or rooms). All subjects reported having normal hearing and cognition with no history of severe mental or physical disorders. This study is approved by the Institutional Review Board at Columbia University.

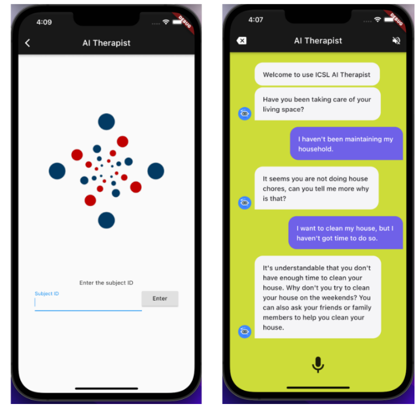

Figure 4: The smartphone version of the "chat room" module in the proposed *AI therapist*.

4.1 Study Setup

The duration of the study was one week. Each subject participated in two in-lab experiment sessions (at the beginning and the end of the week) with various smart home devices. For the in-lab experiment setup, we have:

- 4 pieces of 50 kg half-bridge load cell body scale weighting sensor with an HX711 amplifier embedded in a couch cushion to detect significant weight changes (*Dimension 1*);
- 1 alcohol gas sensor (MQ-3) to detect if the user engages in heavy drinking (*Dimension 20*);
- 1 particulate matter sensor (PMS 5003) to obtain the number of suspended particles in the air and detect if the user smokes alone (*Dimension 21*);
- 1 contact sensor and 1 motion sensor to monitor the mobility of the user and check if the user leaves home regularly (*Dimension 25*);
- 1 temperature and 1 humidity sensor to sense if the user has healthy showering habits (*Dimension 36*);
- 1 Amazon Echo integrated with our "chat room" module.

We used an electric kettle to simulate the showering behavior for the in-lab experiment. Figure 3 shows the sensors we deployed in the in-lab experiments.

In addition, two out of seven subjects brought inertial measurement unit (IMU) devices back home and placed them under their bedsheets to measure sleep quality (*Dimension 10*) for a week. We also created a smartphone version of the "chat room" module to enable the subjects to converse with our "AI therapist" once a day at home. Figure 4 shows the smartphone version of our "chatroom".

4.2 Data Collection

In the first in-lab session, we obtained informed consent from each subject. Subjects were told to go about their daily life as they normally would and chat in our "chat room" via an Alexa Echo. The experiment duration for each subject was 2 hours. Lab members and therapists recorded events with timestamps, such as when the subject got up to leave the "apartment" or drink water. For the two

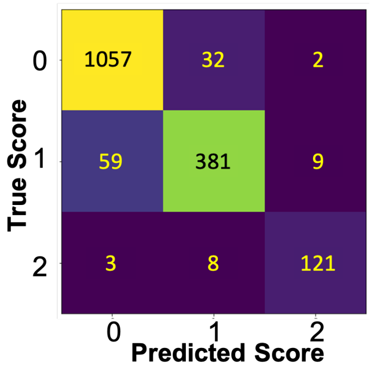

Figure 5: Confusion Matrix of Scoring.

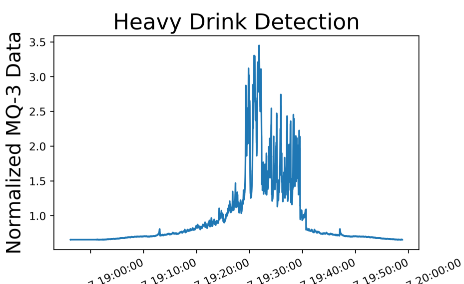

Figure 6: An example of an episode of heavy drinking.

subjects who have placed IMUs under their bedsheets, also wore Oura Rings as a baseline measurement of sleep quality. [32] shows that Oura Ring is one of the most useful consumer wearables for sleep monitoring.

The smartphone version of our "chat room" was introduced to the subject at the end of the first in-lab session. Subjects were instructed to converse with the smartphone "chat room" every day during the study. Conversations between the user and "chat room" were recorded and saved to our lab server in text format only. At the end of the week, subjects came back in for a second in-lab session, where they were given the same instructions as in the first in-lab session. At the end of the second in-lab session, a questionnaire was given to each subject to collect feedback, comments, and qualitative appraisal. In total, we obtained 14 sets of smart home sensor data and 49 sessions of conversation records.

5 EVALUATION

In this section, we evaluate the proposed system from two perspectives: 1) the performance of each module in *AI therapist*, and 2) a user experience study.

5.1 NLP-core Classification and Scoring Accuracy

In the 49 conversation sessions obtained from all of our subjects, there were 1711 question (spoken by the *AI therapist*) and response (given by the subject) pairs. Out of these 1711 questions, each attempting to observe one of the 37 dimensions of daily functioning we proposed, 1672 were correctly classified by the NLP-core into the correct dimension. Figure 5 shows the confusion matrix for scoring

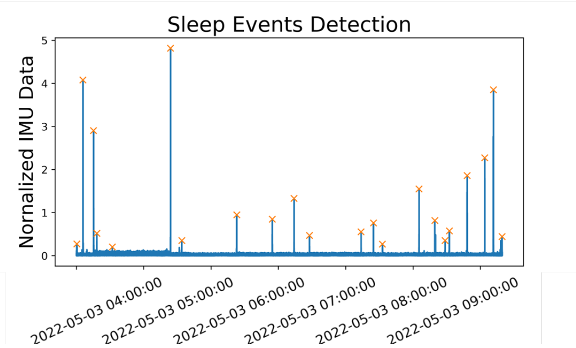

Figure 7: Sleep event (e.g., rolling over) detection. Detected sleep events are marked in orange crosses and the preprocessed IMU data is plotted in blue.

Activity	Correct	Type I Error	Type II Error
(A)	6	0	1
(B)	21	2	3
(C)	14	0	2
(D)	17	0	0
(E)	12	0	0

Table 2: Daily activity detection accuracy (# instances). (Type I Error: detect a false event. Type II Error: fail to detect an event.)

the 1672 responses. NLP-core can achieve over 97% classification accuracy and over 92% scoring accuracy.

5.2 Daily Activity Detection using Smart Home Devices

As mentioned in Section 4.1, using our sensor deployment, our system can detect 6 daily activities including: (A) significant weight changes (over 5% of body weight); (B) episodes of heavy drinking; (C) tobacco smoking instances; (D) home-leaving incidences; (E) showering instances; and (F) sleep events. Due to page limit, we present only the raw signal of a MQ-3 alcohol sensor and an IMU corresponding to a time period where a drinking event and sleep event occurred in Figures 6 and 7.

Compared to the ground-truth events recorded by the lab members and the therapists, the activity detection performance for activity (A) to (E) is shown in Table 2. As for (F) (sleep events), we observed that the frequency of sleep events is highly correlated with the deep, light, and rapid eye movements (REM) measured by the Oura ring.

5.3 User Experience Study

In addition to the quantitative evaluation of the system, we also collect qualitative feedback from the subjects. We asked the 7 subjects to rate *AI therapist* from (poor) 1 to 5 (excellent) in 4 aspects: 1) *overall* experience; 2) *willingness* to continue using the system in the future; 3) if the system is *easy* to use; and 4) rating how well the *AI therapist* or "chat room" *consoled* them. Figure 8 shows the

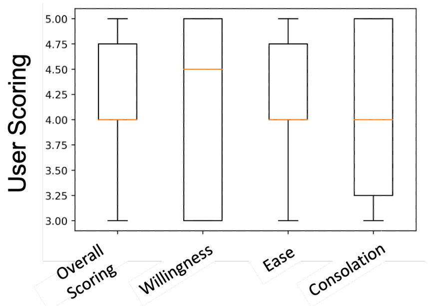

Figure 8: User study results.

results from the user study. We see that the average scores for all categories are above 4, showing that the usability of our system was well-received.

6 CONCLUSION AND FUTURE WORK

We introduce 37-dimensions to screen day-to-day functioning and infer mental status in collaboration with therapists. An *AI therapist* is proposed, which can assess a user's daily functioning and mental health based on a combination of direct dialogue with the user and information acquired through smart home devices. We demonstrate that our "chat room" can have natural conversations with users and assess their responses semantically and emotionally.

In the future, we are going to take advantage of the smart home devices and home robots (e.g., Roomba and drone) to conduct preliminary interventions to interact with the user and help improve the user's wellness. In addition, we also plan to enable the proposed "AI therapist" in multi-occupants households.

ACKNOWLEDGMENTS

This research was partially supported by the National Science Foundation under Grant Numbers CNS-1704899, CNS-1815274, CNS-11943396, and CNS-1837022. The views and conclusions contained here are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of Columbia University, NSF, or the U.S. Government or any of its agencies.

REFERENCES

[1] Census Bureau Releases New Estimates on America's Families and Living Arrangements. https://www.census.gov/newsroom/press-releases/2021/families-and-living-arrangements.html#:~:text=There%20were%2037%20million%20one, over%20the%20last%20two%20decades., 2021. [Online; accessed 24-April-2022].

[2] Jingping Nie, Yanchen Liu, Yigong Hu, Yuanyuting Wang, Stephen Xia, Matthias Preindl, and Xiaofan Jiang. Spiders+: A light-weight, wireless, and low-cost glasses-based wearable platform for emotion sensing and bio-signal acquisition. *Pervasive and Mobile Computing*, 75:101424, 2021.

[3] Stephen Xia, Peter Wei, Jordan Misael Vega, and Xiaofan Jiang. Spindles+: An adaptive and personalized system for leg shake detection. *Smart Health*, 9-10:204–218, 2018. CHASE 2018 Special Issue.

[4] Stephen Xia, Jingping Nie, and Xiaofan Jiang. Csafe: An intelligent audio wearable platform for improving construction worker safety in urban environments. In *Proceedings of the 20th International Conference on Information Processing in Sensor Networks (Co-Located with CPS-IoT Week 2021)*, IPSN '21, page 207–221, New York, NY, USA, 2021. Association for Computing Machinery.

[5] Stephen Xia, Rishikanth Chandrasekaran, Yanchen Liu, Chenye Yang, Tajana Simunic Rosing, and Xiaofan Jiang. A drone-based system for intelligent and autonomous homes. In *Proceedings of the 19th ACM Conference on Embedded Networked Sensor Systems*, pages 349–350, 2021.

[6] Christine A Helfrich, Glenn T Fujiura, and Violet Rutkowski-Kmitta. Mental health disorders and functioning of women in domestic violence shelters. *Journal of Interpersonal Violence*, 23(4):437–453, 2008.

[7] Lisa J Bible, Kristin A Casper, Jennifer L Seifert, and Kyle A Porter. Assessment of self-care and medication adherence in individuals with mental health conditions. *Journal of the American Pharmacists Association*, 57(3):S203–S210, 2017.

[8] Whitney Clausen, Shinobu Watanabe-Galloway, M Bill Baerentzen, and Denise H Britigan. Health literacy among people with serious mental illness. *Community mental health journal*, 52(4):399–405, 2016.

[9] Samuel B Guze. Diagnostic and statistical manual of mental disorders, (dsm-iv). *American Journal of Psychiatry*, 152(8):1228–1228, 1995.

[10] Colleen A McHorney, John E Ware Jr, and Anastasia E Raczek. The mos 36-item short-form health survey (sf-36): Ii. psychometric and clinical tests of validity in measuring physical and mental health constructs. *Medical care*, pages 247–263, 1993.

[11] Ingrid VE Carlier, Denise Meuldijk, Irene M Van Vliet, Esther Van Fenema, Nic JA Van der Wee, and Frans G Zitman. Routine outcome monitoring and feedback on physical or mental health status: evidence and theory. *Journal of evaluation in clinical practice*, 18(1):104–110, 2012.

[12] Blake Anthony Hickey, Taryn Chalmers, Phillip Newton, Chin-Teng Lin, David Sibbritt, Craig S McLachlan, Roderick Clifton-Bligh, John Morley, and Sara Lal. Smart devices and wearable technologies to detect and monitor mental health conditions and stress: A systematic review. *Sensors*, 21(10):3461, 2021.

[13] James Morrison. *The first interview*. Guilford Publications, 2016.

[14] Meredith Wallace, Mary Shelkey, et al. Katz index of independence in activities of daily living (adl). *Urol Nurs*, 27(1):93–94, 2007.

[15] Evidence-Based Mental Health Solutions - MindWise Innovations. https://www.mindwise.org, 2022. [Online; accessed 25-May-2022].

[16] Noom: Stop dieting. Get life-long results. https://www.noom.com, 2022. [Online; accessed 25-May-2022].

[17] Aaron T Beck, Robert A Steer, Gregory K Brown, et al. *Beck depression inventory*. Harcourt Brace Jovanovich New York:, 1987.

[18] Aaron T Beck, Norman Epstein, Gary Brown, and Robert Steer. Beck anxiety inventory. *Journal of consulting and clinical psychology*, 1993.

[19] Ziad S Nasreddine, Natalie A Phillips, Valérie Bédirian, Simon Charbonneau, Victor Whitehead, Isabelle Collin, Jeffrey L Cummings, and Howard Chertkow. The montreal cognitive assessment, moca: a brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, 53(4):695–699, 2005.

[20] Andrew KW Wood and Chandra M Sehgal. A review of low-intensity ultrasound for cancer therapy. *Ultrasound in medicine & biology*, 41(4):905–928, 2015.

[21] Arpan Srivastava, Sonakshi Jain, Ryan Miranda, Shruti Patil, Sharnil Pandya, and Ketan Kotecha. Deep learning based respiratory sound analysis for detection of chronic obstructive pulmonary disease. *PeerJ Computer Science*, 7, 2021.

[22] Shih-Hong Li, Bor-Shing Lin, Chen-Han Tsai, Cheng-Ta Yang, and Bor-Shyh Lin. Design of wearable breathing sound monitoring system for real-time wheeze detection. *Sensors*, 17(1), 2017.

[23] Takahiro Emoto, Udantha R. Abeyratne, Kenichiro Kawano, Takuya Okada, Osamu Jinnouchi, and Ikuji Kawata. Detection of sleep breathing sound based on artificial neural network analysis. *Biomedical Signal Processing and Control*, 41:81–89, 2018.

[24] Ivan Miguel Pires, Rui Santos, Nuno Pombo, Nuno M Garcia, Francisco Flórez-Revuelta, Susanna Spinsante, Rossitza Goleva, and Eftim Zdravevski. Recognition of activities of daily living based on environmental analyses using audio fingerprinting techniques: A systematic review. *Sensors*, 18(1):160, 2018.

[25] Rosa Ma Alsina-Pagès, Joan Navarro, Francesc Alías, and Marcos Hervás. homesound: Real-time audio event detection based on high performance computing for behaviour and surveillance remote monitoring. *Sensors*, 17(4):854, 2017.

[26] Georgios Siantikos, Theodoros Giannakopoulos, and Stasinos Konstantopoulos. Monitoring activities of daily living using audio analysis and a raspberrypi: a use case on bathroom activity monitoring. In *International Conference on Information and Communication Technologies for Ageing Well and e-Health*, pages 20–32. Springer, 2016.

[27] Jiaxuan Wu, Yunfei Feng, and Peng Sun. Sensor fusion for recognition of activities of daily living. *Sensors*, 18(11):4029, 2018.

[28] Muhammad Muzammel, Alice Othmani, Himadri Mukherjee, and Hanan Salam. Identification of signs of depression relapse using audio-visual cues: A preliminary study. In *2021 IEEE 34th International Symposium on Computer-Based Medical Systems (CBMS)*, pages 62–67. IEEE, 2021.

[29] Dongkeon Lee, Kyo-Joong Oh, and Ho-Jin Choi. The chatbot feels you-a counseling service using emotional response generation. In *2017 IEEE international conference on big data and smart computing (BigComp)*, pages 437–440. IEEE, 2017.

[30] Speech-to-Text. https://cloud.google.com/speech-to-text, 2022. [Online; accessed 24-April-2022].

[31] Build next-gen apps with OpenAI's powerful models. https://openai.com/api/, 2022. [Online; accessed 24-April-2022].

[32] Stanisław Saganowski, Przemysław Kazienko, Maciej Dziezyc, Patrycja Jakimów, Joanna Komoszynska, Weronika Michalska, Anna Dutkowiak, A Polak, Adam Dziadek, and Michal Ujma. Review of consumer wearables in emotion, stress, meditation, sleep, and activity detection and analysis. *arXiv preprint arXiv:2005.00093*, 2020.