

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2020.DOI

Smartphone Sensing for the Well-being of Young Adults: A Review

LAKMAL MEEGAHAPOLA^{1,2}, DANIEL GATICA-PEREZ^{1,2}

¹Idiap Research Institute, Martigny, Switzerland

²École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

Corresponding author: Lakmal Meegahapola (e-mail: lmeegahapola@idiap.ch)

This work was funded by the European Commission's Horizon 2020 project "WeNet: The Internet of Us", under grant agreement 823783.

• **ABSTRACT** Over the years, mobile phones have become versatile devices with a multitude of capabilities due to the plethora of embedded sensors that enable them to capture rich data unobtrusively. In a world where people are more conscious regarding their health and well-being, the pervasiveness of smartphones has enabled researchers to build apps that assist people to live healthier lifestyles, and to diagnose and monitor various health conditions. Motivated by the high smartphone coverage among young adults and the unique issues they face, in this review paper, we focus on studies that have used smartphone sensing for the well-being of young adults. We analyze existing work in the domain from two perspectives, namely Data Perspective and System Perspective. For both these perspectives, we propose taxonomies motivated from human science literature, which enable to identify important study areas. Furthermore, we emphasize the importance of diversity-awareness in smartphone sensing, and provide insights and future directions for researchers in ubiquitous and mobile computing, and especially to new researchers who want to understand the basics of smartphone sensing research targeting the well-being of young adults.

• **INDEX TERMS** smartphone sensing, mobile sensing, health, well being, young adults, survey, review

I. INTRODUCTION

Smartphones have been rapidly evolving during the past decade due to advancement of technology in a multitude of disciplines such as hardware (CPU, GPU) [1], sensing [2], [3], computer vision [4], [5], [6], deep learning [7], [8], and human-computer interaction [9]–[11]. These advancements, together with the benefits they provide have made smartphones integral components of the lives of people. A study shows that smartphone coverage in young adults aged 18-29 in USA is 96% [12], and other analyses provide insights as to how smartphones are affecting human behavior while elucidating the close connection of smartphones to millions of people [13]. Smartphones are more user friendly and interactive, while also capable of collecting and processing contextual data in real-time [14]. In addition, app distribution platforms such as Google Play Store [15] and Apple AppStore [16] have enabled application developers and researchers to distribute smartphone apps to millions of people worldwide.

Wearable sensing in health monitoring emerged as a trending topic two decades ago with researchers focusing on using wearable sensors for monitoring behavioral patterns,

health conditions, and lifestyles [17]–[19]. The use of that research in real world settings was rare due to many reasons such as the high cost involved in creating wearable devices, the mindset of people regarding wearable devices, and the inability to distribute devices to wide populations. Hence, many of the research efforts were done in controlled lab settings. With the widespread use of mobile phones in the 2000s, several of these issues went out of the equation as more people, specifically *young adults*, embraced mobile phones. This spanned an emerging literature regarding the utility of mobile phone sensing for large scale applications. A pioneering study on this regard is Reality Mining [20] that demonstrated the utility of phone sensing to collect contextual data passively (gps, bluetooth traces, app usage, charging events), in the wild (out of the lab setting, with 100+ people), for an extended period of time (1 year). Works such as UbiFit Garden [21] and MyExperience [22] further demonstrated the capabilities of mobile phones in processing data obtained with external and internal sensors combined with self-reports for behavioral analysis. The emergence of smartphones with more sensing capabilities and interactions compared to traditional phones injected new momentum into

mobile sensing research with a shift towards *Smartphone Sensing* [23].

A. THE WELL-BEING OF YOUNG ADULTS

People of various age groups have different lifestyles, behavioral patterns, thought processes, and biological characteristics [24], [25]. Young adults (even though there is not a unique definition for the age of young adults, we follow the age range 18-35 suggested by Petry et al. [26]) go through different circumstances in life compared to older generations, and this is reflected in the activities, social interactions, dietary habits, and even the physical and mental health conditions they have to face [27]–[32]. Many young adults are doing their studies, in their early career, in the first few years of marriage, unemployed, or a combination of the aforementioned. Considering this stage in life, it is known that stress, anxiety, depression, obesity, alcohol/smoking/drug addictions, and unhealthy food habits are common among young adults [31] [33]–[35], and the reasons why they face these issues might be different to why someone from another age group would face the same type of issue. For example, it is common to see undergraduate students having depression/stress due to steep work load they have in university, while for a person aged 40-50, they might have depression/stress due to family or job related conditions. Further, young adults use social media and smartphones far more than older generations [36], and prior research suggests differences on the way people use the phone depending on age [37], [38]. In addition, even international organizations such as UNICEF have emphasized the need for human-centered design of products and services for young people [39]. Hence, the criteria to quantify various health/well-being conditions of people using smartphone sensing would in principle be different, and needs special consideration for a range of smartphone-related issues starting from sensor selection, app design, and deployment strategies to data analysis techniques, while keeping in mind the *diversity* existing within each age group. Therefore, in this paper, we focus on smartphone sensing research that has dealt with health and well-being, specifically of *young adults*. We detail the state-of-the-art in smartphone sensing studies and discuss the strengths and weaknesses of approaches. Further, we propose a taxonomy for smartphone sensing studies that reflects changes in smartphone usage behaviors to highlight areas in which young adults can be focused. Moreover, we emphasize the need for diversity-awareness in building health-related sensing paradigms, identify research gaps, and discuss future directions of research to advance the field.

B. BACKGROUND AND MOTIVATION

Previous Surveys on Smartphone Sensing. After Chen and Kotz's early survey [40] regarding mobile sensing, one of the best known reviews in smartphone sensing is by Lane et al. [23] where they provided details regarding sensors of the smartphone that can be utilized for sensing, application eco-systems, and sensing paradigms. McCracken et

al. [41] studied the use of optical smartphone sensing in resource-limited settings, reviewing optical sensors available in smartphones and highlighting important aspects regarding the limitations of this domain. In another survey, Christin et al. [42] analyzed privacy aspects of smartphone sensing in existing applications and identified some of the limitations. Some other reviews in the recent past that involve smartphone sensing discuss domains such as internet of things [43], incentive mechanisms for smartphone based crowd sensing [44], dynamics of mobile cloud computing [45], usage of smartphone and mobile sensing data for urban studies [46], monitoring human movement [47], and smartphone sensing in different application domains [48].

Previous Surveys on Smartphone Sensing for Well-being.

Triantafyllidis et al. [49] provided a comprehensive analysis regarding smartphone usage in healthcare where they emphasized the trend of pervasive healthcare (health anywhere, anytime using smartphones) instead of the current institutional based model which is a burden for older generations due to the high prevalence of chronic diseases. Because the review is about general health/well-being, studies on older generations, infants, and young adults are discussed, and these studies use sensing modalities in different ways to capture different behaviors of these user groups. Cornet et al. [50] conducted a similar survey regarding smartphone based passive sensing modalities for general health and well-being. Going a step further from [49], the review in [50] categorized the well-being literature based on different types such as mental health, sleep, general health/well-being, and other domains. This analysis does not include details regarding how sensing modalities can be used in tandem with other self-reported data, and lacks in-depth analysis regarding the use of sensing modalities. Harari et al. [51] discussed usage of smartphone sensing for studying behavior of everyday life in a brief review. They have categorized several sensing modalities based on the type of sensed behavior (e.g.: physical movements, social interactions, daily activities). Aung et al. [52] surveyed the literature regarding sensing mental health related symptoms and intervention techniques. Trifan et al. [53] described smartphone sensor types used in different health and well-being related conditions in their survey. Further, Mohr et al. [54] surveyed machine learning and feature transformation techniques used in mobile sensing research targeting mental health. The focus of this survey was wider because it considered all types of ubiquitous sensors. Finally, Can et al. [55] surveyed literature about detecting stress using smartphone and wearable sensing techniques.

As discussed above, no review/survey paper on smartphone sensing and mobile sensing discusses studies focusing on the diversity attributes of people (e.g.: age, gender, country, etc.). In our view, this is an important domain for researchers, as diversity-aware smartphone sensing paradigms are essential to design comprehensive studies about the behavior of different user groups, conduct more in-depth analyses, and leverage human diversity to create rich, human-centered applications. With smartphones becoming more so-

p sophisticated and feature rich, the group of people who have benefited and been affected the most are undoubtedly young adults, while smartphone coverage among young adults is an ever increasing figure [12]. In our opinion, leveraging this smartphone ubiquity among young adults by careful *data* acquisition and *system* design considerations would enable to conduct studies to deepen our understanding of young adults's practices and needs, and the development of better informed systems for support and interaction.

C. SCOPE OF THE PAPER

The following inclusion criteria (IC) were chosen in order to define the scope of the paper.

IC1: Passive Sensing – The smartphone sensing studies in discussion should have used at least one passive sensing modality. Smartphone based studies that only utilize diaries, recommendation engines, or questionnaires without any passive sensing were discarded in order to focus on the sensing capabilities of modern smartphones.

IC2: Using other data sources – Studies might have used external sensing modalities (e.g.: smartwatch, activity trackers, wrist bands, etc.) or external data sources (e.g.: foursquare, instagram, weather data, etc.). However, the main theme of the study should be based on a smartphone sensing app, and external data sources serve as complementary data to enhance the analysis or to improve performance.

IC3: Venue of publication – We considered publications in computer science, behavioral science, medicine, and psychology venues. While analyzing computer science applications from the contribution they make to the domain, we draw findings from literature of other domains to understand how these findings can be adapted towards improving ubiquitous and mobile computing research.

IC4: Targeting Young Adults – Young adults are the main beneficiary of the study, or at least have been part of the study cohort. We included studies that have not particularly mentioned a target audience, but used a young adult population in the cohort. This criteria also narrows the scope of the review in terms of the type of health and well-being issues addressed (e.g. not including aspects such as heart diseases, Alzheimer's, stroke, etc.).

IC5: Datasets – Some smartphone sensing datasets have been used in several studies over the course of 5 - 6 years. In such cases, some studies that used the same dataset were not included unless the corresponding contributions are significant compared to previous studies on the same data.

IC6: Timeliness – We observed that certain studies in smartphone sensing carried out in the first decade of the century are not necessarily relevant in today's context. Hence, to make sure that the discussion presented here are timely, we review studies from the last decade (2010 - 2020). This also ensured that the research was conducted in a time period when smartphone usage among young adults in the world was high, and that smartphone sensing related studies appeared in different sub-domains.

IC7: Language – We only considered papers written in English.

IC8: Number of Participants – We used 10 as the minimum number of participants for a study to be considered in this review.

The approach we used to search and select papers for this review was the following:

Search: We used Google Scholar, PubMed, and Scopus to search for studies with different search terms related to the domain such as mobile phone sensing, smartphone sensing, smartphone sensing health, mobile sensing well-being, health young adults, etc.

Screening: We screened titles, abstracts, and contributions of more than 400 articles.

Eligibility: We used IC1-IC8 to check studies for eligibility. This process left us with 26 studies for the primary analysis.

It should be noted that in cases where certain details were not clear, we contacted the original authors to ensure that we include correct information in this review. Moreover, this study is not a systematic review, but a literature review that does not attempt to answer any specific research question¹.

In Section II, we lay out the organization of this study, and in Section III, we provide a summary regarding the selected set of papers. Further, under Data Perspective, in Section IV and Section V, passive sensing and self-report modalities respectively are discussed from a human science perspective. Then, in Section VI, we examine the System Perspective. Finally, in Section VII, we discuss regarding insights and directions for future work.

II. ORGANIZATION OF THE REVIEW

A. THE HUMAN SCIENCE PERSPECTIVE

1) Pillars of Data for Smartphone Sensing.

For purposes of organization of the review, we first borrow from Lewin's Field Theory [56]. This classic model proposed that human behavior is affected by personal and environmental factors, broadly expressed as $B = f(P, E)$, where B , P , and E stand for behavior, person, and environment, respectively. While oversimplifying the complexity and nuances of all the involved processes, this model allows for an organization of the mobile sensing work reviewed here. Descriptions of the three terms are given below.

Behavior: Changes in the life space of an individual as a result from changes in either the environment or the person. Hence, behavior represents activities and actions performed by people.

Person: This aspect is defined according to: (a) attributes, traits and characteristics of the individual such as beliefs, abilities, values, emotions, needs, and other person aspects; (b) static characteristics that intrinsically affect such traits like age, sex, bodily characteristics (height, weight), etc.; and (c) individuals perception of the context that they are in.

¹https://guides.libraries.psu.edu/ld.php?content_id=36146097

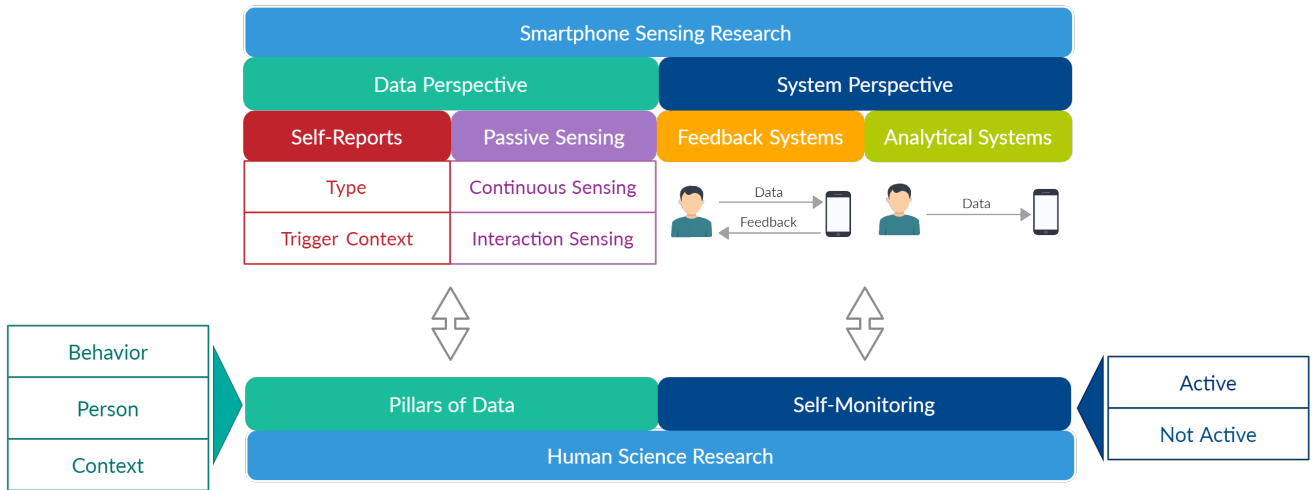


FIGURE 1: Taxonomy of Smartphone Sensing Studies for the Well-being of Young Adults from a Human Science Perspective

Environment: Conscious as well as unconscious entities in a human’s environment or in other terms, the objective context at which the person perceives and acts.

Drawing motivation from Lewin’s field theory, we consider three “pillars of data” to study smartphone sensing research for the well-being of young adults, namely (1) **Behavior (B)**, (2) **Person (P)**, and (3) **Context (C)**. We mapped *environment* from Lewin’s theory into context data pillar in the context of smartphone sensing. We mapped person to P to represent the physical, mental, and social aspects of people in addition to attributes such as age, gender, and this mapping allows us to represent sensors and data sources in smartphone sensing literature as proxies to different pillars of data. Hence, we use these pillars of data in the context of smartphone sensing as a framework throughout this paper. This enables us to analyze studies in the domain using a well-grounded framework, where all the sensing data collected during the study falling into one or more of the pillars. As an example, data from the accelerometer sensor of a smartphone can be taken as a proxy regarding physical activity done by an individual, hence falling under the behavior pillar. Moreover, location coordinates fall under context pillar. If they are further processed to obtain a semantic meaning, location can be used to determine travel patterns of the users which can be categorized under behavior pillar, because changes in location could also mean a distance traveled by users, hence a behavior. Further, a self-report could reflect the emotion of a user, and it belongs to the person pillar. Hence, if a study uses features generated from accelerometer and location sensors to infer/analyze emotions of a person (target domain of the study), that reflects the use of behavior and context features to infer an attribute belonging to the person pillar. As explained above, various features generated from a single sensor might provide data regarding one or more pillars depending on the way data is processed and features are engineered. Hence, we use these three pillars of data to

study the current body of research from the data perspective as explained in Section II-B.

2) Self-Monitoring as part of Mobile Sensing

There are many well-known theories in social and cognitive sciences regarding self-monitoring and behavioral change [57]–[62]. Generally, self-monitoring is the capacity to quantify, observe, and evaluate one’s behavior to facilitate behavioral change. Michie et al. [58] defined self-monitoring as keeping a record of a specified behavior as a method for changing behavior. The Social Cognitive Theory (SCT) of Self-Regulation by Bandura [61] outlined that when people observe a behavior and consequences of the behavior, they learn from the events to guide their subsequent behaviors, with the potential to enable behavioral change. In addition, this body of literature presents factors that affect the quality of self-monitoring such as informativeness, regularity, proximity, and accuracy. Drawing from the essence of these ideas, we argue that mobile sensing systems that allow users to view statistics and feedback on apps, facilitate self-monitoring, thus assisting users interested in specific behavioral changes. Hence, there are two types of mobile sensing systems, one in which self-monitoring is active; and another one in which self-monitoring is not active. By processing the behavior, person, and context related data captured via smartphones, apps with active self-monitoring allow users to monitor themselves and reflect upon their behavior using feedback given in the app. Hence, while simplifying the nuances behind complex theories, we use *self-monitoring* as a concept to study the current body of research from the system perspective as explained in Section II-C.

B. THE DATA PERSPECTIVE

The Data Perspective considers the data flow in smartphone sensing studies. Under this perspective, we primarily consider an existing classification: **Passive Sensing** (Section IV)

and **Self-Reports** (Section V). However, we analyze each of the components from a new perspective using pillars of data. Further, we propose a taxonomy to segregate passive sensing into: (1) Continuous Sensing, that involves embedded sensors that acquire data (e.g. accelerometer, gyroscope, location, ambient light, etc), and (2) Interaction Sensing, that captures user interactions/usage with/of the phone without utilizing any embedded sensor (e.g. app usage, phone calls, messages, typing events, etc.).

The data collected in a study could belong to either one or more pillars mentioned in Section II-A. Even though continuous and interaction sensing have been named as just passive sensing in prior literature, we make this distinction to highlight some aspects regarding how smartphone sensing studies on health and well-being can benefit young adults.

It is important to understand how data are used in smartphone sensing. In all the studies, data from different sensors are used as proxies to different phenomena [63] related to person, context, or behavior aspects. For example, if the sensed phenomena are activity level, brightness of the surrounding, sociability, and indoor mobility, sensing modalities could be accelerometer, ambient light sensor, messaging app usage, and WiFi access point connectivity. Hence, the data perspective of a smartphone sensing study involves obtaining data from the smartphone regarding the three pillars of data in behavioral modeling. Passive sensing directly obtains data regarding these pillars from users unobtrusively, while self-report data act as proxies and ground truth for the three pillars, and are obtained explicitly from users. Hence, as a summary, smartphone sensing studies involve obtaining data from young adults belonging to three pillars using passive sensing and self-reports in order to analyze an unknown attribute belonging to one of the pillars.

C. THE SYSTEM PERSPECTIVE

Smartphone Sensing studies involve smartphone app based systems. Using self-reports and passive sensing techniques mentioned in Section II-B, smartphones acquire multi-dimensional data. The majority of such systems store this information for off-the-shelf analysis that is done at the end of the deployment phase. However, some systems process these data on-device or in cloud to provide feedback to users regarding their health and well-being state, hence allowing self-monitoring. If we combine (a) data acquisition by the smartphone and (b) value-added feedback given to users, together with theories regarding self-monitoring and behavior change, it is possible to divide literature as (a) Feedback Systems, where self-monitoring is active, and (b) Analytical Systems, where self-monitoring should not be active.

Feedback systems (FSys) do data analysis, training or inference (either in-device or on servers; done manually or in an automated manner) during the deployment phase. These systems often use results of the data analysis or inference primarily to give users in-app feedback. Some systems provide feedback regarding user behavior, mental health, and routines in order to motivate them to use the app further [64],

[65], develop self insight [66], [67], provide behavioral intervention strategies [68], and for clinical interventions [69]. In other terms, this type of systems that provide feedback affect normal user behavior, and can be used for interventions as well. However, it should also be noted that not all feedback systems are used for interventions because some apps do not necessarily provide any active intervention even though they summarize sensed details in the app. A feedback system with active self-monitoring should provide app users with useful feedback, and observe how people benefit from the feedback by helping them change their behavior in a positive course. Moreover, from a behavioral science perspective, systems of this sort cannot be used to evaluate a hypothesis regarding the presence or absence of a certain attribute in people because self-monitoring is active, and hence it could create biases and behavioral change in people.

Analytical systems (ASys) are systems that do data analysis and inference off-the-shelf after collecting the data from the study. These systems often run with no (or less) data analysis during the deployment (with the exception being activity inference from accelerometer data [7], [68]–[70]), and even if data analysis is done, results are not directly conveyed to app users, hence avoiding any behavioral change via self-monitoring. A ASys should not have allow self-monitoring, and hence, smartphone app users should not be motivated to change their behavior. Moreover, for an ASys, it is better if app users do not necessarily know what specific hypothesis is being tested by researchers, to reduce potential biases in the mind of study participants (who need to understand the general goals of the study as part of the informed consent process). This would help to obtain better quality data which reflect the real behavior of study participants, hence leading to better studies regarding the validation of hypotheses.

TABLE 1: Study Details of Smartphone Sensing Research for the Well-being of Young Adults (YA)

Reference	Year	Domain	Pillar	Study Details					
				Sample Size	Sample Diversity	Target Audience	Study Length	Feedback	Experts/Theory
Rachuri et al. [71] EmotionSense	2010	Emotion	P	18	-Members of CS Department	All	10 days	✓	UK
Moturu et al. [72] Funf	2011	Well being	P,B	54	-Graduate Students and Spouses	All	1 month	✓	USA
Lu et al. [73] StressSense	2012	Stress	P	14	-University students -4 Men and 10 Women -13 Undergrads and 1 PhD	All	Job Interviews	✓	Switzerland
Lin et al. [66] BeWell+	2012	Well being ²	B	27	-Aged 21-37 -16 Men and 11 Women -9% CS Department, 34% Doctors and Medic. Researchers and 57% Grad. Students	All	19 days	✓	USA
Ma et al. [65] MoodMiner	2012	Mood	P	15	-University Students and Teachers, -Urban White Collar Workers	All	30 days	✓	China
LiKamWa et al. [64] Moodscope	2013	Mood	P	32	-24 students and 8 other occupations -25 from China, 7 from USA	All	2 months	✓	China, USA
Sano et al. [74]	2013	Stress	P	18	-Average age 28 -15 Men and 3 Women	All	5 days	✓	USA
Bogomolov et al. [75]	2013	Emotion ³	P	117	-University (Grad) Students -16 countries	YA	18 months	✓	USA
Wang et al. [70] StudentLife	2014	Well Being ⁴	P,B	48	-University students -One CS Class -30 Undergrads and 18 Grad Students -38 Men and 10 Women -23 Caucasians, 23 Asians and 2 Afro-Americans	YA	10 weeks	✓	USA
Bogomolov et al. [76]	2014	Stress	P	111	-University (Grad) Students -16 countries	YA	7 months	✓	USA
Rabbi et al. [68] My Behavior	2015	Well being	B	16	-7 Men and 9 Women -4 aged 18-29, 6 aged 30-39, 3 aged 40-49 and 3 aged >50 -13 experienced in Food Diary -11 experienced in Exercise Diary	All	14 weeks	✓	USA
Canzian et al. [77] MoodTraces	2015	Depression	P	28	-University population -15 Men and 13 Women -7 students, 3 PhD students, and 7 researchers/lecturers	All	10 months	✓	UK
Farhan et al. [69] LifeRhythm	2016	Depression	P	79	-University students -Aged 18-25 -26.1% Men and 73.9% Women -24.6% Asian, 5.8% African American, 62.3% Caucasian, 5.8% multiracial	YA	4 months	✓ ⁵	USA
Abdullah et al. [78] CognitiveRhythms	2016	Alertness	P	20	-University Students -7 Men and 13 Women -Aged 18-29	YA	40 days	✓	USA

Continued on next page

²Physical Activity, Sleep, Sociability

³Happiness

⁴Stress, Sleep, Academic Performance, Sociability

⁵A session with a clinician when user is depressed

TABLE 1
Continued..

Reference	Year	Domain	Pillar	Study Details						
				Sample Size	Sample Diversity	Target Audience	Study Length	Feedback	Experts/Theory	Region
Huang et al. [79]	2016	Anxiety	P	18	-University students -Undergraduates in Psychology	YA	10 days		✓	USA
Murmane et al. [80]	2016	Alertness ⁶	P,B	20	-University students -7 Men and 13 Women -Aged 18-29	YA	40 days		✓	USA
Seto et al. [81] CalFit Chi and Dong	2016	Eating Behavior	B	12	-University students -32.3% Men and 66.7% Women -Ages 18-31 -2 underweight, 2 overweight and 8 otherwise	YA	2 weeks		✓	China
Bae et al. [82]	2017	Alcohol Usage	B	38	-21 Emergency Department (ED) patients and 17 Others -Aged 21-28 -50% Men and 50% Women	YA	28 days		✓	USA
Rodriguez et al. [67]	2017	Mood	P	18000	-Full time employment and University students -Majority born between 1970 and 2000	All	3 years	✓	✓	Worldwide
Wang et al. [7]	2018	Depression	P	83	-University students -Undergraduate CS students -40 Men and 43 Women -26 Asian, 5 African American, 24 Caucasian, 1 multiracial, and 26 not specified	YA	18 weeks		✓	USA
Biel et al. [83] Bites n' Bits	2018	Eating Behavior	B	122	-University students -Aged 18-26	YA	10 days		✓	Switzerland
Santani et al. [84] DrinkSense	2018	Alcohol Usage	B	160	-Aged 18-25 -From Two Cities -85 Men and 74 Women	YA	3 months		✓	Switzerland
Boukhechba et al. [85] DemonicSalmon	2018	Depression Anxiety	P	72	-University students -35% Men and 37% Women -Aged 18-23 -37.5% Asian, 4.17% African American, 41.67% Caucasian, 4.17% Latin, and 12.5% not specified/multi racial	YA	2 weeks		✓	USA
Gong et al. [86]	2019	Anxiety	P	52	-University students -Undergraduates in Psychology -32% Men and 68% Women	YA	2 weeks		✓	USA
Sagbas et al. [87]	2020	Stress	P	46	-University students -Aged 18-39 -35 Men and 11 Women	All	Tasks			Turkey
Obuchi et al. [88]	2020	Brain Functional Connectivity	P	105	-University students -Avg. Age 18.2 -35 Men and 11 Women -56% Caucasian, 21% Asian, 16% multiracial, 3% African, 1% Hispanic	YA	79 days		✓	USA

⁶Human Body Clock

III. SUMMARY OF STUDIES

We analyzed 26 studies, and summarized the study details under different areas as seen in Table 1. All the studies have been arranged in chronological order to understand the evolution of the body of literature. The domain column represents the target area of analysis in each of the studies. We have segregated the domain of each individual study based on the three pillar categorization under the column “pillar”. However, in all the studies, even though the target variable belonged to one specific pillar, data belonging to other pillars have been collected using passive sensing and self-reports in order to find relationships with or to infer the target variable. As demonstrated by the evolution through the last decade, the majority of the studies are around the person pillar [64], [65], [69], [71], [73]–[78]. This could be because many of the real-world issues faced by young adults such as stress, depression, and emotional instability are in fact socio-psychological. Before 2016, there were studies that considered both person and behavior pillars in terms of overall well-being of young adults [66], [68], [70], [72], and lately, there are studies that primarily focus on the behavior pillar [80]–[84], [89] that are in domains such as eating behavior, alcohol consumption, and circadian behavior.

Sample size is a key design choice in all of the studies that deal with data analysis. In order to make a statistically significant conclusion, any study should have a sufficient sample size, hence why we considered studies with at least ten participants in this review. Considering that the studies discussed here have targeted young adult audiences, and the majority of them used university students as their study cohort [7], [65], [69]–[73], [75]–[81], [83], [85], [86]. This is explainable since most of the studies were conducted within university research labs, and university students are easily accessible in that context.

Depending on the participant recruitment policy (e.g. convenience sampling [71], [77], [79], [84], [86], study administered participant recruitment [68], [78], [82], [90], in-the-wild recruitment [67], [91]) employed by the researchers, studies have different sample sizes as depicted in Table 1. Although there is no concrete trend in the sample size, many recent studies emphasized the need for getting bigger populations for the studies [83], [84]. For example, the majority of studies considered in this review, that were published on or after 2017 have included more than 50 participants [7], [83]–[86], [88], while the majority of studies on or before 2017 (14 out of 19) used less than 50 participants. Even though this is not an absolute figure regarding all the work done during this time, given the criteria we used to filter studies, this provides evidence of ubicomp researchers focusing on increasing sample sizes in order to make more and generalized conclusions from their studies.

Sample Diversity is a term we used to look at the constituents of the samples used for the study. The factors we considered include the type of sample (e.g. university students, researchers, patients, etc.), age group, gender, ethnic diversity, and other diversity attributes. *Target Audience* is

the type of end users of the system (e.g.: “All” was used when the study is for general populations, “YA” was used when the beneficiary of the study are specifically young adults). Even though very few details regarding sample diversity have been disclosed in some early studies [71], [72], later studies are available where a multitude of information regarding the sample diversity including demographics have been disclosed [66], [68]–[70], [85], [88]. This evolution is important in interpreting and evaluating results. For example, predicting drinking episodes of young adults has variations depending on the gender according to [84], while gender might be less crucial in identifying eating episodes [81], [83]. Hence, incorporating diversity data from the sample population to various analyses might help to interpret results in a more meaningful and an accurate manner.

The majority of studies done until 2013 [64]–[66], [71]–[74] have not particularly disclosed a specific target audience (although they used young populations in the sample in varying proportions), while a majority of studies done from 2014 onwards have been specific regarding targeting to specific audiences making it clear that their studies are valid for a specific cohort of people. When examining this in depth by analyzing sample diversity and target audiences as a whole, some observations can be made. Of all the studies which have general target audiences (e.g. “All”), some have specifically used diverse people in the sample populations [64], [66]–[68]; some have chosen less diverse populations [72], [74]; and some have not mentioned the attributes of the sample populations [71]. Furthermore, most of the studies were done with populations from a small set of countries namely; USA [66], [68]–[70], [72], [74]–[76], [88], [88], Switzerland [73], [83], [84], China [65], [81], UK [71], [77], and Turkey [87]. LiKamWa et al. [64] carried out the study with people from both China and USA, while [67] is the only study in which data from several geographical regions was used. Further, the only study in which the diversity of the user location was taken into consideration in the analysis is [84], where they studied drinking behavior of young adults, observing distinct behavioral patterns in two cities. It is clear that, except for [67], most of the work in this domain still lacks diversity in user populations to provide diversity-related behavioral insights in terms of geographical location. Drawing insights from these findings, a discussion regarding diversity-aware mobile computing research is presented in Section VII-A.

On a positive note, all of the studies analyzed in this review have relied on either the help from domain experts [66], [67], [70], [71], [73], [81], [83], [84], [88] or proven theories [64], [68], [72], [79], [87] in passive sensing, pre and post study interviews, surveys, and focus groups. The use of behavioral science experts and theory into mobile computing research is encouraging in terms of the ability of smartphone sensing frameworks in carrying out procedures that were previously done by manual form filling, interviews, etc.

Table 2 shows studies categorized based on the pillar and domain. P is the pillar that had the highest number of studies with domains such as emotion, mood, stress, de-

TABLE 2: Study Goals / Use Cases

Pillar	Domain	Ref.	Goal / Use Case
P	Emotion	[71]	Infer five emotions (happy, sad, neutral, etc.)
		[75]	Infer happiness
	Mood	[65]	Infer three moods (displeasure, tired, tensify)
		[64]	Infer two moods (pleasure, activeness)
		[67]	Infer two moods (valence, arousal)
	Stress	[73]	Infer stressed or not-stressed
		[74]	Infer high or low stress
		[76]	Infer high or low stress
		[87]	Infer stressed or calm
	Depression	[77]	Infer depressive state
[69]		Infer clinical depression	
[7]		Infer depressive state	
Depression & Anxiety	[85]	Analyze depressive state & anxiety	
Anxiety	[79]	Infer SIAS score for anxiety	
	[86]	Analyze anxiety levels	
Brain Functional Connectivity	[88]	Infer higher or lower vmPFC-amygdala RSFC group	
Alertness	[89]	Infer alertness	
P, B	Alertness	[80]	Analyze alertness and sleep (circadian rhythms)
	Well-being (stress, sleep, academic work)	[70]	Analyze relationship between stress, sleep, and academic performance
	Well-being (sleep, mood, sociability)	[72]	Analyze relationships between sleep and (a) mood and (b) sociability
B	Well-being (sleep, activity, sociability)	[66]	Analyze short term behavior change (sleep, activity, social interactions)
	Well-being (eating, activity)	[68]	Analyze behavior change (eating, activity) using a recommendation engine
	Eating Behavior	[81]	Analyze eating behavior
		[83]	Infer meal or snack eating events
Alcohol Usage	[82]	Infer drinking episodes (not drinking, drinking, heavy drinking)	
	[84]	Infer drinking nights (drinking, not drinking)	

pression, anxiety, brain functional connectivity and alertness. Studies included under pillar B focused on aspects such as general well-being, eating behavior, and alcohol usage. Furthermore, there are certain studies that only conducted an analysis around the focus domain and sensing data, while many studies attempted an inference task using sensing data to unobtrusively detect/characterize health and well-being related aspects. In addition, there is a wide array of well-being domains covered in this review with a more or less equal distribution of papers for each domain. To add to that, the set of studies represent well-being domains that are typically associated to young adults. Moreover, none of the studies focused on an attribute that primarily belonged to the pillar C, because in reality, C pillar does not indicate well-being attributes, but the context that affects such well-being attributes. In Section VII-I, we discuss in more depth about the use cases discussed here and other topics related to young adults that could be explored by leveraging smartphone sensing in the future.

IV. PASSIVE SENSING

In this section, we analyze the types of passive sensing in detail from the perspective of three data pillars. Table 3 depicts which passive sensing modality is used in each study, and

to which pillar of data each sensor belonged in the context of the study. In each study, each passive sensing modality can be used as a proxy to either one or more of the data pillars. As an example, in the study by Biel et al. [83], the target variable for inference was related to the eating behavior of people (inference of meal vs. snack eating events), hence it belonged to the behavior pillar. However, to infer the behavioral trait, the study used other sensors that were captured as proxies to other data pillars (e.g.: location that belongs to context pillar, time that belongs to context pillar, sociability that belongs to person pillar, etc.). As another example, Bae et al. [82] inferred the attribute alcohol consumption (behavior pillar) using accelerometer, proximity, and gyroscope sensors that were used as proxies for behavioral aspects; and location, ambient light, bluetooth, and wifi were solely as proxies for contextual aspects. Hence, under passive sensing, we discuss the pillar of data associated for each sensor data type in the context of the study.

A. CONTINUOUS SENSING

Continuous sensing does not require user interactions or system events to generate data. Typically, these sensing modalities use sensors embedded in the smartphone, and provide continuous streams of data while sampling rates and frequency of logging are determined by researchers depending on study requirements and performance considerations.

1) Accelerometer

Because of habits and behaviors of young adults, they often keep smartphones with them in the pocket, bag or in hand [92]. This fact is leveraged when using accelerometer as a sensor for continuous sensing. The majority of studies discussed in this review used accelerometers to derive activity levels of users, and we observed 3 main usage types: (1) The study used raw accelerometer data to obtain various statistics (e.g. mean, median, standard deviation for sensor traces from all three axes) for different time windows [83]–[85]; (2) The study used external APIs such as Google Activity Recognition API [93] to infer activities from accelerometer data [7], [69]; and (3) The study used activity recognition algorithms developed by researchers [68], [70]. In these studies, the activity inference is done on the phone, making it suitable for real-time smartphone apps that use activity inferences to provide user feedback [68].

From the perspective of data pillars, accelerometer data has often been used as a direct proxy to behavior pillar [7], [65]–[70], [74], [81], [82], [84]–[86]. This distinction is mainly due to the way in which accelerometer data is processed to generate features. In many studies, accelerometer data (using it as a proxy regarding behavior pillar) has been used together with other data to infer person traits such as stress and emotions. In some studies, they are used together with other sensors to infer other behavioral patterns such as eating or drinking behavior.

TABLE 3: Passive sensing in smartphone studies for well-being of young adults. Table depicts the pillar of data each sensor was used as a proxy for.

Pillars of Data	Continuous Sensing	Interaction Sensing
Behavior	Accelerometer: [7], [65]–[70], [74], [74], [81], [82], [84]–[88], Proximity Sensor: [82], Location: [67], [70], [74], [79], [85], [86], [88], Ambient Light: [70], Audio: [66], [70], [71], [73], [88], Gyroscope: [66], [82], [87], Bluetooth: NA , WiFi: [67], Other: [7], [73], [83]	Phone Calls: [64], [65], [67], [74]–[76], [82], [85], [86], Messages: [64], [65], [67], [74]–[76], [82], [85], [86], Email: [64], App Usage: [64], [80], [82], [84], Browsing History [64], Calendar: NA , Typing Events: [82], [87], Touch Events: [87], Lock/Unlock or Screen On/Off Events: [7], [74], [78], [82], [84], [88], Push Notifications: NA , Battery Events: [66], [84], Other: NA
Person	Accelerometer: NA , Proximity Sensor: NA , Location: NA , Ambient Light: NA , Audio: NA , Gyroscope: NA , Bluetooth: [75], WiFi: NA	Phone Calls: [75], [85], [86], Messages: [75], [85], [86], Email: NA , App Usage: [64], [80], Browsing History: NA , Calendar: NA , Typing Events: NA , Touch Events: NA , Lock/Unlock or Screen On/Off Events: NA , Push Notifications: NA , Battery Events: NA , Other: [73]
Context	Accelerometer: [69], [82], Proximity Sensor: [82], Location: [7], [64], [65], [68]–[70], [74], [77], [79], [81]–[86], [88], Ambient Light: [7], [82], Audio: [7], [66], [67], [70], [71], [88], Gyroscope: [82], Bluetooth: [70], [72], [75], [76], [82], [84], WiFi: [82], [84]	Phone Calls: NA , Messages: [65], Email: NA , App Usage: [65], Browsing History: NA , Calendar: NA , Typing Events: [82], Touch Events: NA , Lock/Unlock or Screen On/Off Events: NA , Push Notifications: NA , Battery Events: [82], Other: [73]

2) Proximity

This sensor measures whether anything is close to the screen of the phone, while making sure that the screen is turned off if necessary to save battery, and to avoid accidental touches to the screen. Proximity sensor has not been used often in the literature for determining activity levels of users or in that matter, for any other use-case. However, Bae et al. [82] used it to identify whether the phone is in the pocket or not, hence, obtaining data with regard to the context and the behavior. They used proximity data together with other information such as app usage (behavior) and battery events (context) to generate features related to the overall device activity of the individual (behavior). There is ample opportunity to explore how this sensor can be used in creative ways to use it as a proxy to behavior (e.g.: calling using earphones or calling from the phone, how close the phone is to the face of the user might have details regarding the eye conditions of users etc.) or person pillars.

3) Location

Location is the most common sensor found in literature with regard to continuous sensing. Even though in some cases, location has been used as a proxy to sociability (person) [71], it has been primarily used as a proxy for context or behavior pillars of data as depicted in Table 3. One of the most common uses of location is to derive location-based behavioral patterns of users. Techniques such as Latent Dirichlet Allocation (LDA) and Hierarchical Dirichlet Process (HDP) have been commonly used for such cases [94]. Further, there are a number of features that can be derived from location such as distance traveled [77], [82], time at home and the university [82], points of interest [79], [82], entropy [82], radius of gyration [77], [82], number of different places visited [77], etc. which makes this sensing modality versatile

in terms of the features that can be generated. There are three main techniques to determine the location in smartphones: (1) GPS Sensor [70], [77], [83] - This sensor has been used in many studies as the primary source of location sensing. Even though the accuracy of determining the location is high, it is not power-efficient; (2) Phone Signals [67], [69], [71] - Phone signal information also carries location information. Even though this technique is far less battery consuming, the accuracy of determining the location is low; and (3) WiFi Access Points [67], [82], [84] - It is possible to determine the location (moderately higher accuracy compared to phone signals) using WiFi access points. However, it is required that the WiFi connection of the phone is kept turned on. This kind of location sensing is crucial for indoor positioning and localization too.

According to the Table 3, location has not been directly used as a proxy to person attributes in smartphone sensing literature, even though it has been used (as a data source belonging to context pillar of data) together with other data types. It is known that location can affect the mood and stress levels of people [95]. Considering semantic labels regarding the location of users (e.g. in a restaurant, night club, home, school, etc) and associating those semantic labels to stress or emotional variables of young adults would help to use location directly as a proxy to person data. This is another avenue in which future research can be conducted (e.g. inferring whether a person would consume alcohol in the evening, considering stress levels and other factors related to locations he has been to during the day).

4) Ambient Light

This sensor is used in the smartphone for the purpose of adjusting screen brightness depending on the surrounding lighting condition. Hence, it allows measuring the lighting

condition the user is in. Only four out of 26 studies used this sensor in their analyses. In [65], [82], it was used to measure well-lit environments or the position of the phone (e.g. in the bag, pocket, on the table, etc.), hence providing data belonging to the context pillar, while in [7], [66], it was used in determining sleeping patterns, providing data corresponding to the behavior pillar. While there are numerous studies [96], [97] that discuss the effect of ambient light conditions in stress, fatigue, and mental well-being, this connection has not been extensively studied in smartphone sensing literature on the well-being of young adults. It is a worthwhile problem to look into because many young adult students tend to be awake at night for longer time periods, and this might have a considerable effect on their mental well-being.

5) Audio

Audio can be passively recorded given the consent of the users. This sensor is used to extract features that can be used for multiple purposes: human voice [71], [73], background noise and the sound of the environment [67], and conversations [70]. Hence, studies used audio sensor (mic) to capture for varying purposes including inferring mental well-being, stress, and sociability (all of which belonging to the person pillar). In most of these studies, what audio directly measures is the contextual data regarding the users (e.g. background noise, social context), and they use these contextual information to infer the target variable belonging to the person pillar. Rachuri et al. [71] used passively sensed audio as a proxy regarding the context and behavior, and used the data to infer emotions from speech, while in [67], they used background noise as a proxy regarding the environmental conditions (context). Further, *StressSense* [73] only used audio from the microphone to infer stress, and specifically, they considered how people speak to measure stress levels of university students who are facing job interviews. [66], [67] further discussed the use of audio for determining sleeping behavior. In *StudentLife* [70], two classifiers were used for human voice and conversations in a single pipeline, with the goal of determining how long a conversation lasts. They used this as a measure of sociability and context of students. In addition, they use audio data in determining sleep patterns too. An important factor when using audio is the privacy of users. On a positive note, some studies [67], [70], [71] have made sure they followed certain policies such as: (1) never record audios, (2) never send audio to the cloud, and (3) never analyze speech content to make sure that the privacy of users is preserved. For people to trust audio-based smartphone apps, this kind of privacy-preserving techniques are extremely important to be incorporated. This is clearly a current problem with popular devices and systems such as Alexa and Google Home that have been shown to involve people listening to individual audio snippets without user knowledge [98], [99].

6) Gyroscope

Gyroscope is used to measure the orientation of the smartphone, and is used often in mobile games, augmented reality and virtual reality related applications as part of the trio of sensors (accelerometer, gyroscope, and magnetometer) that make up the inertial measurement unit (IMU unit) [100]. It is another sensor that has been rarely used in smartphone sensing research in general, although some activity recognition algorithms use gyroscope data in addition to accelerometer data [93]. When it comes to research targeting the well-being of young adults, we came across only one study which used the capabilities of a gyroscope. Bae et al. [82] used the rotation of the phone derived from the gyroscope in order to study the drinking behavior of young adults. They concluded that movement and location-based features (e.g.: accelerometer mean magnitude, maximum magnitude of rotation, radius of gyration) can contribute to determining drinking episodes. There is plenty of opportunity to test the validity of using Gyroscope sensor data as a proxy to person and context pillars. For example, can smartphone sensing studies determine phone usage behavior of people (e.g.: whether they are using the phone while keeping it on table, on the hand, or while on the bed where gyroscope readings would be significantly different from each other) and leverage these features to infer person and behavior attributes (e.g. are people using smartphone more while on bed more stressed/healthy compared people who do so while being on a table, etc.)?

7) WiFi and Bluetooth

WiFi is the primary way through which most smartphone users access the internet. Even though many smartphone studies reviewed here used WiFi connections for transmitting data to servers [68], [72], only a few studies [7], [65], [67], [82], [84] used WiFi related data for a purpose other than data transmission. WiFi has been used as a mode of location sensing in all of these studies, while [65], [67], [82], [84] used WiFi in a hybrid location-sensing strategy accompanied by GPS and Cellphone signals as explained in Section IV-A3. [70] used WiFi signals to get fine-grained indoor positions using the indoor localization system at Dartmouth College.

Bluetooth technology has been used in smartphone sensing research on young adults, primarily to understand the context of users. Several studies [71], [72], [75], [76], [101] used bluetooth logs to infer physical proximity to others, and to determine co-location (context). Other works such as [84] used features derived from bluetooth scans such as number of records, number of unique bluetooth IDs in range, and the percentage of empty bluetooth scans as general proxies to social context. Moreover, [82] used bluetooth network usage and [70] used bluetooth connections to transmit data from wearable bands used in their study. In this case, bluetooth is only used for connectivity rather than as a behavioral cue by itself.

If we consider continuous sensing techniques as a whole, it is visible that they have been often used to infer person

attributes by taking them as proxies for context or behavior pillar data. Avenues for future research in continuous sensing are discussed in Section VII.

B. INTERACTION SENSING

Interaction sensing does not require a physical sensor to be present, even though it captures data regarding users' interaction with a smartphone. These modalities are based on the software used in the system and events triggered in the smartphone. The informativeness of the sensing modality depends on the way people use the phone. Interaction sensing modalities generate data based on events that are triggered by the smartphone user. When an event occurs, the data can be logged to data repositories. e.g.: phone calls, messages, app usage, browsing history, calendar, typing events, touch events, lock/unlock events. These data have been often used as direct proxies to either *behavior* or *person* attributes.

1) Phone Calls and Messages

Several studies have put an effort to leverage call events and message events as virtual sensing modalities for person attributes of young adults. Most of the research that uses these two interaction sensing techniques [64], [67], [74]–[76] are related to mental well-being, while one study dealt with inferring drinking episodes (behavior) [82]. Some general features derived from these two types of modalities include the number of phone calls, number of messages, duration of phone calls, and length of message [64], [67], [74]–[76], while some studies have delved in depth by analyzing types of features derived from these modalities.

[74]–[76], [82] used incoming and outgoing calls as different features in their studies, while Sano et al. [74] found that fewer SMSs sent out correlated with higher stress levels in young adults. Bae et al. [82] used features such as speed of typing, number of emojis in messages, types of emojis, number of unique conversations and they found positive relationships between features such as number of deletions, number of key-press insertions, average time between key-presses on one hand, and drinking episodes on the other hand. In a study about the anxiety of college students, Gong et al. [86] used accelerometer and location data in and around phone calls or messages, and were able to infer anxiety episodes using these features. Moreover, Murnane et al. [80] used phone calls and SMSs as app usage events instead of isolated call or message events and measured phone app usage time and SMS app usage time instead of the usual event log details. Since most smartphones include phone and SMS features in the form of apps, this strategy could be said to be beneficial, while the concrete added value of these measurements over usual measures such as the number of calls and messages needs further research. There is also the question of whether messages or voice calls are useful features in the current day and age, especially among young adults who frequently use apps such as Whatsapp, Skype, Viber and Facetime for communication purposes. Hence, this needs further research.

2) App Usage and Browsing History

Even though young adults tend to use smartphone apps regularly for different usage scenarios, and a body of research regarding smartphone app usage behaviors in other research domains exist [102]–[105], this aspect has not been used extensively in smartphone sensing related research targeting young adults with few exceptions. We found only 3 studies [64], [80], [84] that directly leverage app usage behavior for their analyses by generating behavioral features from app usage.

There are different types of apps used by people for different purposes, and usage patterns of these apps can portray a picture regarding daily routines (behavior) as per Murnane et al. [80]. In their work, they emphasize the importance of app usage features related to apps running in the foreground instead of the total app running times (including when running in the background), as the former type of features are more informative. Further, they categorize smartphone apps into entertainment, games, communication, browsing, etc. to gain a fine-grained understanding regarding app usage behavior, and circadian rhythms of people. Santani et al. [84] and LiKamWa et al. [64] used similar approaches for their studies in the domains of alcohol usage (behavior) inference and mood (person) inference. In both studies, they consider the most used apps and categorize them into app groups using different criteria. Further, they consider app launch events and app usage duration in their analyses. Moreover, [84] reports that they were able to discriminate drinking episodes with an accuracy of 61% using app usage alone. The most frequently used app in their study was Whatsapp (for messaging and communication), while [80] found out that communication apps are frequently used between 9 a.m. - 9 p.m., whereas social media apps are the top used category before sleeping and after waking up (after 9 p.m. and before 9 a.m.). As depicted in Table 3, app usage has not been used in the context of contextual pillar that much. Some interesting avenues to explore are: (a) Driven by the fact that majority of young adults use mobile fitness applications [106], explore whether smartphone fitness app usage does have correlations with the real physical activities or fitness, hence making it possible to use app behavior as a proxy for behavior pillar; and (b) whether the apps that are used at a particular moment are correlated with the context/location of a user. e.g.: whether people use the same kind of apps while they are traveling, eating, in class, in a bar, etc. High correlations might enable us to directly use highly used app categories as a proxy to the context of users.

Browsing History is another modality that can be derived from smartphones. Even though internet browsers have been used as an application type in certain studies [80], [84], they have not particularly analyzed browsing history-related features in connection to the problem domain in which they have worked on. On the other hand, LiKamWa et al. [64] used unique website domains visited as a feature (behavior) in their model to infer the mood of individuals (person). As depicted in Table 3, there is plenty of opportunities to use

internet browsing as proxies for all three pillars. An interesting area to look forward would be to use a categorization of websites (similar to app categorizations) and to understand person attributes. E.g.: number of visits to websites related to education, knowledge discovery such as StackOverflow, Google Scholar might have correlations with the intellectual capabilities of individuals and mental well-being.

3) Phone Usage Events

This relates to sensing modalities and event triggers related to phone usage such as typing, touching, screen on and off, locking and unlocking the phone, battery events. Sano et al. [74] demonstrated in their study that screen on/off times (behavior) have correlations with stress levels (person) of young adults. They derived a set of features, out of which standard deviation of the percentage of the screen on events between 6 pm - 9 pm demonstrated correlations with stress, and also worked with 75% accuracy in predicting stress events. Further, Abdulla et al. [78] used the screen on/off related features (behavior) to determine the overall alertness of university students (behavior/person). They also used touch events (behavior) to determine alertness more as a mechanism of self-reporting (discussed in Section V) in accordance with a standard testing mechanism known as Psychomotor Vigilance Task (PVT).

Bae et al. [82], who analyzed message typing behavior, also paid attention to other typing related dynamics. They showed that there is a correlation between features related to general and keypress typing (behavior) with drinking episodes (behavior). They further utilized screen on/off times to determine phone usage behavior and showed correlations between those metrics and drinking episodes. Moreover, they concluded that screen unlocks per minute is likely to be lower when a user is drinking in comparison to non-drinking episodes. They used battery charging event-related metrics in their analysis too. In another study on alcohol drinking episode prediction, Santani et al. [84] used screen on/off related features to determine phone usage. They concluded that the percentage of screen on time (behavior) is negatively correlated to alcohol consumption (behavior). Further, they were able to predict drinking episodes with 61.1% using only phone usage behavior. Wang et al. [7] used phone usage in classrooms (behavior) as a proxy for lack of concentration (person) in university students. The study was built on previous literature regarding smartphone usage and concentration [107], [108], while they specifically measured smartphone usage events in classrooms, dorms, and throughout the day as a whole using lock and unlock events of the phone.

As a whole, in terms of generating features belonging to pillars of data, interaction sensing modalities have been extensively used to generate features as proxies to person and behavior attributes. As some social science research shows [109], there are connections between phone usage, physical activity, and sedentary behavior in young adults. Further research in the smartphone sensing domain is required to validate these claims, and to directly generate features that

can be used as proxies for attributes belonging to behavior pillar.

V. SELF-REPORTS

In this section, we discuss regarding data captured via self-reports used in smartphone sensing studies for the well-being of young adults as given in Table 4. While passive sensing data are used as proxies for traits of different pillars, self-report data are often used as ground truth events [110], and they too can belong to either one or more pillars. There are some studies in which ground truth data are obtained manually, before or after the deployment phase (pre/post study), while many studies use the ubiquity of smartphones in obtaining real-time self reports in the form of Ecological Momentary Assessments (EMA) and Questionnaires [67], [83], [84]. In Table 4, we have divided self-reports into two sections: (1) *Type* - What is the type of method used to capture self-report data from the user; and (2) *Trigger Context* - In what kind of a situation/context the self-reporting is captured, and whether users are reminded regarding self reports. Further, we consider whether filled surveys, interviews or focus groups are involved in studies, to collect additional data in any of the three study phases. We analyze all of those traits in terms of the pillars of data.

A. TYPES OF SELF-REPORTS

1) Questionnaires, EMA and Diaries

Structured questionnaires [64], [65], [68], [72], [74], [77] and Ecological Momentary Assessments [111] (EMA) [70], [78], [81] are the most common ways of collecting self reports from users. With these questionnaires, users were given multiple choice questions, and yes or no questions depending on the type of ground truth which needs to be obtained. Further, most studies used standard questionnaires or techniques borrowed from social psychology and health related literature in order to make sure that the questions asked in the self reporting phase are of a solid foundation [65], [69], [70], [72], [74], [77], [79], [80]. Some studies have even used the expertise of personals in respective health, well-being domain in order to create standardized questionnaires [66], [67], [70], [71], [73], [81], [83], [84]. Structured QnA, EMA were commonly used to obtain ground truth related to person pillar, using both in-situ and retrospective approaches. As given in Table 4, questionnaires and EMA are the primary types of self-reports used in the body of literature, and these self-report techniques have been used to capture features regarding all three pillars of data. Rachuri et al. [71] used a diary to explicitly report emotions (person) in their study, and it was mentioned that a group of psychologists and domain experts were involved in the study design, while [80] used a sleep diary for sleep related (behavior) data collection in their study. Moreover, in most studies, self-reports were used to obtain data regarding contextual aspects, as many nitty gritty details regarding the context can not be captured via techniques available in smartphones currently. For example, Santani et al. [84] and Biel et al. [83] in their behavioral stud-

TABLE 4: Self-Report Data in Smartphone Sensing Studies for Well-being of Young Adults that Belong to Different Pillars of Data.

Pillars of Data	Type	Trigger Context	Pre and Post Deployment
Behavior	Structured QnA: [66], [68], [70]–[72], [74], [78], [82]–[84], Pictures from Camera: [68], [83], Videos from Camera: NA , Audio: NA , Diary: [66], [80], Other: [87]	In-situ: [66], [68], [78], [83], [84], [87], Retrospective: [68], [70]–[72], [80], [82]–[84], Self-Initiated: [66], [71], [72], Reminders: [70], [80], [82]–[84]	Interviews/Focus Groups: [64], [80], [82], [85], Filling Survey: [67], [74], Other: NA
Person	Structured QnA: [7], [65], [68]–[72], [74]–[78], [81], [85], Pictures from Camera: NA , Audio: NA , Diary: NA , Other: [64], [67], [68], [70], [78], [80]	In-situ: [7], [64], [65], [67], [70], [78], [80], [81], [85], Retrospective: [70]–[72], [75]–[77], [81], [85], Self-Initiated: [65], [67], [70]–[72], [75], [76], [80], Reminders: [64], [69], [77], [81], [85]	Interviews/Focus Groups [64], [69], [80], Filling Survey [7], [67], [70], [74]–[76], [79], [86], Other: [88]
Context	Structured QnA: [70], [71], [74], [81]–[84], Pictures from Camera: [68], [83], Videos from Camera: [81], Audio: [81], Diary: [66], Other: NA	In-situ: [66], [81], [83], [84], Retrospective: [71], [81]–[84], Self-Initiated: [66], [71], [83], Reminders: [82], [84]	Interviews/Focus Groups: NA , Filling Survey: [67], Other: NA

ies regarding drinking and eating respectively, collected data regarding social context (context) and concurrent activities (behavior) that are done while drinking/eating.

2) Photo, Audio and Video

These modalities are traditionally discussed in multimedia research, and is mainly focused on the content of the multimedia. With the emergence of ubiquitous smartphones that have become primary cameras for many individuals worldwide [112], it has become an important area of discussion in ubiquitous computing as well. While multimedia and computer vision researchers are focused on generating different techniques to obtain features from photos and audio, ubiquitous computing domain has focused on using these contextual and behavioral features for pervasive applications to sense humans. Further, it should be noted that audio has been used as a passive sensing modality in many studies as described in Section IV. In this case, we consider instances where audio data has been reported by users.

In a study regarding the eating behavior of university students, Biel et al. [83] used photographs taken from the smartphone as a way of self reporting. Users were asked to take a picture of the meal/snack they are taking and upload the picture together with the structured questionnaire that they had to fill in. Further, in another eating behavior related study, Seto et al. [81] used videos with voice annotations in order to determine portion sizes and food groups of meals. Some of the considerations [81], [83] mentioned about using photos and videos are the high data usage and the possibility of using crowd-sourcing to annotate videos/photos. Interestingly, even-though audio has been used as a passive sensing modality (see Section IV-A5), we did not find studies where audio has been used for self reporting except for audio annotations of videos in [81].

Another important aspect with regard to photo and video based self-reports is that it is highly likely to be free of human bias compared to answering questions or using EMAs. More-

over, from the perspective of three pillars, photos, videos, and audio were not used extensively as proxies to person attributes. While it is obvious given that the sensing technique uses the smartphone camera, innovative mechanisms might be able to tackle such cases. For example, young adults are obsessed with taking selfies, and leveraging this by asking young adults to upload selfies with their reports might possibly help determine the contextual aspects (e.g. light levels of the room, what is present in the background, etc.) and person aspects (e.g. facial expressions of the users). A more adventurous envisaged application might use facial features such as beard, hair, etc. to generate interesting novel features (e.g. growing beard, cutting hair, emotions, etc.) that could be useful in generating fine-grained behavioral or socio-psychological (person) patterns over time.

B. TRIGGER CONTEXT OF SELF-REPORTS

Trigger context is the context at which the smartphone users initiate the self reporting procedure, and the context of the initiation it self. We have segregated it into 4 main sections as In-situ, Retrospective, Self-Initiated, and Reminders. It should be noted that some studies have not disclosed specific details regarding certain trigger contexts. Hence, we gave our best effort to categorize the studies from the details they have provided.

1) In-situ

In-situ experience sampling is where the smartphone users report details regarding the current situation they are in [64], [70], [78]. This is beneficial for understanding the user behavior, and to obtain accurate ground truth data. Some example questions that can be asked are "How do you feel now?" and "What are you eating now?". LiKamWa et al. [64] used in-situ self-reports to obtain the mood of young adults (person) with an interface they created based on Circumplex Mood Model and Positive and Negative Affect Scale (PANAS). They designed the study such that users reported

their current mood four times a day, each report three hours apart. Wang et al. [70] used EMAs several times per day to obtain details regarding depression and stress (both under person pillar) of university students. In another study regarding alertness of people (behavior/person), Abdulla et al. [78] used Psychomotor Vigilant Task (PVT) several times per day to obtain the in-situ alertness of smartphone users. In another study regarding the eating behavior of university students (behavior), Biel et al. [83] used in-situ reporting of meals and snacks with images taken from the phone which they used to obtain portion sizes, nutritional values and also to enforce compliance. Wang et al. [70] used in-situ self reporting in an interesting manner where they changed the frequency of self reports gathered from users depending on the time period. For example, they mentioned that they used more frequent self reports to obtain more stress related details during time periods closer to assignment deadlines. Hence, in-situ self reporting has been used reasonably well to capture data regarding the person pillar (e.g.: mood, emotion, feelings, and alertness) and behavior pillar (e.g.: eating, drinking, and sleeping habits), hence avoiding the recall bias, because data are captured then and there. Moreover, there is a considerable increase in the use of in-situ experience sampling techniques over time in the set of studies within our scope.

2) Retrospective

Retrospective trigger context considers situations where the self-reporting is done retrospectively, and not regarding the current situation. Some examples for questions of this nature are "How was your mood in the morning today?", "What did you eat for lunch", and "How was your sleep last night?". This type of self-reports are common in most studies according to our findings. Bae et al. [82] used retrospective self-reporting with high completion rates in their study on alcohol usage patterns of young adults. They also reported from one of their previous studies that using hourly EMAs to obtain alcohol consumption had low reporting rates. Some studies also used end of the day retrospective surveys to obtain details from users regarding certain details about the whole day [72], [74], [85]. Further, [83], [84] used forgotten drinks surveys and forgotten meal/snack surveys in order to obtain retrospective details from university students in cases where they have forgotten to complete the in-situ reporting. While retrospective techniques might be convenient for smartphone users because they are not disturbed throughout the day, it does not allow researchers to collect real-time data. Another drawback of this technique is that it might lead to recall bias.

3) Self-Initiated

Self-initiated trigger context refers to instances where users are supposed to initiate the self report them selves without any reminders. Self-Initiated survey can be both *in-situ* or *retrospective*. In Table 4, there are few main types of studies which were categorized under *Self-Initiated*. Namely, studies that use self reporting mechanisms and have not mentioned details regarding user reminders/notifications [71]; studies

TABLE 5: Studies from the System Perspective

Analytical Systems [71], [72], [73], [74], [75], [70], [76], [77], [69], [78], [79], [80], [81], [82], [7], [83], [84], [85], [86], [88], [87]
Feedback Systems [66], [65], [64], [68], [67]

that used both reminders and self-initiated reporting options [64], [67], [76], [84]; studies that used notification mechanisms when self-initiated reporting is not done. [64], [67]; and studies that require self-reports only when an event occurs (e.g.: eating, drinking etc) [83], [84].

4) Reminders

Reminders, or in the context of smartphones, *Push Notifications* are used in studies to remind young adults regarding self-reports. Reminders can be useful in increasing the self-reporting rates and compliance [64], [84]. There are studies in which notifications have been sent several times per day in order to collect in-situ EMAs [64] while in some studies, phone notifications have been used to remind users regarding retrospective surveys. Phone notifications might help users in some cases when it is difficult for them to keep track of the self-reports, and keep completing them. On the other hand, too many notifications might be disturbing for users and this aspect should be considered when choosing notification strategies for real-world deployments [113]. However, studies regarding young adults [114] have shown that young adults respond to location based push notifications in a positive manner. Leveraging this behavior of young adults could benefit smartphone sensing studies, specially since push notifications have not been discussed thoroughly as a proxy to any of the three pillars as given in Table 3. With preliminary finding suggesting the effect of push notifications in the productivity and stress (person) of people [115], [116], [117], [118], it is worth being explored more on how people respond to push notifications, and also how push notifications affect the behavioral and socio-psychological (person) state of individuals.

VI. SYSTEM PERSPECTIVE

In social science and clinical experiments where researchers try to reduce bias imposed on study participants, they use various techniques to make sure that they do not inflict behavioral change (e.g. in randomized controlled trials, a group does not know exactly what they are tested for). This ensures that researchers are able collect unbiased and accurate data from human subjects regarding their normal behavior. When considering smartphone sensing studies by the essence of self-monitoring, we are able to segregate all the studies into two types: (1) Analytical Systems and (2) Feedback Systems.

A. ANALYTICAL SYSTEMS

Analytical Systems run primarily for the purpose of data collection while no (or less) analysis, training or inference is done during the deployment phase. Another common feature of these systems is that they do not provide any feedback to the user regarding health, well-being, or behavioral routines, hence not assisting with self-monitoring. All the data collected during the study, are then used for offline analysis when the study is finished. Majority of studies (except for the once mentioned in Section VI-B) which target the well-being of young adults, fall under this system perspective. Some examples for this type of systems are studies by Santani et al. [84], Biel et al. [83], Wang et al. [7] etc. Further, there is a visible trend in studies in-terms of decreasing number of FSys and increasing number of ASys.

The goal of analytical systems is to test for a specific hypothesis. For example, Wang et al. [70] studied the behavior of students in Dartmouth College using an ASys. During the semester, there were no app based or physical interference with students, and the application was primarily used for data collection, and their objective was to find out the relationship between behavior, emotions, and GPA during the studies. Similar traits can be seen in [72], [74], [83], [84] too. Drawing motivation from self-monitoring related theories, in order to test the hypothesis in an unbiased manner, the participants of the smartphone sensing study should not have been aware of the hypothesis the researchers are testing for and they should not be allowed to change their behavior based on self-monitoring. Sagbas et al. [87] explicitly mention this in their study, pointing out that they specifically did not let the participants know that they are conducting the study focusing on stress. As another example, if participants in [70] knew that the experiment would be used to find connections with their GPA, there is a tendency that they would intentionally alter their behavior to obtain higher GPAs, hence having the effect of an interference. In [84], if students knew that researchers are testing for exactly alcohol drinking behavior, they might feel as if they are drinking less or more alcohol, hence alter the drinking behavior. Further, unlike research methodologies such as action research [119] where researchers engage with test participants to discuss and evaluate research techniques, analytical systems are better off being tested without self-monitoring capabilities (no feedback, interactions or interventions with participants), hence making sure that participants are sensed in their natural environments without any significant behavioral change due the usage of the mobile application.

While it is understood that the ultimate goal of smartphone sensing studies targeting the well-being of young adults is to infer well-being conditions just using passive sensing techniques, in order to get there, self-reports are still required. Hence, it involves smartphone users explicitly reporting conditions such as their emotions, stress levels, drink consumption, food consumption etc. When reporting such data for several days, it might bias the users leading to changes in the normal behavior of users. E.g. a person who reports drinking

behavior might have reported that they did not drink alcohol in 2 weeks. When they realize this because they focus on it on a daily basis, it might motivate them to drink alcohol. Therefore, the real challenge in conducting smartphone sensing studies with analytical systems is that it is very difficult to reduce effects of self-monitoring during the study period. Current body of research does not discuss regarding this aspects in detail, and for smartphone sensing studies regarding well-being to grow in maturity for real world utility, aspects regarding self-monitoring should be thoroughly taken into consideration, and necessary principles should be adopted during participant recruitment phase and deployment phase to make sure that the biases are reduced and interventions are minimal during the study period. This would require novel recruitment policies for smartphone sensing studies that deal with young adults, special attention to reduce self-monitoring by carefully preparing the smartphone application, and its' content (e.g.: not revealing/emphasizing the final goals of a research, not showing statistics that might bias the behavior) and many more innovative techniques by considering the mindset of young adults and their behavior.

B. FEEDBACK SYSTEMS

All the studies considered under this section provided feedback to users, or had behavioral intervention strategies included in the study. In implementation, studies used some sort of data analysis and visual representations that are visible to the users, making the system more of a real-time tool rather than a tool that is just used for data collection. Hence, these systems have allowed self-monitoring to users, making sure that the system is monitoring the progress or the reaction of the users when reflecting upon the feedback they get through the application. Hence, the goal of this kind of systems would be to evaluate the effectiveness of feedback, and the engagement and behavioral change from the perspective of users. It should be understood that when providing users with feedback, they reflect upon it, and alter their behavior more often than not. That is why the system perspective is important to be considered when designing smartphone sensing systems, specially for young adults.

LiKamWa et al. [64] carried out a study regarding moods of young adults. Their study included components such as sharing user mood with others, viewing mood history, and they used a star achievement system to motivate users to self report more often, allowing them to collect a high amount of accurate ground truth data. It should be noted that even though this is not an intervention to the participants, this self-monitoring capability might cause behavioral changes in participants. Ma et al. [65] too provided users with real-time feedback summarizing some information such as distance traveled, SMS frequency, most lasting activity, and general mood. Again, this feedback in the app is not necessarily an intervention, but allows users to self-monitor their behavior. Lin et al. [66], in their study *BeWell+*, deployed a comprehensive real-time system, which involved in-device activity inference while community adapted well-being scores

were computed in servers to provide user feedback. Further, they used a multi-dimensional ambient display with a fish to exhibit how well users are doing in-terms of the well-being score. They mentioned that their system received good feedback from 70% of users at the end of the study. Their off-the-shelf analysis provided further insights saying that users of the system improved their well-being scores while using the app (positive behavioral change), elucidating the benefit of allowing the users to self-monitor. This is an example of how self-monitoring is activated in order to evaluate the usefulness of an end-to-end feedback system that facilitate behavioral change.

Rabbi et al. [68] have done a study regarding the well-being of a group of people including 10 young adults. They analyzed food, exercise, and activity patterns of users to provide user adapted well-being suggestions. They also analyze how smartphone users make use of these feedback to enhance their well-being. Rodriguez et al. [67], in their large scale study regarding smartphone user emotions, provided a simple feedback regarding emotional states of the users. Further, they introduced gamified elements to their feedback where app users were able to receive additional feedback by providing more self report data. This study is a proof that giving comprehensive feedback to obtain more user data in real world, in-the-wild deployments can be successful. Farhan et al. [69] have used real-time self reported data from their smartphone app to make clinical interventions, where they make sure that university students who have higher stress levels are interviewed by clinicians for them to continue to be part of the study.

Smartphone sensing applications are supposed to run unobtrusively as background apps, and it is important that these applications are optimized in-terms of performance for a better user experience. Fewer studies in our scope have done a thorough system analysis with the deployed smartphone application. LiKamWa et al. [64], in their study regarding moods of young adults, have done an extensive analysis regarding their app "MoodScope" in smartphones such as Apple iPhone 4, and Samsung Galaxy S2. For each phone, they computed latencies, computation times, power consumption, data consumption associated with background logging of data, in-phone data pre-processing and inference, and communication with servers. The only other study which has done a performance analysis of system aspects was by Rachuri et al. [71]. In this study regarding emotion recognition with smartphone sensing, they have considered latency, and energy consumption of their system with regard to speaker recognition, and emotion recognition tasks. Even though certain studies [65]–[67], [82] have discussed some performance aspects with regard to accelerometer sensing, location sensing, and battery life, they have not performed extensive analyses regarding these system aspects. As a summary, it could be said that user experience is an important component in feedback systems, because social science research too suggest that young adults have a taste towards interactive, gamified, less buggy, and feedback providing

smartphone applications [120], [121].

VII. DISCUSSION

A. DIVERSITY-AWARE RESEARCH IN UBIQUITOUS AND MOBILE COMPUTING

Diversity in Machine Learning [122] is an important topic that has grown in popularity during the last few years. While traditional machine learning focuses on data, model, and inference; diversity-aware machine learning has components such as data diversification (e.g. age, sex, country, culture, race, etc.), model diversification, and inference diversification. Specially in computer vision domain, these issues have been discussed in depth where some examples are biases in face datasets [123] and facial recognition systems [124], and studies finding that self-driving cars are more likely to hit Black-American people [125], [126]. Considering studies within our scope, we discussed some diversity aspects in Section III. We now identify a set of key limitations and areas that require future work.

1) Lack of diversity in a sample population

Using a population of diverse groups where each group is represented by few individuals is problematic. This might lead to models that poorly learn and lead to wrong conclusions. For example, in [66], a study to understand well-being had a small population of 27 people with 9% from CS department, 34% doctors/medical researchers, and 57% graduate students. If we consider these sample cohorts, the behavioral routines of undergraduates, doctors, medical researchers and graduate students could be different. Hence, the resulting analysis about well-being might differ for different groups. The question would be whether the models could have actually considered these diversity aspects of people. Another issue is when the sample sets are not diverse enough. For example, stress levels of students could be different based on the classes they take, or many other factors. In [79], the authors recruited all the students participating in the study from a psychology class, while the conclusions are made as if they were valid to all university students. Whether this is a fair conclusion remains a question.

2) Gender-Biased training data

Some studies contain rather gender-biased sample populations. As pointed out by Santani et al. [84] and also by a body of specialized literature, diversity aspects such as gender of young adults could play a role in determining their drinking patterns. Yet, some research has used heavily biased gender ratios. For example, [74] had 18 young people out of which 15 were men. [70] used a dataset of 48 people out of which 38 were men. From this perspective, and as gender biases in machine learning systems become more evident [127], having a more in-depth analysis regarding the results with gender diversity in mind would be very informative and conclusive. A body of literature, both classic [128] and recent [129], [130], could guide how researchers can pursue directions related to gender. Mobile sensing research, especially in

domains of mental health should concretely look into these gender biases.

3) Geographical Diversity

Smartphone usage behavior of young adults across countries can differ [102], [129], [131], ranging from the type of apps they use to the time of phone consumption due to plethora of reasons such as cost of phones, unique lifestyles, and culture. For example, for young adults in western countries, Friday night would be a relaxing day where they drink and party, the situation could be totally different in Asian countries such as India where drinking is not socially accepted. This kind of geographical diversity has not been thoroughly studied in the literature so far, and it could be mainly because most studies were done in one or two countries with limited young adult participation. As EmotionSense [71] and Sea Hero Quest [91] have demonstrated, with wide smartphone and internet adoption [132], changes in the attitudes of people regarding using smartphones, and availability of app-based ecosystems, now it has become possible to conduct large-scale smartphone sensing for wider audiences. This also makes it possible to conduct smartphone sensing studies for young adults, specifically considering geographical diversity.

4) Personalization vs. Diversity-Awareness

Another important aspect regarding diversity-awareness is its distinction from personalization. While personalization focuses on building different models for different individuals from personal data, diversity-aware models consider diversity-aware features in building a single machine learning model [122]. While personal models allow learning details regarding a single user, diversity-aware machine learning models allow learning hidden patterns/routines regarding diverse user groups. Moreover, the goal of diversity-aware learning is preventing or at least mitigating racial, age, gender, and other biases in artificial systems while also leveraging this diversity to build better models that benefit communities in different parts of the world.

Unlike computer vision datasets, smartphone sensing datasets in academic research so far have been comparatively smaller. Hence, applying diversity-aware machine learning techniques to these datasets can be challenging, often calling for specific techniques for oversampling or data augmentation [133]. However, as smartphone usage continues to grow among young adults, we believe that there is an opportunity to conduct studies, across many geographical regions, with diverse young populations. Hence, leveraging these technological developments to design smartphone sensing studies that integrate human diversity is timely. Developing new diversity-aware learning and inference techniques for heterogeneous smartphone sensing data could be the key to applications and systems that are both innovative and relevant to society. In summary, two RQs from the above analysis are: **RQ1:** How to design mobile health sensing studies with human diversity as a key value, across geographic regions and population groups? What are the techniques available

to measure diversity in datasets using machine-mediated models?

RQ2: How to build diversity-aware machine learning models for mobile health sensing that take into account the diversity aspects of people to provide them utility, but also to not discriminate based on diversity aspects?

B. ROBUST MACHINE LEARNING MODELS AGAINST SENSOR FAILURE IN FEEDBACK SYSTEMS.

Many studies examined in this review mentioned that there are instances of incomplete data because of sensor failures and lack of self-reports [83], [84]. Yet, the machine learning architectures in the studies reviewed here do not attempt to examine inferences under sensor failure situations. This is a very important aspect relevant to mobile sensing studies in the open world. Further, in ASys, this situation will not cause issues because data cleaning and processing is done later to avoid data segments where some data sources are missing. However, for FSys, having a single model that uses all the features generated from the study to perform an inference can be troublesome in cases of sensor failure where analytics, interventions, and insights must be provided in real-time. Moreover, given that different types of features (e.g. app related features, accelerometer features, location features) have proven to perform the same inference with similar accuracies [83], [84], it would be interesting to train different models using different feature sets, and build ensemble architectures that would select the best set of models to perform an inference, given the availability of data. Extending this line of work, another direction would be the use of different ML models to process different feature types in real-time to perform the same inference (e.g. LSTMs for accelerometer data, random forest classifiers for tabular categorical data from self reports, convolutional neural networks to process images, etc.) and combining them as an ensemble. In summary, two RQs from the analysis we just presented are:

RQ3: How to build ensemble machine learning architectures to make feedback systems more robust against sensor failure?

RQ4: How to use different types of ML models in an ensemble architecture, and what are the benefits of this approach in feedback systems?

C. INTERACTION SENSING TO BETTER UNDERSTAND YOUNG ADULTS

According to our analysis, there are three primary advantages of using interaction sensing over continuous sensing. They are: (1) **Sensor Failure and Calibration.** As mentioned in previous sections, continuous sensing techniques primarily rely on embedded sensors on the smartphone. Some of these sensors might lose their accuracy over time unless properly calibrated (e.g. location, accelerometer, gyroscope, etc.) [134], [135] and some sensors fail over certain situations (e.g. when the phone is in airplane mode, when location is turned off, when bluetooth or Wifi is turned off, when microphone access is not given, etc.). However, such pitfalls

are less frequent in interaction sensing (compared to continuous sensing), because interaction sensing requires user permission less often, and thereafter, such permissions are not turned on or off on a continuous basis, hence reducing the probability of failure; (2) **Heterogeneous Smartphones**. Continuous sensing techniques primarily process data from sensors whose sampling rates and accuracies might vary significantly based on phone type, operating system, quality of the sensor, and availability of certain specialized sensors [136]. Hence, this heterogeneous nature poses challenges in the context of well-being applications that would be run on different phones in different contexts. However, interaction sensing focuses on using low-dimensional representations of data such as app usage, screen on/off, battery events, and typing events that are to some extent agnostic to the heterogeneous nature of smartphones, making it possible to build ML models that require fewer adjustments and calibrations to function under different circumstances; and (3) **Young Adults and Smartphone Usage Behavior**. The current work in the domain reflects that priority has been given to continuous sensing techniques as summarized in Table 3. It should be understood that these sensing techniques act in a similar way regardless of the user, as in whether they are young adults or older people, and hence, the sensing data can be obtained with good quality regardless of the age group, as long as the technological infra-structure supports it (e.g.: wifi access, quality of smartphone, etc.). However, when we consider interaction sensing, a key issue is how people use and interact with the smartphone. Prior research [12], [13] suggests that there are significant differences in smartphone and app usage behavior of young adults compared to older generations. Hence, this might lead to signals of different strengths when used on young adults as compared to older generations. We believe that researchers in ubiquitous computing and human-computer interaction should focus on this aspect. Leveraging this prominent characteristic would allow a paradigm shift from continuous sensing to interaction sensing. While we understand the importance of continuous sensing techniques, interaction sensing offers a lot of room for improvement as mentioned above, and a multitude of open research questions in smartphone sensing, especially when targeting young adults and their well-being. In summary, two RQs from the previous discussion are:

RQ5: How to tackle methodological pitfalls of continuous sensing (e.g. sensor failure, heterogeneous sensors) by using different interaction sensing approaches focused on health and well-being conditions?

RQ6: How to improve interaction sensing techniques to better suit different age groups?

D. MAKING BETTER USE OF KNOWLEDGE IN HUMAN SCIENCES

As per prior research we discussed in this study, it is unclear whether computer science researchers are fully making use of principles and techniques from human psychology, clinical, and behavioral research into smartphone sensing

research. Given that smartphone sensing systems capture the behavior of young adults, it is essential to understand behavioral dynamics and psychology of this age group. When looking at this problem from the system perspective, two points arise: (1) **ASys**: If the requirement is to understand behavior or person aspects in everyday life, if researchers let app users know the test hypothesis explicitly (e.g. that we are testing for stress, happiness, heavy alcohol usage) and allow self-monitoring, it might bias the mindset of people (e.g. I reporting that I am stressed everyday, or that I should relax today). These biases may get reflected in the data if people alter their behavior, resulting in researcher's drawing possibly wrong conclusions from the studies. (2) **FSys**: If we want to understand how the behavior of users changes over time due to the usage of smartphone sensing-based feedback systems, it is necessary to influence users by using self-monitoring mechanisms using interventions and feedback. A clear example for this is BeWell system [66], which measure how well young people adapt their lifestyle based on the feedback they get. Another example is the work of Farhan et al. [69], which generated real-time clinical interventions for participants who reported higher stress levels.

These two types of systems are the two corners of the spectrum, and there can be studies which have features of both. We believe that proper attention should be given to system design and experimentation by drawing principles from behavioral research, keeping in mind what each type of system offers in terms of testing hypotheses. Researchers should be aware of how experiments get affected by human bias, behavior, and psychology.

It should also be highlighted that FSys are better used when the basic hypothesis regarding human behavior has already been established using ASys or clinical methodologies. For example, if the experiments are done with a FSys to analyze the relationship between well-being and physical activity levels, the setting might lead to wrong conclusions because self-monitoring capability of the applications alter the behavior of young adults. This kind of relationships are better to be examined first with ASys to establish relationships, and then test for behavioral change that is occurring using FSys. A summary of some research questions that can be derived from this section are:

RQ7: How to systematically decide when ASys or FSys are the best methodological choice for a given problem?

RQ8: How to leverage knowledge in human sciences to quantify, control, and reduce possible influences on behavior due to the use of ASys?

E. UNDERUSED PILLARS OF DATA AND SENSOR TYPES

Table 3 examines which sensor (passive sensing modality) is used as a proxy for a variable belonging to each pillar of data, and Table 4 examines which self-report modality has been used to get data. Both tables provide interesting perspectives to understand research gaps in the mobile sensing applications that target young adults. Table 3 shows that

accelerometer has been directly used as a proxy to variables belonging to the behavior pillar. The table also shows an opportunity as accelerometer has not been used as a direct proxy to the person pillar. Moreover, it also shows that interaction sensing modalities have rarely been used as proxies for the person or context pillars, that opens up opportunities for mobile health research to leverage such relationships (established in clinical research), and to use such findings in applications that target well-being aspects of young adults. Further, Table 4 also opens up interesting avenues for future mobile sensing research. As examples, pictures and videos, as well as audio have been used less frequently as proxies for variables belonging to the person or context pillars. This provides opportunities to assess whether such relationships exist in other research domains (e.g. inferring ambiance of a surrounding place using a video/photo, etc), and use such relationships in mobile sensing studies. As a summary, two RQs for future work are:

RQ9: Can continuous sensing techniques (e.g. accelerometer, proximity, gyroscope, location, ambient light, audio, bluetooth, wifi, etc.) be used as proxies to attributes that belong to the person pillar (e.g. stress, emotions, depression, sociability, mood, happiness, etc.)?

RQ10: Can interaction sensing techniques (e.g. app usage behavior, typing events, touch events, battery levels, phone calls, messages, screen events, etc.) be used as proxies to attributes that belong to the context pillar (social context, semantic location, ambient light, temperature, crowdedness, etc.)?

F. PARTICIPANT RECRUITMENT STRATEGIES

We have highlighted several participant recruitment strategies used in this set of studies in Section III. As mentioned in that section, it is evident that computer science researchers have not paid enough systematic attention to participant recruitment strategies. Most studies used convenience samples without a more formal recruitment strategy. We believe that one open area of research would be to systematically compare different recruitment strategies from the perspective of utility for mobile sensing-based research. Research along this line would not only help research targeting young adults, but also mobile sensing research as a whole. While smartphone sensing studies in our scope have focused mainly on cost in terms of money and time when recruiting participants, human science research focuses on aspects such as scientific validity and ties with related clinical research [137], [138]. Moreover, in social science research, in addition to proper recruitment strategies, corrections are used to unbalanced variables such as age, gender, or ethnicity with rigorous statistical analysis [108]. This makes sure that results are scientifically valid.

RQ11: Do datasets collected from mobile health sensing studies from young adults reflect different recruitment strategies? What are the implications on data when not using systematic recruitment strategies as compared to clinical studies?

G. USING CONTEXTUAL TRIGGERS TO COLLECT SELF-REPORTS

When understanding the notions of each of these strategies from the perspective of pillars of data, some interesting observations can be made. Aspects that are often subjective, such as emotions, feelings, and stress, which belong to the person pillar, are difficult to be captured retrospectively. For example, asking a young adult how did they feel when they were in the lecture hall might introduce recall bias because socio-psychological notions are often subjective. On the other hand, self-reporting aspects such as eating or drinking behaviors that lead to objective answers might be easier to collect retrospectively, and in some cases they might be cumbersome to collect in-situ. For example, if young adults are asked to explicitly report their drinking behavior (data belonging to behavior and context pillars), in-situ reporting would be cumbersome to them if they are dancing in the club and it might lead to low response volumes. In fact, Bae et al. [82], have experimentally reported that retrospective reporting of alcohol drinking behavior have provided better results compared to in-situ reporting. In general, it can be said that, the more intimate the self-reported activity is, the more thought should be given to the subjectivity of such data, the privacy of users, and the ease of use of the application. While subjective data are better off being collected in-situ to avoid bias, the trade-off is that the application might be adding a layer of complexity. Another consideration regards the actual motivations of young adults in using the application. If the users are motivated to achieve a personal goal (e.g.: lose weight, reduce alcohol usage, eat healthy food), such users might not need push notifications reminding them to report their details. On the other hand, if the study participants are not intrinsically motivated to report certain aspects, then reminders might be useful.

RQ12: How can contextual triggers be used to improve mobile sensing data collection? What are the effects of these methodological choices on the final outcomes of a study?

H. TRANSFER LEARNING AND META-LEARNING

In computer vision problems, transfer learning is commonly used with the expectation that convolutional neural networks would learn certain features that would help classify other related datasets [139]. However, mobile health sensing data that are time series and tabular in nature are sparse and complex, and features of datasets are manually crafted, hence making transfer learning and dataset concatenation difficult [136], [140]. As an example, mobile health applications contain passive sensing information (high dimensional and high resolution) and self-reports (low dimensional). Moreover, different research studies analyzing similar aspects (e.g. food consumption [81], [83], stress [73]–[75], depression [69], [77], [85], and drinking episodes [82], [84]) have collected these behavioral traces using different protocols, phone types, and techniques resulting in datasets with contrasting characteristics and different types and number of features.

However, the fundamental types of data that are collected in behavioral smartphone sensing are under three main pillars of data P, B, and C according to the taxonomy we proposed. The general expectation is that these pillars would show high correlation across datasets that are similar in nature and analyze the same behavioral aspect. As an example, two studies regarding *stress* could use raw accelerometer traces or steps counts (derived in the smartphone using IMU sensors) belonging to B pillar to indicate the physical activity levels. Even though the collected data are different in the number and type of features, what they capture is the physical activity level, and we could expect those features to have similar correlations with stress. Hence, one approach could be to craft intermediate data representations from features belonging to similar pillars to seek the possibility of transfer learning, hence accommodating transfer learning among datasets in the same pillar and domain, while the collected datasets come from using slightly different protocols.

In addition to the challenging use-case mentioned above, there could be instances where training and testing distributions could differ, even if the same set of features are collected using the same procedure in training and deployments phases [136], [140], [141]. An example is a mobile sensing inference pipeline to count steps in android phones developed using data collected in the United States, being deployed in China. One could expect differences in the type of phones that are used (e.g. Samsung, Huawei, Sony, Xiomi, etc.), and also the way people use the phone, hence leading to different and inaccurate step count results when deployed in the wild, even if the same set of features are collected. In a recent study, Gong et al. emphasized this issue and provided a solution to the problem using meta-learning [136], showing that their technique MetaSense achieves state-of-the-art results even compared to other meta learning and transfer learning techniques for activity recognition and speech recognition tasks. Studies along this line would help this research in making sure that datasets and models in the domain are reusable, hence accelerating the growth of the research field. While most of the recent work [136], [140] are focused on speech and accelerometer sensors, it would be worth to look at other commonly used modalities in mobile health sensing. Moreover, this could be another way of accommodating diversity-awareness (discussed in Section VII-A) in smartphone sensing research through machine learning.

RQ13: How to build intermediate representations from mobile sensing data to allow transfer learning to different tasks, even when datasets are collected using different protocols?

RQ14: How can meta-learning be used for heterogeneous data collected in mobile sensing applications, when datasets are collected using the same protocol but from diverse population groups?

I. OTHER ISSUES

In this study, we came across literature in pillars P and B, for domains such as stress, depression, anxiety, alcohol consumption, eating behavior, etc. as shown in Table 2. In

addition, there exists opportunities for researchers to expand upon the current work that are focused on P and B pillars, and to examine whether smartphone sensors could be used to characterize contextual attributes that are linked to domains in P and B. Consider the following example regarding eating behavior for clarity. Smartphone sensing involving eating behavior takes the form of mobile applications that combine passive smartphone sensing with mobile food diaries: (1) Detecting eating events - this type of studies could be done with the aim of detecting the time of eating. Hence, a smartphone should be able to understand the time at which a user would eat. This task has been previously attempted using wearable sensing [142], [143]. Hence, given the time of the day, the smartphone should be able to determine whether a user is eating or not; (2) Characterizing eating events - food types and categories are the most important aspects regarding the food consumption. A smartphone application that could unobtrusively infer categories of food taken by users depending on sensed contextual cues could be a powerful tool in providing users with valuable interventions and analytics. The study by Biel et al [83] belongs to this category. In addition, this step could go beyond the consumed food type, and attempt to characterize the eating event in more depth by understanding food consumption levels and the contextual aspects that affect food consumption. For example, aspects such as the social context of eating (an contextual attribute belonging to pillar C) is known to affect the food consumption behavior of people according to concepts such as social facilitation and impression management [131], [144], [145]. The social context also affects the amount of food people eat – i.e. highly social contexts can lead to overeating [131], [146]. If smartphone sensing capabilities could be used to sense the social context of eating unobtrusively, such inferences could be used to provide users with context-aware interventions – e.g. if a highly social eating context is sensed, notify the users to be aware of the food intake.

As shown by the two steps above, even for other domains mentioned in Table 2, a similar framework could be used to explore research directions for future work. In general terms, the two main steps that we identified are: (1) **Event Detection** - given the *time* of the day, smartphone should be able to unobtrusively infer whether an event is occurring. Some examples for events are feeling happy, feeling stressed, not feeling depressed, eating food, and drinking alcohol; once the events are detected, the next step is to (2) **Event Characterization** - given that an event is detected (the time of the event is known), smartphones could infer attributes related to the context of the event. For a domain like depression, the characterizations could be related to the ambiance of the environment or the social context that might affect a person. For a domain such as alcohol consumption, characterizations could be related to social context of drinking or types of drinks that are consumed (e.g. beer, wine, other spirits, etc.). Similarly, understanding smartphone sensing research done in various domains using the above mentioned two-fold categorization would allow future research to go into

more depth in each of the domains. Hence, according to findings in Table 2, smartphone sensing research could look into characterizing events in domains where detection has been shown to be feasible.

There are several other well-being related areas that are common among young adults, but discussed relatively less in smartphone sensing research according to Table 2. For example, aspects such as anger and violence, have been identified as important aspects that affect young adults according to prior literature [147], but have not been given proper attention in smartphone sensing research. In addition, there are certain domains that have been discussed in human-computer interaction (HCI) literature, but have not been studied in smartphone sensing. For example, gender specific issues such as menstruation, while being discussed in HCI literature for sometime [148], has not been studied in smartphone sensing research. Moreover, the study by Epstien et al. [148] emphasized that currently available menstrual cycle tracking apps (that do not use passive sensing) do not take into account diversity aspects that affect the menstrual cycle length such as young adulthood, pregnancy, and menopause. This again is an example of proof of why smartphone sensing techniques should be designed with diversity-awareness in mind. Furthermore, self-harm is another well-being aspect that has been discussed in HCI literature as a common issue among young people [149], but has not been given enough emphasis in smartphone sensing studies. Finally, some other issues faced by young adults that have not been addressed using smartphone sensing research include self-esteem [150], obsessive compulsive disorder [151], and addictive behaviors [152], [153], for which there are both a body of literature and active work in public health research.

RQ15: Even though mood instability and stress have been discussed in smartphone sensing studies regarding young adults, aspects such as self-harm, anger, and violence have not been given proper attention in the domain. Can smartphone sensing and machine learning models, integrated with ethical and privacy-by-design principles, be used to support young adult populations regarding these situations?

J. ETHICAL ISSUES ON FUTURE SMARTPHONE SENSING

As the discussion above highlights, the use of smartphone sensing for studies involving young adults raises a number of ethical issues. First, personal data is involved in the studies reviewed here, including both sensor data like location, as well as survey data where participants disclose information including their habits, state of mind, and even sensitive information such as age, gender, etc. [129]. Furthermore, some of the reviewed studies address rather sensitive issues like mental health. Clearly, this research requires both legal and ethical review procedures, which are well established in many countries, although not in the whole world. The existence of recent regulations like the European General Data Protection Regulation (GDPR), and similar efforts in other parts of the world, provide frameworks to address some

of these issues, but also uncover potential inconsistencies across world regions. This is especially important for studies that aim to engage diverse populations across countries. Second, many of the current concerns about machine learning-based decision making systems are also applicable to the mobile sensing setting, including issues of fairness, accountability, and transparency. This is an issue that will become increasingly important. Finally, the commercial use of smartphone sensing apps to monitor aspects of young adults well-being also poses ethical questions, where economic interests of companies may not always align with the objective of supporting young adults to achieve and sustain their own goals regarding well-being (see for instance recent journalistic coverage of mobile health apps for college students in the US [154]–[156]). These fundamental issues need to be considered by researchers as future studies and systems are designed and deployed.

VIII. CONCLUSION

In this paper, we extensively analyzed studies that leverage smartphone sensing for the well-being of young adults. By specifically targeting young people, we narrowed down the type of well-being related aspects, and hence identified how self-reports and passive sensing of smartphones have been applied in a very specific context. First, we used human-science concepts and theories to frame behavioral research in smartphone sensing using a solid framework. We analyzed the Data Perspective of smartphone sensing using pillars of data, and the System Perspective using self-monitoring related concepts. We emphasized aspects including, but not limited to the importance of focusing on interaction sensing to better understand young adults; feedback systems to test for behavioral change; analytical systems to test for hypothesis regarding specific behavioral aspects; diversity-awareness of smartphone sensing to support diverse groups of people; developing feedback systems that are robust to sensor failure; and using systematic recruitment strategies for mobile sensing studies regarding young adults.

REFERENCES

- [1] Y. Deng, "Deep learning on mobile devices - A review," *CoRR*, vol. abs/1904.09274, 2019.
- [2] L. Meegahapola, T. Kandappu, K. Jayarajah, L. Akoglu, S. Xiang, and A. Misra, "Buscope: Fusing individual and aggregated mobility behavior for live smart city services," in *Proceedings of the 17th Annual International Conference on Mobile Systems, Applications, and Services*, ser. *MobiSys '19*. New York, NY, USA: ACM, 2019, pp. 41–53.
- [3] K. Jayarajah, V. Subbaraju, N. Athaide, L. Meegahapola, A. Tan, and A. Misra, "Can multimodal sensing detect and localize transient events?" in *Ground/Air Multisensor Interoperability, Integration, and Networking for Persistent ISR IX*, M. A. Kolodny, D. M. Wiegmann, and T. Pham, Eds., vol. 10635, International Society for Optics and Photonics. SPIE, 2018, pp. 331 – 343.
- [4] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "Mobilenets: Efficient convolutional neural networks for mobile vision applications," *ArXiv*, vol. abs/1704.04861, 2017.
- [5] M. Vidanapathirana, L. Meegahapola, and I. Perera, "Cognitive analysis of 360 degree surround photos," *CoRR*, vol. abs/1901.05634, 2019. [Online]. Available: <http://arxiv.org/abs/1901.05634>

- [6] D. Rathnayake, A. de Silva, D. Puwakdandawa, L. Meegahapola, A. Misra, and I. Perera, "Jointly optimizing sensing pipelines for multimodal mixed reality interaction," in *Proceedings of the 17th International Conference on Mobile Ad Hoc and Sensor Systems (MASS)*, ser. MASS '20. IEEE, 2020.
- [7] R. Wang, W. Wang, A. daSilva, J. F. Huckins, W. M. Kelley, T. F. Heatherton, and A. T. Campbell, "Tracking depression dynamics in college students using mobile phone and wearable sensing," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 2, no. 1, pp. 43:1–43:26, Mar. 2018.
- [8] L. Meegahapola, V. Subramaniam, L. M. Kaplan, and A. Misra, "Prior activation distribution (PAD): A versatile representation to utilize DNN hidden units," *CoRR*, vol. abs/1907.02711, 2019. [Online]. Available: <http://arxiv.org/abs/1907.02711>
- [9] M. Raento, A. Oulasvirta, and N. Eagle, "Smartphones: An emerging tool for social scientists," *Sociological Methods & Research*, vol. 37, no. 3, pp. 426–454, 2009.
- [10] L. Coles-Kemp, R. B. Jensen, and R. Talhouk, "In a new land: Mobile phones, amplified pressures and reduced capabilities," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, ser. CHI '18. New York, NY, USA: ACM, 2018, pp. 584:1–584:13.
- [11] L. Meegahapola and I. Perera, "Enhanced in-store shopping experience through smart phone based mixed reality application," in *2017 Seventeenth International Conference on Advances in ICT for Emerging Regions (ICTer)*, 2017, pp. 1–8.
- [12] (2019) Mobile fact sheet. [Online]. Available: <https://www.pewinternet.org/fact-sheet/mobile/>
- [13] (2019) Smartphone addiction facts & phone usage statistics. [Online]. Available: <https://www.bankmycell.com/blog/smartphone-addiction/>
- [14] (2019) Qualcomm's new chip brings a.i. smarts (and 5g!) to 2019 flagship android phones. [Online]. Available: <https://www.digitaltrends.com/mobile/qualcomm-snapdragon-855-news/>
- [15] (2019) Google playstore. [Online]. Available: <https://play.google.com/store?hl=en>
- [16] (2019) Apple appstore. [Online]. Available: <https://www.apple.com/ios/app-store/>
- [17] J. Farringdon, A. J. Moore, N. Tilbury, J. Church, and P. D. Biemond, "Wearable sensor badge and sensor jacket for context awareness," in *Digest of Papers. Third International Symposium on Wearable Computers*, Oct 1999, pp. 107–113.
- [18] S. Park, C. Gopalsamy, R. Rajamanickam, and S. Jayaraman, "The wearable motherboard: a flexible information infrastructure or sensate liner for medical applications," *Studies in health technology and informatics*, vol. 62, p. 252–258, 1999.
- [19] A. Pentland, "Looking at people: sensing for ubiquitous and wearable computing," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 1, pp. 107–119, Jan 2000.
- [20] N. Eagle and A. (Sandy) Pentland, "Reality mining: sensing complex social systems," *Personal and Ubiquitous Computing*, vol. 10, no. 4, pp. 255–268, May 2006.
- [21] S. Consolvo, D. W. McDonald, T. Toscos, M. Y. Chen, J. Froehlich, B. Harrison, P. Klasnja, A. LaMarca, L. LeGrand, R. Libby, I. Smith, and J. A. Landay, "Activity sensing in the wild: A field trial of ubifit garden," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '08. New York, NY, USA: ACM, 2008, pp. 1797–1806.
- [22] J. Froehlich, M. Y. Chen, S. Consolvo, B. Harrison, and J. A. Landay, "Myexperience: A system for in situ tracing and capturing of user feedback on mobile phones," in *Proceedings of the 5th International Conference on Mobile Systems, Applications and Services*, ser. MobiSys '07. New York, NY, USA: ACM, 2007, pp. 57–70.
- [23] N. D. Lane, E. Miluzzo, H. Lu, D. Peebles, T. Choudhury, and A. T. Campbell, "A survey of mobile phone sensing," *IEEE Communications Magazine*, vol. 48, no. 9, pp. 140–150, Sep. 2010.
- [24] D. Rizzuto, N. Orsini, C. Qiu, H.-X. Wang, and L. Fratiglioni, "Lifestyle, social factors, and survival after age 75: population based study," *BMJ*, vol. 345, 2012.
- [25] L. Cohen, G. C. Curhan, and J. P. Forman, "Influence of age on the association between lifestyle factors and risk of hypertension," *J Am Soc Hypertens*, 2012.
- [26] N. M. Petry, "A Comparison of Young, Middle-Aged, and Older Adult Treatment-Seeking Pathological Gamblers," *The Gerontologist*, vol. 42, no. 1, pp. 92–99, 02 2002.
- [27] V. Patel, A. J. Flisher, S. Hetrick, and P. McGorry, "Mental health of young people: a global public-health challenge," *The Lancet*, vol. 369, no. 9569, pp. 1302 – 1313, 2007.
- [28] K. Houston, K. Hawton, and R. Shepperd, "Suicide in young people aged 15–24: a psychological autopsy study," *Journal of Affective Disorders*, vol. 63, no. 1, pp. 159 – 170, 2001.
- [29] M. D. Newcomb and P. M. Bentler, "Impact of adolescent drug use and social support on problems of young adults: A longitudinal study," *Journal of Abnormal Psychology*, 1988.
- [30] M. Tabuchi, T. Nakagawa, A. Miura, and Y. Gondo, "Generativity and Interaction Between the Old and Young: The Role of Perceived Respect and Perceived Rejection," *The Gerontologist*, vol. 55, no. 4, pp. 537–547, 11 2013.
- [31] H. Sung, R. L. Siegel, P. S. Rosenberg, and A. Jemal, "Emerging cancer trends among young adults in the usa: analysis of a population-based cancer registry," 2019.
- [32] M. De Nadai, A. Cardoso, A. Lima, B. Lepri, and N. Oliver, "Strategies and limitations in app usage and human mobility," *Scientific Reports*, vol. 9, p. 10935, 07 2019.
- [33] HealthLinkBC, "Alcohol and drug use in young people," 2019. [Online]. Available: <https://www.healthlinkbc.ca/health-topics/tp17749>
- [34] HealthLink-BC, "Statistics of alcohol and drug use in young people," 2019. [Online]. Available: <https://www.healthlinkbc.ca/health-topics/tp17749#uq2408>
- [35] J. E. Pelletier, D. J. Graham, and M. N. Laska, "Social norms and dietary behaviors among young adults," *American Journal of Health Behavior*, vol. 38, no. 1, pp. 144–152, 2014.
- [36] pewinternet, "Social media fact sheet," 2019. [Online]. Available: <https://www.pewinternet.org/fact-sheet/social-media/>
- [37] I. Andone, K. Błazkiewicz, M. Eibes, B. Trendafilov, C. Montag, and A. Markowitz, "How age and gender affect smartphone usage," in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*, ser. UbiComp '16. New York, NY, USA: Association for Computing Machinery, 2016, p. 9–12.
- [38] A. J. van Deursen, C. L. Bolle, S. M. Hegner, and P. A. Kommers, "Modeling habitual and addictive smartphone behavior: The role of smartphone usage types, emotional intelligence, social stress, self-regulation, age, and gender," *Computers in Human Behavior*, vol. 45, pp. 411 – 420, 2015.
- [39] (2019) Human-centred design in the field - how might we design with and for young people? [Online]. Available: <https://www.unicef.org/innovation/reports/human-centred-design-field>
- [40] G. Chen and D. Kotz, "A survey of context-aware mobile computing research," Hanover, NH, USA, Tech. Rep., 2000.
- [41] K. E. McCracken and J.-Y. Yoon, "Recent approaches for optical smartphone sensing in resource-limited settings: a brief review," *Anal. Methods*, vol. 8, pp. 6591–6601, 2016.
- [42] D. Christin, A. Reinhardt, S. S. Kanhere, and M. Hollick, "A survey on privacy in mobile participatory sensing applications," *Journal of Systems and Software*, vol. 84, no. 11, pp. 1928 – 1946, 2011, mobile Applications: Status and Trends.
- [43] A. Kamilaris and A. Pitsillides, "Mobile phone computing and the internet of things: A survey," *IEEE Internet of Things Journal*, vol. 3, no. 6, pp. 885–898, Dec 2016.
- [44] X. Zhang, Z. Yang, W. Sun, Y. Liu, S. Tang, K. Xing, and X. Mao, "Incentives for mobile crowd sensing: A survey," *IEEE Communications Surveys Tutorials*, vol. 18, no. 1, pp. 54–67, Firstquarter 2016.
- [45] N. Fernando, S. W. Loke, and W. Rahayu, "Mobile cloud computing: A survey," *Future Generation Computer Systems*, vol. 29, no. 1, pp. 84 – 106, 2013, including Special section: AIRCC-NetCoM 2009 and Special section: Clouds and Service-Oriented Architectures.
- [46] F. Calabrese, L. Ferrari, and V. D. Blondel, "Urban sensing using mobile phone network data: A survey of research," *ACM Comput. Surv.*, vol. 47, no. 2, pp. 25:1–25:20, Nov. 2014.
- [47] M. B. del Rosario, S. J. Redmond, and N. H. Lovell, "Tracking the evolution of smartphone sensing for monitoring human movement," *Sensors*, 2015.
- [48] W. Z. Khan, Y. Xiang, M. Y. Aalsalem, and Q. Arshad, "Mobile phone sensing systems: A survey," *IEEE Communications Surveys Tutorials*, vol. 15, no. 1, pp. 402–427, First 2013.
- [49] A. K. Triantafyllidis, C. Velardo, D. Salvi, S. A. Shah, V. G. Koutkias, and L. Tarassenko, "A survey of mobile phone sensing, self-reporting, and social sharing for pervasive healthcare," *IEEE Journal of Biomedical and Health Informatics*, vol. 21, no. 1, pp. 218–227, Jan 2017.

- [50] V. P. Cornet and R. J. Holden, "Systematic review of smartphone-based passive sensing for health and wellbeing," *Journal of Biomedical Informatics*, vol. 77, pp. 120–132, 2018.
- [51] G. M. Harari, S. R. Müller, M. S. Aung, and P. J. Rentfrow, "Smartphone sensing methods for studying behavior in everyday life," *Current Opinion in Behavioral Sciences*, vol. 18, pp. 83–90, 2017, big data in the behavioural sciences.
- [52] M. H. Aung, M. Matthews, and T. Choudhury, "Sensing behavioral symptoms of mental health and delivering personalized interventions using mobile technologies," *Depression and Anxiety*, vol. 34, no. 7, pp. 603–609, 2017.
- [53] A. Trifan, M. Oliveira, and J. L. Oliveira, "Passive sensing of health outcomes through smartphones: Systematic review of current solutions and possible limitations," *JMIR Mhealth Uhealth*, vol. 7, no. 8, p. e12649, Aug 2019.
- [54] D. C. Mohr, M. Zhang, and S. M. Schueller, "Personal sensing: Understanding mental health using ubiquitous sensors and machine learning," *Annual Review of Clinical Psychology*, vol. 13, no. 1, pp. 23–47, 2017.
- [55] Y. S. Can, B. Amrich, and C. Ersoy, "Stress detection in daily life scenarios using smart phones and wearable sensors: A survey," *Journal of Biomedical Informatics*, vol. 92, p. 103139, 2019.
- [56] K. Lewin, in *Principles of topological psychology*, 1936.
- [57] J. M. Snyder, "Self-monitoring of expressive behavior," *Journal of Personality and Social Psychology*, vol. 30, pp. 526–537, 1974.
- [58] S. Michie, S. Williams, F. Sniehotta, S. Dombrowski, A. Bishop, and D. French, "A refined taxonomy of behaviour change techniques to help people change their physical activity and healthy eating behaviours: The calo-re taxonomy," *Psychology & health*, vol. 26, pp. 1479–98, 06 2011.
- [59] S. Compemolle, A. DeSmet, L. Poppe, G. Crombez, I. Bourdeaudhuij, G. Cardon, H. van der Ploeg, and D. Dyck, "Effectiveness of interventions using self-monitoring to reduce sedentary behavior in adults: a systematic review and meta-analysis," *International Journal of Behavioral Nutrition and Physical Activity*, vol. 16, 12 2019.
- [60] (2020) How to: Teach students to change behaviors through self-monitoring. [Online]. Available: <https://www.interventioncentral.org/node/961544>
- [61] A. Bandura, "Social cognitive theory of self-regulation," *Organizational Behavior and Human Decision Processes*, vol. 50, no. 2, pp. 248–287, 1991, theories of Cognitive Self-Regulation.
- [62] tim boone, anthony j. reilly, and M. Sashkin, "Social learning theory albert bandura englewood cliffs, n.j.: Prentice-hall, 1977. 247 pp., paperbound," *Group & Organization Studies*, vol. 2, no. 3, pp. 384–385, 1977.
- [63] M. B. Morshed, K. Saha, R. Li, S. K. D' Mello, M. De Choudhury, G. D. Abowd, and T. Plötz, "Prediction of mood instability with passive sensing," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 3, no. 3, Sep. 2019.
- [64] R. LiKamWa, Y. Liu, N. D. Lane, and L. Zhong, "Moodscope: Building a mood sensor from smartphone usage patterns," in *Proceeding of the 11th Annual International Conference on Mobile Systems, Applications, and Services*, ser. MobiSys '13. New York, NY, USA: ACM, 2013, pp. 389–402.
- [65] Y. Ma, B. Xu, Y. Bai, G. Sun, and R. Zhu, "Daily mood assessment based on mobile phone sensing," in *2012 Ninth International Conference on Wearable and Implantable Body Sensor Networks*, May 2012, pp. 142–147.
- [66] M. Lin, N. D. Lane, M. Mohammad, X. Yang, H. Lu, G. Cardone, S. Ali, A. Doryab, E. Berke, A. T. Campbell, and T. Choudhury, "Bewell+: Multi-dimensional wellbeing monitoring with community-guided user feedback and energy optimization," in *Proceedings of the Conference on Wireless Health*, ser. WH '12. New York, NY, USA: ACM, 2012, pp. 10:1–10:8.
- [67] S. Servia-Rodríguez, K. K. Rachuri, C. Mascolo, P. J. Rentfrow, N. Lathia, and G. M. Sandstrom, "Mobile sensing at the service of mental well-being: A large-scale longitudinal study," in *Proceedings of the 26th International Conference on World Wide Web*, ser. WWW '17. Republic and Canton of Geneva, Switzerland: International World Wide Web Conferences Steering Committee, 2017, pp. 103–112.
- [68] M. Rabbi, M. H. Aung, M. Zhang, and T. Choudhury, "Mybehavior: Automatic personalized health feedback from user behaviors and preferences using smartphones," in *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, ser. UbiComp '15. New York, NY, USA: ACM, 2015, pp. 707–718.
- [69] Farhan, Yue, Morillo, Ware, Lu, Bi, Kamath, Russell, Bamis, and Wang, "Behavior vs. introspection: refining prediction of clinical depression via smartphone sensing data," in *2016 IEEE Wireless Health (WH)*, Oct 2016, pp. 1–8.
- [70] R. Wang, F. Chen, Z. Chen, T. Li, G. Harari, S. Tignor, X. Zhou, D. Ben-Zeev, and A. T. Campbell, "Studentlife: Assessing mental health, academic performance and behavioral trends of college students using smartphones," in *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, ser. UbiComp '14. New York, NY, USA: ACM, 2014, pp. 3–14.
- [71] K. K. Rachuri, M. Musolesi, C. Mascolo, P. J. Rentfrow, C. Longworth, and A. Aucinas, "Emotionsense: A mobile phones based adaptive platform for experimental social psychology research," in *Proceedings of the 12th ACM International Conference on Ubiquitous Computing*, ser. UbiComp '10. New York, NY, USA: ACM, 2010, pp. 281–290.
- [72] S. T. Moturu, I. Khayal, N. Aharony, W. Pan, and A. S. Pentland, "Using social sensing to understand the links between sleep, mood, and sociability," in *Proceedings of IEEE International Conference on Social Computing*, September 2011.
- [73] H. Lu, D. Frauendorfer, M. Rabbi, M. S. Mast, G. T. Chittaranjan, A. T. Campbell, D. Gatica-Perez, and T. Choudhury, "Stressense: Detecting stress in unconstrained acoustic environments using smartphones," in *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, ser. UbiComp '12. New York, NY, USA: ACM, 2012, pp. 351–360.
- [74] A. Sano and R. W. Picard, "Stress recognition using wearable sensors and mobile phones," in *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction*, Sep. 2013, pp. 671–676.
- [75] A. Bogomolov, B. Lepri, and F. Pianesi, "Happiness recognition from mobile phone data," in *2013 International Conference on Social Computing*, Sep. 2013, pp. 790–795.
- [76] A. Bogomolov, B. Lepri, M. Ferron, F. Pianesi, and A. S. Pentland, "Daily stress recognition from mobile phone data, weather conditions and individual traits," in *Proceedings of the 22Nd ACM International Conference on Multimedia*, ser. MM '14. New York, NY, USA: ACM, 2014, pp. 477–486.
- [77] L. Canzian and M. Musolesi, "Trajectories of depression: Unobtrusive monitoring of depressive states by means of smartphone mobility traces analysis," in *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, ser. UbiComp '15. New York, NY, USA: ACM, 2015, pp. 1293–1304.
- [78] S. Abdullah, E. L. Murnane, M. Matthews, M. Kay, J. A. Kientz, G. Gay, and T. Choudhury, "Cognitive rhythms: Unobtrusive and continuous sensing of alertness using a mobile phone," in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, ser. UbiComp '16. New York, NY, USA: ACM, 2016, pp. 178–189.
- [79] Y. Huang, H. Xiong, K. Leach, Y. Zhang, P. Chow, K. Fua, B. A. Teachman, and L. E. Barnes, "Assessing social anxiety using gps trajectories and point-of-interest data," in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, ser. UbiComp '16. New York, NY, USA: ACM, 2016, pp. 898–903.
- [80] E. L. Murnane, S. Abdullah, M. Matthews, M. Kay, J. A. Kientz, T. Choudhury, G. Gay, and D. Cosley, "Mobile manifestations of alertness: Connecting biological rhythms with patterns of smartphone app use," in *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services*, ser. MobileHCI '16. New York, NY, USA: ACM, 2016, pp. 465–477.
- [81] E. Seto, J. Hua, L. Wu, V. Shia, S. Eom, M. Wang, and Y. Li, "Models of individual dietary behavior based on smartphone data: The influence of routine, physical activity, emotion, and food environment," *PLOS ONE*, vol. 11, no. 4, pp. 1–16, 04 2016.
- [82] S. Bae, D. Ferreira, B. Suffoletto, J. C. Puyana, R. Kurtz, T. Chung, and A. K. Dey, "Detecting drinking episodes in young adults using smartphone-based sensors," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 1, no. 2, pp. 5:1–5:36, Jun. 2017.
- [83] J.-I. Biel, N. Martin, D. Labbe, and D. Gatica-Perez, "Bites‘n’bits: Inferring eating behavior from contextual mobile data," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 1, no. 4, pp. 125:1–125:33, Jan. 2018.
- [84] D. Santani, T. Do, F. Labhart, S. Landolt, E. Kuntsche, and D. Gatica-Perez, "Drinksense: Characterizing youth drinking behavior using smartphones," *IEEE Transactions on Mobile Computing*, vol. 17, no. 10, pp. 2279–2292, Oct 2018.

- [85] M. Boukhechba, A. R. Daros, K. Fua, P. I. Chow, B. A. Teachman, and L. E. Barnes, "Demonicsalmon: Monitoring mental health and social interactions of college students using smartphones," *Smart Health*, vol. 9-10, pp. 192–203, 2018, cCHASE 2018 Special Issue.
- [86] J. Gong, Y. Huang, P. I. Chow, K. Fua, M. S. Gerber, B. A. Teachman, and L. E. Barnes, "Understanding behavioral dynamics of social anxiety among college students through smartphone sensors," *Information Fusion*, vol. 49, pp. 57–68, 2019.
- [87] E. Sagbas, S. Korukoglu, and S. Balli, "Stress detection via keyboard typing behaviors by using smartphone sensors and machine learning techniques," *Journal of Medical Systems*, vol. 44, 02 2020.
- [88] M. Obuchi, J. F. Huckins, W. Wang, A. daSilva, C. Rogers, E. Murphy, E. Hedlund, P. Holtzheimer, S. Mirjafari, and A. Campbell, "Predicting brain functional connectivity using mobile sensing," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 4, no. 1, Mar. 2020.
- [89] S. Abdullah, M. Matthews, E. L. Murnane, G. Gay, and T. Choudhury, "Towards circadian computing: "early to bed and early to rise" makes some of us unhealthy and sleep deprived," in *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, ser. UbiComp '14. New York, NY, USA: Association for Computing Machinery, 2014, p. 673–684.
- [90] R. Whittaker, H. McRobbie, C. Bullen, A. Rodgers, and Y. Gu, "Mobile phone-based interventions for smoking cessation," *Cochrane Database of Systematic Reviews*, no. 4, 2016.
- [91] AlzheimersResearchUK, "Play sea hero quest and #gameforgood," 2019. [Online]. Available: <https://www.alzheimersresearchuk.org/our-research/what-we-do/sea-hero-quest/>
- [92] A. K. Dey, K. Wac, D. Ferreira, K. Tassini, J.-H. Hong, and J. Ramos, *Getting closer: an empirical investigation of the proximity of user to their smart phones*, ser. Proceedings of the 13th ACM International Conference on Ubiquitous Computing - UbiComp'11. New York: ACM, 2011, pp. 163–172, iD: unige:72664.
- [93] Google, "Activity recognition api," 2019. [Online]. Available: <https://developers.google.com/location-context/activity-recognition/>
- [94] T.-T. Phan, S. Muralidhar, and D. Gatica-Perez, "#drink or #drunk: Multimodal signals and drinking practices on instagram," in *Proceedings of the 13th EAI International Conference on Pervasive Computing Technologies for Healthcare*, ser. PervasiveHealth'19. New York, NY, USA: Association for Computing Machinery, 2019, p. 71–80.
- [95] G. M. Sandstrom, N. Lathia, C. Mascolo, and P. J. Rentfrow, "Putting mood in context: Using smartphones to examine how people feel in different locations," *Journal of Research in Personality*, vol. 69, pp. 96–101, 2017, within-Person Variability in Personality.
- [96] N. Kamarulzaman, A. Saleh, S. Hashim, H. Hashim, and A. Abdul-Ghani, "An overview of the influence of physical office environments towards employee," *Procedia Engineering*, vol. 20, pp. 262–268, 2011, 2nd International Building Control Conference.
- [97] K. C. Smolders and Y. A. de Kort, "Bright light and mental fatigue: Effects on alertness, vitality, performance and physiological arousal," *Journal of Environmental Psychology*, vol. 39, pp. 77–91, 2014, light, lighting, and human behaviour.
- [98] S. Sarkar, "Is google home listening to me?" 2019.
- [99] M. Day, G. Turner, and N. Drozdziak, "Amazon workers are listening to what you tell alexa," 2019.
- [100] N. Ahmad, R. A. Raja Ghazilla, N. Khairi, and V. Kasi, "Reviews on various inertial measurement unit (imu) sensor applications," *International Journal of Signal Processing Systems*, vol. 1, pp. 256–262, 01 2013.
- [101] T. M. T. Do and D. Gatica-Perez, "Groupus: Smartphone proximity data and human interaction type mining," in *2011 15th Annual International Symposium on Wearable Computers*, June 2011, pp. 21–28.
- [102] A. Mathur, L. M. Kalanadhabhatta, R. Majethia, and F. Kawsar, "Moving beyond market research: Demystifying smartphone user behavior in india," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 1, no. 3, pp. 82:1–82:27, Sep. 2017.
- [103] M. Böhmer, B. Hecht, J. Schöning, A. Krüger, and G. Bauer, "Falling asleep with angry birds, facebook and kindle: A large scale study on mobile application usage," in *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services*, ser. MobileHCI '11. New York, NY, USA: ACM, 2011, pp. 47–56.
- [104] T. M. T. Do, J. Blom, and D. Gatica-Perez, "Smartphone usage in the wild: A large-scale analysis of applications and context," in *Proceedings of the 13th International Conference on Multimodal Interfaces*, ser. ICMi '11. New York, NY, USA: ACM, 2011, pp. 353–360.
- [105] T.-M.-T. Do and D. Gatica-Perez, "By their apps you shall understand them: Mining large-scale patterns of mobile phone usage," in *Proceedings of the 9th International Conference on Mobile and Ubiquitous Multimedia*, ser. MUM '10. New York, NY, USA: ACM, 2010, pp. 27:1–27:10.
- [106] J. Comstock, "Survey: Two thirds of teens, young adults have used a health app," in *HopeLab*, 2018. [Online]. Available: <https://www.mobihealthnews.com/content/survey-two-thirds-teens-young-adults-have-used-health-app>
- [107] N. Park and H. Lee, "Social implications of smartphone use: Korean college students' smartphone use and psychological well-being," *Cyberpsychology, Behavior, and Social Networking*, vol. 15, no. 9, pp. 491–497, 2012.
- [108] J. D. Elhai, J. C. Levine, R. D. Dvorak, and B. J. Hall, "Non-social features of smartphone use are most related to depression, anxiety and problematic smartphone use," *Computers in Human Behavior*, vol. 69, pp. 75–82, 2017.
- [109] A. Lepp, J. E. Barkley, G. J. Sanders, M. J. Rebold, and P. S. Gates, "The relationship between cell phone use, physical and sedentary activity, and cardiorespiratory fitness in a sample of u.s. college students," in *International Journal of Behavioral Nutrition and Physical Activity*, 2013.
- [110] V. Pejovic, N. Lathia, C. Mascolo, and M. Musolesi, "Mobile-based experience sampling for behaviour research," *Emotions and Personality in Personalized Services*, p. 141–161, 2016.
- [111] S. Shiffman, A. Stone, and H. MR., "Ecological momentary assessment," ser. Annual Review of Clinical Psychology, 2008, pp. 1–32.
- [112] C. Cakebread, "People will take 1.2 trillion digital photos this year — thanks to smartphones," in *Business Insider*, 2017. [Online]. Available: <https://www.businessinsider.com/12-trillion-photos-to-be-taken-in-2017-thanks-to-smartphones-chart-2017-8?r=US&IR=T>
- [113] A. Mehrotra and M. Musolesi, "Intelligent notification systems: A survey of the state of the art and research challenges," 2017.
- [114] Warc, "Millennials act on app push notification," in *Warc*, 2015. [Online]. Available: https://www.warc.com/newsandopinion/news/millennials_act_on_app_push_notification/35626
- [115] M. Pielot and L. Rello, "Productive, anxious, lonely: 24 hours without push notifications," in *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services*, ser. MobileHCI '17. New York, NY, USA: ACM, 2017, pp. 11:1–11:11.
- [116] R. Gibb, "How consumers perceive push notifications in 2018," 2018. [Online]. Available: <http://info.localytics.com/blog/push-notification-survey-2018>
- [117] A. Zantal-Wiener, "7 types of push notifications users actually enjoy," 2019. [Online]. Available: <https://blog.hubspot.com/marketing/push-notification-types>
- [118] L. G. Morrison, C. Hargood, V. Pejovic, A. W. A. Geraghty, S. Lloyd, N. Goodman, D. T. Michaelides, A. Weston, M. Musolesi, M. J. Weal, and L. Yardley, "The effect of timing and frequency of push notifications on usage of a smartphone-based stress management intervention: An exploratory trial," *PLOS ONE*, pp. 1–15, 01 2017.
- [119] K. Lewin, "Action research and minority problems," *Journal of Social Issues*, vol. 2, no. 4, pp. 34–46, 1946.
- [120] M. Park, H. Yoo, J. Kim, and J. Lee, "Why do young people use fitness apps? cognitive characteristics and app quality," *Electronic Commerce Research*, vol. 18, no. 4, pp. 755–761, Dec 2018.
- [121] M. J. K. Z. F.-C. S. D. P. W. R. and D. C., "User preferences for content, features, and style for an app to reduce harmful drinking in young adults: Analysis of user feedback in app stores and focus group interviews," *JMIR Mhealth Uhealth*, vol. 4, no. 2, 2016.
- [122] Z. Gong, P. Zhong, and W. Hu, "Diversity in machine learning," *IEEE Access*, vol. 7, p. 64323–64350, 2019.
- [123] J. Vincent, "Ibm hopes to fight bias in facial recognition with new diverse dataset," 2018. [Online]. Available: <https://www.theverge.com/2018/6/27/17509400/facial-recognition-bias-ibm-data-training>
- [124] I. D. Raji and J. Buolamwini, "Actionable auditing: Investigating the impact of publicly naming biased performance results of commercial ai products," in *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, ser. AIES '19. New York, NY, USA: ACM, 2019, pp. 429–435.
- [125] R. Brandom, "Self-driving cars are headed toward an ai roadblock," 2018. [Online]. Available: <https://www.theverge.com/2018/7/3/17530232/self-driving-ai-winter-full-autonomy-waymo-tesla-uber>

- [126] A. Cuthbertson, "Self-driving cars more likely to drive into black people, study claims," 2019. [Online]. Available: <https://www.independent.co.uk/life-style/gadgets-and-tech/news/self-driving-car-crash-racial-bias-black-people-study-a8810031.html>
- [127] M. K. Scheuerman, J. M. Paul, and J. R. Brubaker, "How computers see gender: An evaluation of gender classification in commercial facial analysis services," *Proc. ACM Hum.-Comput. Interact.*, vol. 3, no. CSCW, Nov. 2019.
- [128] L. M. Brown and C. Gilligan, "Meeting at the crossroads: Women's psychology and girls' development," *Feminism & Psychology*, vol. 3, no. 1, pp. 11–35, 1993.
- [129] L. Meegahapola, S. Ruiz-Correa, and D. Gatica-Perez, "Protecting mobile food diaries from getting too personal," in *Proceedings of the 19th International Conference on Mobile and Ubiquitous Multimedia*, ser. MUM '20. New York, NY, USA: Association for Computing Machinery, 2020.
- [130] D. Cirillo, S. Catuara-Solarz, C. Morey, E. Guney, L. Subirats, S. Mellino, A. Gigante, A. Valencia, M. J. Rementeria, A. S. Chadha, and N. Mavridis, "Sex and gender differences and biases in artificial intelligence for biomedicine and healthcare," *npj Digital Medicine*, vol. 3, no. 1, p. 81, Jun 2020.
- [131] L. Meegahapola, S. Ruiz-Correa, and D. Gatica-Perez, "Alone or with others? understanding eating episodes of college students with mobile sensing," in *Proceedings of the 19th International Conference on Mobile and Ubiquitous Multimedia*, ser. MUM '20. New York, NY, USA: Association for Computing Machinery, 2020.
- [132] V. Mallawaarachchi, L. Meegahapola, R. Madhushanka, E. Heshan, D. Meedeniya, and S. Jayarathna, "Change detection and notification of web pages: A survey," *ACM Comput. Surv.*, vol. 53, no. 1, Feb. 2020.
- [133] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "Smote: Synthetic minority over-sampling technique," *Journal of Artificial Intelligence Research*, vol. 16, p. 321–357, Jun 2002.
- [134] (2019) How to calibrate your smartphone's sensors for increased accuracy. [Online]. Available: <https://www.nextpit.com/how-to-calibrate-sensors-on-android-smartphones>
- [135] W. Yao, Y. Liu, D. Zhou, Z. Pan, M. Till, J. Zhao, L. Zhu, L. Zhan, Q. Tang, and Y. Liu, "Impact of gps signal loss and its mitigation in power system synchronized measurement devices," 07 2017, pp. 1–1.
- [136] T. Gong, Y. Kim, J. Shin, and S.-J. Lee, "Metasense: Few-shot adaptation to untrained conditions in deep mobile sensing," in *Proceedings of the 17th Conference on Embedded Networked Sensor Systems*, ser. SenSys '19. New York, NY, USA: Association for Computing Machinery, 2019, p. 110–123.
- [137] L. K. Berger, A. L. Begun, and L. L. Otto-Salaj, "Participant recruitment in intervention research: scientific integrity and cost-effective strategies," *International Journal of Social Research Methodology*, vol. 12, no. 1, pp. 79–92, 2009.
- [138] M. X. Patel, V. Doku, and L. Tennakoon, "Challenges in recruitment of research participants," *Advances in Psychiatric Treatment*, vol. 9, no. 3, p. 229–238, 2003.
- [139] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 10, pp. 1345–1359, 2010.
- [140] Y. Ma, A. T. Campbell, D. J. Cook, J. Lach, S. N. Patel, T. Ploetz, M. Sarrafzadeh, D. Spruijt-Metz, and H. Ghasemzadeh, "Transfer learning for activity recognition in mobile health," 2020.
- [141] N. Hermandex-Cruz, J. Lundstram, J. Favela, I. McChesney, and B. Arnrich, "Literature review on transfer learning for human activity recognition using mobile and wearable devices with environmental technology," *SN Computer Science*, vol. 1, 03 2020.
- [142] T. Rahman, M. Czerwinski, R. Gilad-Bachrach, and P. Johns, "Predicting "about-to-eat" moments for just-in-time eating intervention," in *Proceedings of the 6th International Conference on Digital Health Conference*, ser. DH '16. New York, NY, USA: Association for Computing Machinery, 2016, p. 141–150.
- [143] E. Thomaz, I. Essa, and G. D. Abowd, "A practical approach for recognizing eating moments with wrist-mounted inertial sensing," in *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, ser. UbiComp '15. New York, NY, USA: Association for Computing Machinery, 2015, p. 1029–1040.
- [144] C. P. Herman and J. Polivy, "Dieting as an exercise in behavioral economics." *Time and decision: Economic and psychological perspectives on intertemporal choice*, 2003.
- [145] C. Herman, J. Polivy, and T. Leone, "6 - the psychology of overeating," in *Food, Diet and Obesity*, ser. Woodhead Publishing Series in Food Science, Technology and Nutrition, D. J. Mela, Ed. Woodhead Publishing, 2005, pp. 115 – 136.
- [146] M. M. Hetherington, A. S. Anderson, G. N. Norton, and L. Newson, "Situational effects on meal intake: A comparison of eating alone and eating with others," *Physiology & Behavior*, vol. 88, no. 4, pp. 498 – 505, 2006.
- [147] (2019) Anger, hostility, and violent behavior. [Online]. Available: <https://www.uofmhealth.org/health-library/anger>
- [148] D. A. Epstein, N. B. Lee, J. H. Kang, E. Agapie, J. Schroeder, L. R. Pina, J. Fogarty, J. A. Kientz, and S. Munson, "Examining menstrual tracking to inform the design of personal informatics tools," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, ser. CHI '17. New York, NY, USA: Association for Computing Machinery, 2017, p. 6876–6888.
- [149] M. Honary, B. Bell, S. Clinch, J. Vega, L. Kroll, A. Sefi, and R. McNaney, "Shaping the design of smartphone-based interventions for self-harm," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, ser. CHI '20. New York, NY, USA: Association for Computing Machinery, 2020, p. 1–14.
- [150] (2019) How to boost confidence as a young adult. [Online]. Available: <https://www.verywellfamily.com/parents-help-confidence-young-adult-women-4092900>
- [151] M. M. Weissman, "Cross-national epidemiology of obsessive-compulsive disorder," *CNS Spectrums*, vol. 3, no. S1, p. 6–9, 1998.
- [152] (2006) Teenagers and pornography addiction: Treating the silent epidemic. [Online]. Available: <https://www.counseling.org/knowledge-center/vistas/by-subject2/vistas-children/docs/default-source/vistas/teenagers-and-pornography-addiction-treating-the-silent-epidemic>
- [153] A. D. F. Lauvsnes, M. Langaas, P. Toussaint, and R. W. GrÅwwe, "Mobile sensing in substance use research: A scoping review," *Telemedicine and e-Health*.
- [154] R. Siegel, "Smoking and depression apps are selling your data to google and facebook, study finds," 2020. [Online]. Available: <https://www.washingtonpost.com/business/2019/04/22/smoking-depression-apps-are-selling-your-data-google-facebook-study-finds/>
- [155] D. Paul, "Colleges want freshmen to use mental health apps. but are they risking students' privacy?" 2020. [Online]. Available: <https://www.washingtonpost.com/technology/2019/12/27/colleges-want-freshmen-use-mental-health-apps-are-they-risking-students-privacy/>
- [156] D. Harwell, "Colleges are turning students' phones into surveillance machines, tracking the locations of hundreds of thousands," 2019. [Online]. Available: <https://www.washingtonpost.com/technology/2019/12/24/colleges-are-turning-students-phones-into-surveillance-machines-tracking-locations-hundreds-thousands/>



LAKMAL MEEGAHAPOLA was born in Kandy, Sri Lanka in 1993. He received his B.Sc. in Computer Science and Engineering from the Faculty of Engineering, University of Moratuwa, Sri Lanka in 2018. Currently, he is a Doctoral Candidate specializing in Mobile Sensing at the School of Engineering, École polytechnique fédérale de Lausanne (EPFL), Switzerland and a Research Assistant in the Social Computing Group at Idiap Research Institute, Switzerland. He is interested

and experienced in using machine learning and data mining techniques for research in the intersection of (a) mobile health sensing, (b) context-awareness, and (c) social computing.



DANIEL GATICA-PEREZ directs the Social Computing Group at Idiap Research Institute and is a Professor at EPFL in Switzerland. His research interests span social computing, ubiquitous computing, and crowdsourcing for social good. He is a member of the IEEE.

...