

## DEPARTMENT: VISUALIZATION VIEWPOINTS

# Visual Analytics of Smartphone-Sensed Human Behavior and Health

Hamid Mansoor , Walter Gerych, Abdulaziz Alajaji , Luke Buquicchio , Kavin Chandrasekaran , Emmanuel Agu , and Elke A. Rundensteiner, Worcester Polytechnic Institute, Worcester, MA, 01609, USA

*Smartphone health sensing tools, which analyze passively gathered human behavior data, can provide clinicians with a longitudinal view of their patients' ailments in natural settings. In this Visualization Viewpoints article, we postulate that interactive visual analytics (IVA) can assist data scientists during the development of such tools by facilitating the discovery and correction of wrong or missing user-provided ground-truth health annotations. IVA can also assist clinicians in making sense of their patients' behaviors by providing additional contextual and semantic information. We review the current state-of-the-art, outline unique challenges, and illustrate our viewpoints using our work as well as those of other researchers. Finally, we articulate open challenges in this exciting and emerging field of research.*

**T**HE CURRENT HEALTHCARE SYSTEM is under-resourced and schedule-driven with patients receiving little care outside of appointments. Consequently, patient assessments are infrequent, typically months apart, and often result in late diagnoses that worsen their prognoses. Emerging research is exploring the use of sensor-rich smartphones that are now owned by over 80% of the U.S. population,<sup>†</sup> to passively detect various ailments, and continuously gather valuable health behavior information and corresponding contexts. Ailments such as depression<sup>5,17</sup> and influenza<sup>9</sup> can be detected early by analyzing sensor data collected from smartphones using machine learning models. This novel paradigm is called *smartphone health sensing or smartphone ailment phenotyping*. Early detection can significantly improve health outcomes.<sup>‡</sup> Passive smartphone phenotyping provides clinicians with an objective, contextualized

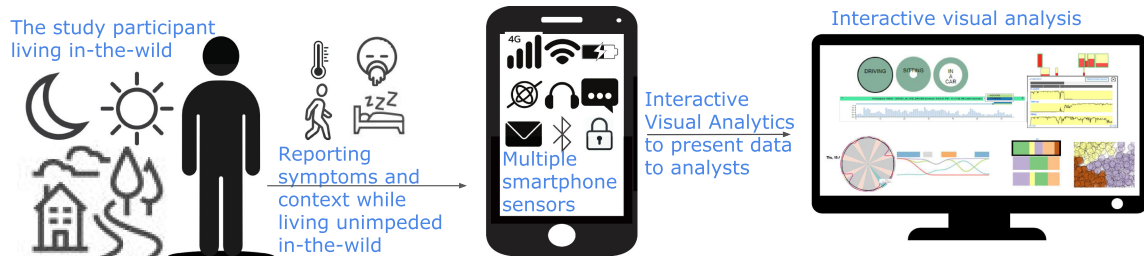
picture of their patients' lives in the real world. Such evidence can then be used to support treatment decisions as patient self-reports may be inaccurate due to recall bias and exaggeration. However, analysis of real world smartphone-sensed health data is challenging.

In addition to traditional issues such as the highly multivariate and complex nature of such spatio-temporal data, smartphone-sensed data analysis faces unique challenges such as the need to disambiguate noisy ground truth labels of health behaviors and context in natural settings. While passive data gathering in natural settings yields realistic data, it also means that users often provide wrong or no labels when they are busy with their lives. Such labeling issues in turn lead to weak supervision for machine learning modeling. Smartphone health inference and phenotyping also faces unique challenges as user behaviors indicative of health status are often intertwined with other unrelated real world activities. Ultimately, the smartphone user's specific situations, contexts, and health status at any point in time are not always clear. Moreover, multiple ailments can have the same smartphone signature or phenotype, leading to confounding effects. Finally, the degree to which users express each symptom of the same underlying ailment vary a lot, making intersubject comparisons challenging.

Prior work on IVA for health related data were typically on structured data from sources such as

<sup>†</sup><https://www.pewresearch.org/internet/fact-sheet/mobile/>

<sup>‡</sup><https://www.webmd.com/depression/guide/untreated-depression-effects>



**FIGURE 1.** During user studies to gather labeled data for developing smartphone health inference models, a user lives in-the-wild while their smartphone passively gathers data continuously. Periodically, the user provides ground truth labels of their context (situation) and health symptoms on their phone. Users often provide wrong or no labels when they become busy with their lives, which presents a challenge for supervised machine learning methods. Visual analytics can assist data scientists in correcting wrong or missing labels, cleaning such data before machine learning modeling, and debugging of such models. For clinicians using the final machine learning-based passive health inference models, IVA provides additional patient context and semantic information on the health symptoms, and comparisons with other patients, which enhances interpretability and trust of models' outputs.

Electronic Health Records (EHRs) with clear definitions of data fields and coding of health related events and relevant patient information.<sup>6,15</sup> IVA systems for sensor-based health data have visualized low level health and wellness variables such as meals, sleep patterns, and step counts, which were gathered using health wearable devices such as Fitbits.<sup>18</sup> We focus on IVA systems for making higher level inferences [e.g., detecting Traumatic Brain Injury (TBI)] passively from these low level health variables and symptoms captured in-the-wild using smartphones (Figure 1).

A prior survey by Polack *et al.* reviewed IVA methods to analyze *Mobile Health (mHealth)* data in general.<sup>15</sup> Topics covered included the visual representation of complex and multivariate temporal data, interactive cohort selection, and trend mining. Our viewpoints are more focused on specific IVA support for passive smartphone health sensing or phenotyping, which typically utilize machine learning. We add to this exciting and emerging field by distilling novel viewpoints, defining key stakeholders, outlining unique challenges arising from this new method of monitoring health, and providing concrete examples to illustrate how IVA methods can provide additional insights and assist specific stakeholders. Our viewpoints focus on the unique challenges that stem from the weakly and sometimes incorrectly supervised nature of *in-the-wild* smartphone health sensing studies and the dynamic, ambiguous, and sometimes confounding behaviors of the monitored user. Aspects our viewpoints cover include support for correcting user-provided annotations, enhancing the understanding, contextualization, and sensemaking of smartphone-sensed health behaviors, and population-level visualizations and subject health status comparisons.

## SMARTPHONE HEALTH SENSING—STAKEHOLDERS AND CHALLENGES

There are generally two groups of stakeholders for smartphone health sensing or phenotyping: 1) *Clinicians and Health Professionals* who seek to use such systems to understand their patients' ailment-related real world behaviors and symptom trajectories in the real world better, and identify concerning behaviors early; 2) *Data Scientists* who develop computational and machine learning models to passively detect ailments using weakly supervised smartphone-sensed data with noisy labels. Their ultimate goal is to deploy those models to continuously assess and monitor the patient and detect ailments early in a completely unsupervised fashion. These two groups of stakeholders face different but related challenges, which we now summarize.

### Health Behavior Understanding Challenges

While inferences about health and contexts made using machine learning on smartphone data can be accurate,<sup>16,17</sup> they are typically not *explainable* nor do they incorporate expert knowledge. For example, while machine learning can detect sleep duration and quality accurately using smartphone data,<sup>1</sup> it does not capture a comprehensive picture of potential causes of sleep disruptions which are important for healthcare professionals. Such disruptions may be *explainable* if additional *human understandable* information and the occurrence of comorbidities such as participants' increased reported stress were *visually* linked to such

inferences. Clinicians who use smartphone health sensing systems to monitor their patients as well as data scientists who develop them both face health behavior understanding challenges.

### Symptom and Health Behavior Contextualization Challenges

Human life involves multiple, intertwined experiences. Clinician end users as well as data scientist model developers typically would like temporal information regarding their subjects' situation and trajectory leading up to a smartphone-sensed assessment, concurrently and afterward in order to fully contextualize it. Health analysts can utilize *visual* methods to contextualize such information<sup>15</sup> to disambiguate confounding scenarios and improve the specificity of diagnoses. For instance, while reductions in a smartphone user's step count may be caused by depression, it may also be a short-term reduction caused by fatigue because the user engaged in strenuous exercise the previous day.

### Smartphone Data Labeling Challenges

Data scientists creating smartphone health assessment models using machine learning require labeled, *real world* datasets. Typically, an app installed on users' smartphones, continuously gathers sensor data as they live their lives. To provide ground truth labels, users periodically respond to questions to report their health condition<sup>17</sup> symptoms, as well as corresponding contexts visited, activities performed, and social situations experienced.<sup>16</sup> Label data collection studies can be disruptive leading to two data science issues: 1) *Missing Labels*: participants fail to provide health or context labels when they are busy or distracted. Participant response rates to ground truth questions also vary leading to imbalanced datasets, and 2) *Wrong Labels*: participants make human errors in providing labels due to carelessness or recall bias.<sup>12</sup>

## OUR POSTULATIONS ON HOW IVA CAN ENHANCE SMARTPHONE HEALTH SENSING

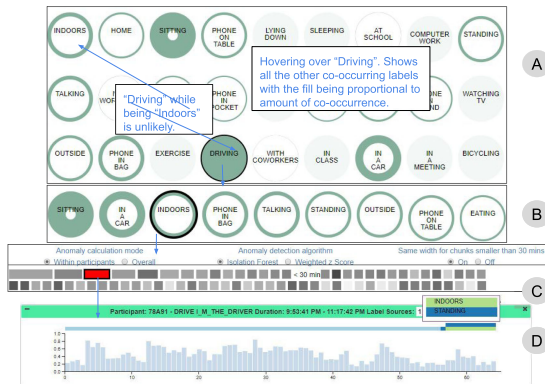
Interactive data visualizations are useful for analyzing multivariate data.<sup>7,14</sup> Polack *et al.*<sup>15</sup> previously highlighted some research directions for IVA for mHealth data broadly including visualizing its multi-scale, temporal nature. We build on and extend this work by summarizing specific *Viewpoints* on how IVA can be useful for the different stakeholders of

smartphone health sensing and phenotyping, and present concrete, illustrative examples.

### IVA Support for Data Scientists

*V1: Visual support for detecting mislabeled data* to improve the model development. Traditionally, IVA works for data cleaning in complex domains such as multimedia, trajectory, and textual data,<sup>8</sup> rely on metrics such as anomaly scores to alert users of poor quality data. In-the-wild gathered smartphone-sensed data is more multicontextual in nature, which makes it difficult to rely solely on such computational methods to discover poorly labeled data. IVA is well suited to present the complex characteristics of smartphone-sensed health data<sup>15</sup> and can leverage multiple visual metaphors to display automatically derived metrics and other intuitive cues for anomalous data indicative of mislabeling effectively. For instance, activities labeled as occurring simultaneously but which are unlikely to truly be co-occurring (e.g., standing while driving) can be visually highlighted to make them easy to discover.

Context Mislabeled EXplorer (COMEX)<sup>11</sup> is an IVA framework that facilitates the discovery of mislabeled smartphone sensed data by incorporating visuals that provide additional context. COMEX analyzes continuously gathered, in-the-wild smartphone data<sup>16</sup> with participant-provided ground truth health and context labels. COMEX combines the visualization of computed label anomaly scores with visual metaphors designed to address the multicontextual and continuous nature of real world, labeled smartphone sensor data. For instance, COMEX highlights unlikely co-occurrence of activities as a clue for discovering wrong labels [e.g., "Driving" while "Indoors" Figure 2(A) and (B)], alerting analysts about potential mislabels. Unlikely context and activity durations can also provide another clue for wrong user labels [e.g., "Walking" for 15 h, Figure 2(C)]. However, users tend to label data at frequencies that vary over time, providing more labels when they are free and less when busy. Consequently, detecting the start-end and continuation of contexts over time can be confusing. COMEX deals with temporal variations in labeling frequency by presenting visual indicators ("chunks") of contextual continuity to make them easy to identify continuing contexts, compare user-reported context durations, and identify likely wrong reports. The intuitive use of simple metaphors to highlight suspicious context co-occurrence and duration demonstrate that IVA is well-suited to assist in label correction of smartphone sensed data.

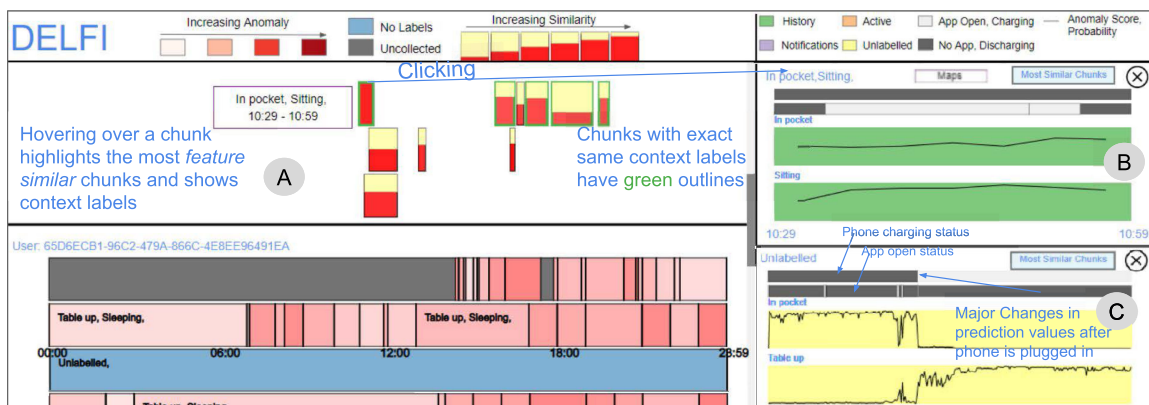


**FIGURE 2.** COMEX. (A) Showing labels across participants, ordered by occurrence. Hovering over a circle shows its most commonly co-occurring labels with the fill being proportional to co-occurrence. (B) Clicking on a circle shows Chunks of the selected label with length encoding time. The chunks are ordered by duration and their opacity encodes their average anomaly score. Clicking on a chunk shows its details as a histogram. The bars above the histogram are the co-occurring labels for the sessions.

V2: Visual support for labeling unlabeled data to improve the model development. In order to discern the likely labels of unlabeled data, some IVA methods integrate contextual visual cues of multivariate data to enable analysts to assign labels to data more confidently.<sup>3</sup> However, such approaches require data with verifiable ground truth labels, which is not the case in

in-the-wild smartphone data, where the user may not have provided labels and exact ground truth labels are not known. In cases where smartphone data are partially labeled, a semisupervised visual paradigm can be utilized, wherein the probable labels of unlabeled data are discerned from their visual similarity to labeled data. To make labeling easy, IVA tools can provide contextual details from passively sensed data and highlight similarities between labeled and unlabeled data. For instance, time periods during which the smartphone logs show calls have been received as well as high noise levels can be labeled as “conversation” based on similarity between features corresponding to this time period and other instances of data labeled as “conversation.”

Detecting Erroneous Labels using Feature-linking Insights (DELFI) (see Figure 3)<sup>12</sup> is an IVA framework to highlight unlabeled data, suggest similarities to labeled data instances in terms of *sensor feature values*, and facilitate the assignment of labels with confidence. DELFI utilized a Multi-Feature Similarity Linking paradigm (inspired by Nguyen *et al.*<sup>14</sup>) to visually link feature-similar data with an overlay of contextual information, enhancing intuition. Interacting with continuous “chunks” of labeled or unlabeled data shows the most feature-similar chunks (based on Euclidean distance between feature values) for visual linking [see Figure 3(A)], to discover potentially mislabeled data and assign labels to unlabeled data. The values of soft sensors such as the apps running, call, and SMS logs, and charging status add another layer of



**FIGURE 3.** DELFI: (A) Habit View : Shows every user’s participation days as Continuous Context Chunks. Hovering over a chunk hides all others except those that are most feature similar. (B) Chunk Detail View : Shows chunk details like the labeling mechanism (these labels were provided using “History”). The labels making up the context are split as bars and the lines show the respective anomaly scores or probability values (for unlabeled data) for the comprising data. The two gray bars represent the charging status (first bar) and app usage status (second bar) for the data sessions. (C) Unlabeled chunk. The labels with the highest average probability values are shown and the lines in the bars represent probability values for individual sessions.



explainability. For instance, around the time the phone was plugged in, the unlabeled chunk in Figure 3(C) had an increased likelihood of having the label “Table up” (phone position) and a decline for the label “In pocket.” Visual overlays increase the analyst’s confidence by presenting intuitive contextual information to enhance interpretation. Such intuition is difficult to generate with nonvisual and purely computational methods.

V3: *Visual clustering of smartphone-sensed data* to improve scalability and steer the development of machine learning classifiers for health applications. Groups of data similar in terms of sensor feature values can be visualized with overlays of analyst-understandable semantic information that can help in building classification models. For instance, grouping days with poor sleep reports can help identify user phenotypes and patterns in smartphone-sensed data features that can be predictive of them. In addition, as the scale of such studies grows, *Population level* analysis becomes necessary for understanding differences between groups and subpopulations of people. IVA is well-suited for the task of cohort selection and comparison as noted by Polack *et al.*<sup>15</sup> For instance sleep problems in late shift workers versus early shift workers can be contextualized based on the distribution of sensor-detected sleep hours, and can help analysts in assigning chronotypes (groups with similar sleep-wake cycles)<sup>1</sup> (e.g., morning person versus night owl) to participants. Presenting inferred as well as reported symptom data with contextual information such as phone interactions, mobility patterns, and days of the week can enable analysts to understand the differences in sensed data between participants and allow for larger scale, longer term analyses.

Prior work enabled unsupervised analysis of multi-feature data by presenting results from multiple clustering and dimension reduction algorithms.<sup>7</sup> We are researching and developing INTERactive Observation of Smartphone Inferred Symptoms (INTOSIS), an IVA tool which adopted a similar approach for a large-scale smartphone-sensed dataset in an ongoing study. The data include sensors such as anonymized geolocation and activity levels along with sparse daily and weekly symptom labels. A ranked list of clustering results (based on quality of clustering) was generated using various algorithms such as *K*-Means and spectral distancing [see Figure 6(E)] from clustering all days across users based on sensor features. The clusters are then visualized across a plane using *t*-distributed stochastic neighbor embedding (*t*-SNE), a visual dimension reduction technique. This enables analysts to see similarity of features between symptomatic

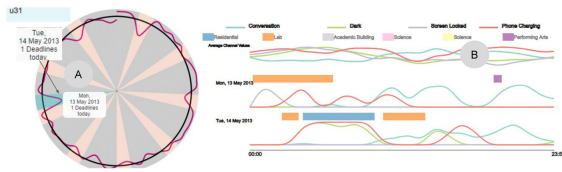
days and provide explanations for the distributions of clusters. In addition, this shows the sensor-detected factors that may be indicative of symptoms. For example, a cluster with several days labeled as “Poor sleep” may be correlated with less time at the participant’s primary location at night.

The “Feature Averages Heatmap” [see Figure 6(D)] shows the feature value distribution, ordered by their importance (ANOVA *F*-statistic). This shows the defining characteristics of clusters [e.g., purple cluster shows days with more than usual time spent at home, Figure 6(D)] and assign semantic meaning to objective sensor data. IVA enables easier interpretation of such unsupervised data in a way that lets analysts make such important associations.

## IVA Support for Health Professionals

V4: *Visualizing anomalous, passively sensed unhealthy patterns of user behaviors.* Deviations from routines can provide health analysts with clues about causes of symptoms/concerning behaviors. Human Bio behavioral Rhythms (HBR) such as sleep-wake cycles or circadian rhythms and their disruptions have health ramifications and are detectable from smartphone data.<sup>1</sup> While such computational approaches are accurate, they usually provide little *explainability*, which is an issue as the scope of such studies grows and as users provide more sparse labels with increased study durations. Alternate nonvisual analysis methods exploit the multiscale temporality of behavioral rhythms derived from smartphone-sensed data. IVA techniques can combine multiple views that facilitate contextualization of multiscale temporal data for trend and pattern mining, which can reveal important, health-relevant information<sup>15</sup> such as sleep patterns and stress levels during hectic times such as weekdays versus more relaxed times such as weekends, etc.

ARGUS<sup>10</sup> is an IVA framework that displays smartphone-sensed HBRs and disruptions in them, using multiple visual concepts to assist analysts in not only identifying but also explaining HBR disruptions. ARGUS utilizes a novel Rhythm Deviation Score (RDS) that quantifies the degree of periodicity of the underlying sleep wake-cycle based on sensor data. Each participant day is assigned an RDS score, which can then be visualized effectively in conjunction with other contextual information. ARGUS uses a glyph based on the Z-glyph,<sup>4</sup> a visual metaphor to present disruptions from the norm. The black circle [see Figure 4(A)] represents the overall rhythm (larger circle means more rhythmic) and the dips in the purple line represent days with disruptions. Bigger dips toward the center

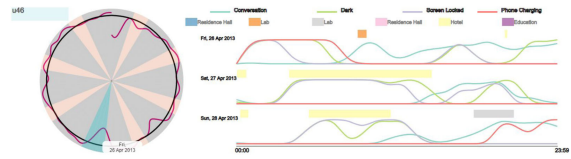


**FIGURE 4.** (A) Larger black circle means more rhythm. The purple line shows daily disruption in rhythm (closer to center means higher disruption). Every slice is a day, with the beige slices being weekends. Days can be selected for analysis in (B). (B) displays the daily distribution across sensor values (lines) along with durations spent in geoclusters (colored bars over lines are place types).

indicate bigger HBR disruptions. Contextual data about possible causes of HBR disruptions such as academic and project deadlines and the places visited are overlaid on the glyph. Presenting additional contextual information such as the type of places visited enables analysts to assign contextual meaning to the sensed HBR (see Figures 4 and 5).

For instance, a university student's<sup>17</sup> smartphone-sensed rhythm was disrupted for two days by deadlines [see Figure 4(A)]. Exploration shows that they were in a "Lab" during the early hours of the 2 days [see Figure 4(B)]. Such detailed visual analysis not only shows concerning data, but also helps to *explain* and *contextualize* it. Nonvisual methods are limited in their ability to make use such connections in order to explain the degree of HBR periodicity.

V5: *Visually overlay health markers and symptom reports to assign semantic values to objective sensor data* for health analysis. Smartphone-sensed data are often anonymized, with loss of information that can potentially explain symptoms. For instance, not knowing semantic information about a participant's work-life routine may make it harder to explain depressive symptoms since staying longer at work has been linked with depression.<sup>2, 13</sup> Consequently, the ability to *semantically* label unlabeled/anonymized smartphone-sensed data can be valuable (e.g., labeling the place where a participant "Stays" most as their home<sup>5</sup>). IVA can provide views to show temporal trends in smartphone-sensed data<sup>15</sup> that can enable semantic labeling of objective sensor data. INTO-SIS utilizes the same large-scale dataset from the ongoing study previously mentioned, to visualize day-level mobility features, based on location data per day [daily values shown for a user in Figure 6(A) and (C)]. Figure 6 (B) shows a list of the geoclusters (clustered using DBSCAN), sorted by the duration of participants' "Stays" in them. For each geocluster, the average 24-h



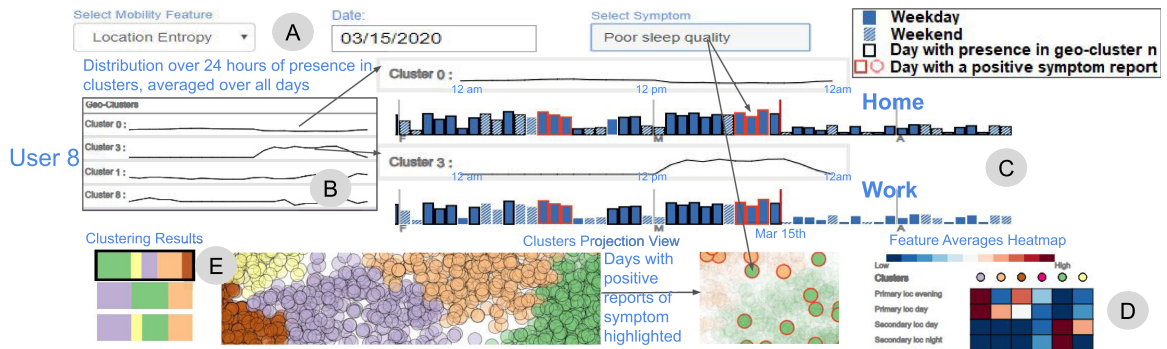
**FIGURE 5.** Contextualizing a rhythm break by showing human understandable information. For instance there was a large break in rhythm detected over a weekend where we see the participant in a "Hotel."

distribution of presence is shown (flatter lines indicate uniform presence). Days with positive symptom responses can be highlighted [see Figure 6(A)] for drill-down. For instance, the user was in geocluster 0 approximately every day [see Figure 6(B)], with uniform distribution, suggesting this is a residence. They are in geocluster 3, the second longest amount of time. This geocluster is likely their workplace, as they were never in that geocluster before 12 pm nor on weekends and also not since mid-March, when social distancing went into effect in the U.S. due to COVID-19 and workplaces were closed. Also, after mid-March, there was a decrease in location entropy, which measures how much the user visited popular locations. The ability to flexibly define and assign *semantic labels* can help analysts assign meaningful/ predictive labels to unlabeled data and improve inference.

## CALL TO ACTION: OPEN RESEARCH CHALLENGES

While significant progress has been made toward realizing the vision of using IVA to enhance smartphone-sensed health assessment tools, several open challenges still need to be solved. As the number of studies grow, data scientists and health experts will need to address challenges associated with increases in both the duration of studies and the number of users monitored including

- ▶ *Scalable visualizations:* As the scope of in-the-wild studies broadens and larger groups of smartphone users are analyzed, visual scalability must be considered. Specifically, the ability to visualize larger numbers of users with more data and over longer periods of time becomes important. The wealth of data available on the current COVID-19 pandemic is a case in point. While this may be tackled by using longer time windows (weeks instead of days) for analysis, this could inversely affect the quality or precision of the inferred results as ailments may not manifest



**FIGURE 6.** INTOSIS: Interacting with a geocluster distribution (B) shows the days where the user was in that cluster for some time with a black stroke. Flatter lines mean even presence for all hours of the day. Specifying a symptom (A) shows days with positive instances with a red stroke. Based on sensor features, data were analyzed using multiple clustering and dimension reduction techniques (E) to project each day (circle = day) of data across all participants on a 2-D plane. The clustering results are ordered by quality. Average feature value distribution across clusters is shown as a heatmap. Red and blue represent high and low values, respectively (D).

clearly on larger time scales. The visualization techniques themselves have to be scalable and interactive. While some of our work has begun to address this challenge, more research is required.<sup>15</sup> Visual clustering can also be used along with multiscale visual interactive techniques such as zooming, filtering, and details on demand which can facilitate drill-downs, high-level meta views, and inter and intragroup analysis as the number of participants grows.

- ▶ *Visualizing causative presymptom patterns:* To facilitate retrospective analysis and the discovery of behaviors that caused illness, it may be useful to support visual lookbacks on data and highlight users' behavior patterns that occurred frequently leading up to specific symptom reports. These can serve as clues for potential ailment causes. For instance, it may be informative to display a pattern of higher than usual activity levels on days preceding the time when a subject reported fatigue.
- ▶ *Visual encodings that align with experts training and build on domain knowledge.* Many users of IVA tools for smartphone sensing such as psychiatrists, doctors, and nurses will already have substantial prior domain training that they could bring to bear in interpreting various visual clues. Working with experts to develop visual encodings that build on their domain knowledge is important. For instance, psychologists and health experts are trained that a patient's behavior differs on weekdays versus weekends. In our work with them, they liked encodings of user

behaviors on weekdays versus weekends along with specific smartphone sensed patterns such as mobility or app usage during day versus night. Such visuals will enable the experts to apply their domain expertise to the analysis task. Therefore, any data visualizations for smartphone data should explicitly encode contextual information accepted by the health domain.

- ▶ *Multifacet aggregation of contextual variables.* A smartphone user's context has multiple facets including their activity, app being used and the type of place they are at. Displaying these disparate bits of information as separate visual streams across long periods of time can quickly become overwhelming. Future work may consider multicontext visual metaphors that represent analyst-specified aggregations of various sensor channels or those accepted by the health domain in a human-understandable fashion. For instance, aggregating low user activity, and environmental light and sounds levels as well as information that the user stayed at the same geolocation at nights to be visually displayed as "sleeping at night." Once labeled, subsequently collected data can then be automatically labeled with the "sleeping at night" label as appropriate.
- ▶ *User-friendly visual metaphors for variable visualization literacy* among health experts and general public. Smartphone-sensed health detection is a multidisciplinary topic, including researchers with varying levels of visualization, computer, and data science literacy. Future

work can develop universally comprehensible visual metaphors to lower the learning curve for researchers across the spectrum of expertise in the health domain, who use IVA frameworks. Additionally, smartphone-sensed studies may also benefit from deploying on-device, user-friendly visualizations, and *gamification* strategies to increase participant compliance such as encouraging accurate and frequent labeling. Visualizing smartphone-sensed health data on a personal level may also mitigate the privacy issues that are inherent in most smartphone data gathering projects by giving participants access to their own data. Likewise, health recommendations generated by automated models can be provided to the user without analyst intervention.

## CONCLUSION

Smartphones-sensed human behavior and health data are rapidly increasing in both quantity and complexity. Smartphone health sensing and phenotyping analyses try to utilize objective data to passively assess the health of the smartphone user and derive meaningful insights. However, labeling issues and the complex nature of smartphone-sensed, real world data present challenges to analysts who utilize predominately computational approaches. Interactive data visualizations are a viable alternative that can assist health professionals in such phenotyping analysis and improve the understanding and contextualization of health behaviors. IVA can also provide data scientists with valuable tools for mitigating data labeling issues and debugging machine learning health inference models during model development. In this visualization viewpoints article, we posited that interactive data visualization is an exciting approach to empower health scientists to discover smartphone-sensed human behaviors and symptoms that are predictive of health ailments and presented illustrative examples.

## ACKNOWLEDGMENT

This material is based on research sponsored by DARPA under agreement number FA8750-18-2-0077. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as

necessarily representing the official policies or endorsements, either expressed or implied, of DARPA or the U.S. Government.

## REFERENCES

1. S. Abdullah, E. L. Murnane, M. Matthews, and T. Choudhury, "Circadian computing: Sensing, modeling, and maintaining biological rhythms," in *Mobile Health*. New York, NY, USA: Springer, pp. 35–58, 2017.
2. P. Afonso, M. Fonseca, and J. F. Pires, "Impact of working hours on sleep and mental health," *Occup. Med.*, vol. 67, no. 5, pp. 377–382, 2017.
3. B. Jurgen *et al.*, "VIAL: A unified process for visual interactive labeling," *Vis. Comput.*, vol. 34, no. 9, pp. 1189–1207, 2018.
4. N. Cao, Y.-R. Lin, D. Gotz, and F. Du, "Z-glyph: Visualizing outliers in multivariate data," *Inf. Vis.*, vol. 17, no. 1, pp. 22–40, 2018, doi: [10.1177/1473871616686635](https://doi.org/10.1177/1473871616686635).
5. W. Gerych, E. Agu, and E. Rundensteiner, "Classifying depression in imbalanced datasets using an autoencoder-based anomaly detection approach," in *Proc. IEEE 13th Int. Conf. Semantic Comput.*, 2019, pp. 124–127.
6. Z. Jin *et al.*, "Carepre: An intelligent clinical decision assistance system," *ACM Trans. Comput. Healthcare*, vol. 1, no. 1, pp. 1–20, 2020.
7. B. C. Kwon *et al.*, "Clustervision: Visual supervision of unsupervised clustering," *IEEE Trans. Vis. Comput. Graph.*, vol. 24, no. 1, pp. 142–151, Jan. 2018.
8. L. Shixia *et al.*, "Steering data quality with visual analytics: The complexity challenge," *Vis. Inform.*, vol. 2, no. 4, pp. 191–197, 2018.
9. A. Madan, M. Cebrian, S. Moturu, K. Farrahi, and A., "Sensing the" health state" of a community," *IEEE Pervasive Comput.*, vol. 11, no. 4, pp. 36–45, Oct./Dec. 2011.
10. H. Mansoor *et al.*, "AR-GUS: Interactive visual analytics framework for the discovery of disruptions in bio-behavioral rhythms," in A. Kerren, C. Garth, and G. E. Marai, eds., in *EuroVis 2020 - Short Papers*. Genève, Switzerland: The Eurographics Assoc., 2020, doi: [10.2312/evs.20201043](https://doi.org/10.2312/evs.20201043).
11. H. Mansoor, W. Gerych, L. Buquicchio, K. Chandrasekaran, E. Agu, and E. Rundensteiner, "Comex: Identifying mislabeled human behavioral context data using visual analytics," in *Proc. IEEE 43rd Annu. Comput. Softw. Appl. Conf.*, vol. 2, pp. 233–238, 2019, doi: [10.1109/COMPSAC.2019.10212](https://doi.org/10.1109/COMPSAC.2019.10212).



12. H. Mansoor, W. Gerych, L. Buquicchio, K. Chandrasekaran, E. Agu, and E. Rundensteiner, "DELFI: Mislabeled human context detection using multi-feature similarity linking," in *Proc. IEEE Vis. Data Sci.*, 2019, pp. 11–19, doi: [10.1109/VDS48975.2019.8973382](https://doi.org/10.1109/VDS48975.2019.8973382).
13. Martin, M., R. Weibel, C. Röcke, and S. M. Boker, "Semantic activity analytics for healthy aging: Challenges and opportunities," *IEEE Pervasive Comput.*, vol. 17, no. 3, pp. 73–77, Jul./Sep 2018.
14. P. H. Nguyen, C. Turkey, G. Andrienko, N. Andrienko, O. Thonnard, and J. Zouaoui, "Understanding user behaviour through action sequences: From the usual to the unusual," *IEEE Trans. Vis. Comput. Graph.*, vol. 25, no. 9, pp. 2838–2852, Sep. 2019.
15. P. Polack, M. Sharmin, K. de Barbaro, M. Kahng, S. Chen, and D. Chau, "Exploratory visual analytics of mobile health data: Sensemaking challenges and opportunities," in *Mobile Health*. Cham, Switzerland: Springer, 2017, pp. 349–360.
16. Y. Vaizman, K. Ellis, G. Lanckriet, and N. Weibel, "Ex- trasensory app: Data collection in-the-wild with rich user interface to self-report behavior," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, 2018, pp. 1–12.
17. R. Wang *et al.*, "Studentlife: Assessing mental health, academic performance and behavioral trends of college students using smartphones," in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput.*, 2014, pp. 3–14.
18. Z. Liang *et al.*, "SleepExplorer: A visualization tool to make sense of correlations between personal sleep data and contextual factors," *Pers. Ubiquitous Comput.*, vol. 20, no. 6, 2016, pp. 985–1000.

**HAMID MANSOOR** is currently working toward the Ph.D. degree in computer science with Worcester Polytechnic Institute, Worcester, MA, USA. His research interests include building interactive data visualizations to analyze multivariate data. He is the corresponding author of this article. Contact him at [hmansoor@wpi.edu](mailto:hmansoor@wpi.edu).

**WALTER GERYCH** is currently working toward the Ph.D. degree in data science with Worcester Polytechnic Institute,

Worcester, MA, USA. His research interests include developing machine learning algorithms to learn from incompletely labeled data. Contact him at [wgerych@wpi.edu](mailto:wgerych@wpi.edu).

**ABDULAZIZ ALAJAJI** is currently working toward the Ph.D. degree in data science with Worcester Polytechnic Institute, Worcester, MA, USA. His research interests include developing machine learning algorithms for mobile health. Contact him at [asalajaji@wpi.edu](mailto:asalajaji@wpi.edu).

**LUKE BUQUICCHIO** is currently working toward the Ph.D. degree in data science with Worcester Polytechnic Institute, Worcester, MA, USA. His research interests include developing machine learning algorithms to understand the emergence of new, unknown classes. Contact him at [lbuquicchio@wpi.edu](mailto:lbuquicchio@wpi.edu).

**KAVIN CHANDRASEKARAN** is currently working toward the Ph.D. degree in data science with Worcester Polytechnic Institute, Worcester, MA, USA. His research interests include developing machine learning algorithms for recognizing complex human activities. Contact him at [kchandrasekaran@wpi.edu](mailto:kchandrasekaran@wpi.edu).

**EMMANUEL AGU** is currently a Professor in the Computer Science Department, Worcester Polytechnic Institute, Worcester, MA, USA. His research interests include computer graphics, mobile computing, image analysis, and machine learning especially applications in healthcare. Contact him at [emmanuel@wpi.edu](mailto:emmanuel@wpi.edu).

**ELKE A. RUNDENSTEINER** is currently a Professor of computer science and the founding Director of the interdisciplinary Data Science program at Worcester Polytechnic Institute, Worcester, MA, USA. As an internationally recognized expert in big data analytics, her research spans data science, data stream analytics, machine learning, visual and computational big data infrastructures, and digital health. Contact her at [rundenst@wpi.edu](mailto:rundenst@wpi.edu).

Contact department editor Theresa-Marie Rhyne at [theresamarierhyne@gmail.com](mailto:theresamarierhyne@gmail.com).